

## Research Article

# Identification of Wheat Inflorescence Development-Related Genes Using a Comparative Transcriptomics Approach

Lingjie Ma, Sheng-Wei Ma, Qingyan Deng, Yang Yuan, Zhaoyan Wei, Haiyan Jia , and Zhengqiang Ma

The Applied Plant Genomics Laboratory of Crop Genomics and Bioinformatics Centre, Nanjing Agricultural University, Jiangsu 210095, China

Correspondence should be addressed to Haiyan Jia; [hyjia@njau.edu.cn](mailto:hyjia@njau.edu.cn)

Received 3 August 2017; Revised 26 November 2017; Accepted 3 December 2017; Published 8 February 2018

Academic Editor: Marco Gerdol

Copyright © 2018 Lingjie Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Inflorescence represents the highly specialized plant tissue producing the grains. Although key genes regulating flower initiation and development are conserved, the mechanism regulating fertility is still not well explained. To identify genes and gene network underlying inflorescence morphology and fertility of bread wheat, expressed sequence tags (ESTs) from different tissues were analyzed using a comparative transcriptomics approach. Based on statistical comparison of EST frequencies of individual genes in EST pools representing different tissues and verification with RT-PCR and RNA-seq data, 170 genes of 59 gene sets predominantly expressed in the inflorescence were obtained. Nearly one-third of the gene sets displayed differentiated expression profiles in terms of their subgenome orthologs. The identified genes, most of which were predominantly expressed in anthers, encode proteins involved in wheat floral identity determination, anther and pollen development, pollen-pistil interaction, and others. Particularly, 25 annotated gene sets are associated with pollen wall formation, of which 18 encode enzymes or proteins participating in lipid metabolic pathway, including fatty acid  $\omega$ -hydroxylation, alkane and fatty alcohol biosynthesis, and glycerophospholipid metabolism. We showed that the comparative transcriptomics approach was effective in identifying genes for reproductive development and found that lipid metabolism was particularly active in wheat anthers.

## 1. Introduction

Inflorescences are the reproductive architectures of plants, composed of stems, stalks, bracts, and flowers. *Poaceae* (also called *Gramineae*) is one of the largest families in the monocotyledonous flowering plants, including the major cereal crops, such as maize (*Zea mays* L.), rice (*Oryza sativa* L.), and wheat (*Triticum aestivum* L.). Inflorescences of this family are characterized by their panicle or spike shapes [1], complex branches, and unique spikelets, as well as inconspicuous and anemophilous flowers without obvious petals and sepals [2, 3]. Species in the genus *Triticum* takes the shape of the spike with the spikelets, each containing one or more florets, attached to rachis nodes. The wheat floret consists of a carpel with its ovary, style, and stigma, with three anthers attached to the base through slender filaments, which are enclosed by bract-like organs called lemma and palea.

Inflorescence development and regulation have attracted great attention from plant biologists and crop breeders since they are crucial for reproduction of flowering plants and for production of food grains in cereal crops. Although the key genes regulating the flower initiation and development are conserved in higher plants [4, 5], the diversity of reproductive structure and behaviors is still not well explained. Transcriptomes reflect the complete set of RNA transcripts expressed by the genome under developmentally or physiologically distinct states; therefore, its comparison allows identification of genes under common regulation. High-throughput methods, such as serial analysis of gene expression (SAGE) [6], microarray technology [7], and next-generation sequencing [8] have enabled transcriptome studies in an unprecedented scale in many plant species, especially in those with full genome sequence information available. This has led to the discovery of a number of

genes involved in flower development. These genes display tempo-spatial expression patterns not only in transcriptome level [9–12] but also at the proteome level, such as those in pollen development of tomatoes [13], indicating their strictly regulated functions. Anther tissues were used in most of these studies because they are easy to isolate and specific in biological roles. The availability of these data has shed some light on the gene networks that contribute to the formation of unique reproductive structures and flower development, although it might still not be enough because of the particularity of different plant species and highly specific flower components.

In bread wheat (*T. aestivum* L.), an allohexaploid species ( $2n=42$ ) with three closely related subgenomes, that is, A, B, and D [14], large-scale transcriptomic investigations have been conducted in some tissue types. Crismani et al. performed microarray-based expression analysis of anthers across various stages of meiosis in an attempt to build the link between the model and wheat and identify early meiotic genes [15]. McIntosh et al. investigated the transcriptome in wheat caryopsis development using data generated by SAGE [16]. Yang et al. compared the RNA-seq tags of pistillody stamen and pistil from a pistillody wheat mutant and stamen from the wild-type control to identify differentially expressed genes [17]. In these studies, the annotation of the identified transcripts was based on short tags or sequence reads which is impossible to ensure accuracy for species such as wheat that has a complex genome and does not have detailed sequence information so far. Moreover, gene identity determination according to short sequence fragments could be problematic due to functional diversification of the homoeologous genes in polyploidy species.

Expressed sequence tags (ESTs) are single-pass sequence reads by sequencing cDNA libraries and usually have >400-base read length. They represent part of the transcriptome in a given tissue and/or at a given developmental stage. As of January 1, 2013, over one million ESTs are available for wheat in GenBank of the National Center for Biotechnology Information (NCBI) and are valuable for identifying genes involved in biotic and abiotic stress [18–22], kernel development and quality [23, 24], and development [21, 25, 26].

In this paper, we reported mining of 170 genes predominantly expressed in floral organs through comparative transcriptomic profiling using ESTs of 67 wheat cDNA libraries deposited in GenBank genes and the identification of a few metabolic pathways involved in anther development.

## 2. Materials and Methods

**2.1. Wheat ESTs and Contigs.** ESTs were downloaded from the NCBI dbEST database (<ftp://ncbi.nlm.nih.gov/repository/dbEST>). The ESTs used in this study were produced with 67 cDNA libraries prepared using inflorescences (spike at flowering or before flowering, anther, pistil, ovary, palea, and lemma), roots, stems, leaves (including seedlings and crown tissues), and seeds (matured or immature embryos) from normally grown seedlings or plants and represented

434,658 cloning events (Table S1). Libraries subjected to enrichment or normalization treatment and those with less than 1000 ESTs were not included in the expression profiling analysis. The 163rd release of unique wheat transcripts, including 77,657 contigs, was downloaded from the PUT database (<http://www.plantgdb.org>).

**2.2. Gene Mining.** To identify putative genes predominantly expressed in wheat flowers, BLAST search was conducted using the PUT contig sequences as queries against the EST database consisting of the 67 cDNA libraries. ESTs matching each PUT contig with  $\geq 90\%$  homology were recorded and classified according to their library origins from the five aforementioned tissue types. The contigs with matched ESTs solely from inflorescences and those with significantly more matched ESTs from inflorescence tissues than from any of the other four types of tissues were considered to be the putative genes predominantly expressed in inflorescences. The probability to achieve the ESTs matching to a contig in the noninflorescence libraries was estimated based on the random sampling principle using the equation:  $P = C_N^n f^n (1-f)^{N-n}$ , as described by Ding et al. [27], where  $f$  is the EST frequency of a contig in the inflorescence libraries estimated by dividing the number of matched ESTs by the total number of ESTs in the libraries and  $n$  and  $N$  are the number of matched ESTs and the total number of ESTs in libraries of other tissues types, respectively. A sampling probability  $\leq 0.0001$  was considered an indication of significant difference in expression levels between the inflorescence tissues and noninflorescence tissues.

Chromosome-specific genomic DNA sequences corresponding to the retained contigs were obtained by retrieving the chromosome-assigned homologous scaffolds in the TGAC database [28] (<http://www.tgac.ac.uk/grassroots-genomics>) that showed 95% homology to the contig sequences. Coding DNA sequences (CDS) corresponding to the contigs in the scaffolds were predicted using the gene-finding program FGENESH [29] (<http://linux1.softberry.com>) and verified by alignment with the homologous ESTs. Scaffolds that did not contain the full-length target CDS were extended via in silico walking, using a parameter of 100% match, with the Roche 454 sequence reads of Chinese Spring [30] (<http://www.cerealsdb.uk.net>).

To confirm the expression predominance of the identified genes in wheat flowers, ESTs with  $\geq 99\%$  homology to the individual CDS identified from the sixty-seven libraries were subjected to further analysis. The sampling probability to achieve the EST frequency of a CDS in the noninflorescence libraries was again calculated. A sampling probability  $\leq 0.01$  was used as the threshold for declaration of significant difference. Expression specificity of the candidate genes was estimated similarly.

The identified genes were coded in numerical order with the prefix “IDG” (inflorescence development-related gene). Orthologous genes were given a common name but marked with a chromosome assignment suffix. Multiple copies derived from duplication of an ancestral gene at the same chromosome were differentiated by a numerical suffix.

**2.3. RNA-Seq Data Analysis.** RNA-data sets used in the analysis included the developmental time course series in five tissues (spike or inflorescence, root, leaf, grain, and stem) each with three different developmental stages. The stages of inflorescence included Z32, two nodes stage; Z39, meiosis; Z65, anthesis of C.S. [31] (<http://wheat.pw.usda.gov/WheatExp>), those for C.S. pistil and stamen by Yang et al. [17], and those for anther at meiosis data from the URGI public database (<https://wheat-urgi.versailles.inra.fr/Seq-Repository/Expression>). After trimmed adapters and removing vectors and low quality reads with Adapter-Removal in a setting of quality base = 33 [32], the RNA-seq reads were mapped to the sequence database consisting of genomic DNA sequences of each identified candidate gene using HISAT2 [33]. FPKM of a gene was estimated with reads matching to each gene with 100% identity, which were counted using featureCounts with both readExtension5 and readExtension3 set at 70 [34].

To estimate the relative expression specificity (RES) of a gene in a certain tissue (A), the tissue (B) in which the gene was most highly expressed among all but tissue A was identified. RES was then measured by dividing the difference of the normalized expression values between A and B by the expression value in A. The higher the RES value is, the more specific the gene expression is in A. Significance of the expression difference between two tissues was examined via  $\chi^2$  test. A gene was considered to be differentially expressed when the difference reached significance at  $P = 0.05$ .

To reflect the relative expression abundance of each gene in each tissue among the identified genes, a heat map was drawn with the quotients obtained from dividing the FPKM value (the maximal one if two or more developmental stages of a tissue were involved) of a specific gene in a tissue by the FPKM means of all genes. A log-transformation was applied to facilitate the mapping drawing.

**2.4. RT-PCR.** Root, node, internode, flag leaf, glume, lemma, palea, lodicule, stamen, pistil, and rachis tissues of the common wheat landrace “Wangshuibai,” grown in a field during the normal growing season, were collected at the heading stage, and developing kernels were collected at the 9th day postanthesis for RNA extraction. RNA was extracted using TRIzol reagent (Invitrogen, CA) following the manufacturer’s protocol and quantified with an Ultraspec 2100 Pro spectrometer (Amersham Pharmacia, UK). To eliminate DNA contamination, RNA samples were treated with RNase-free DNase I (Fermentas, Canada), following the product manual. First-strand cDNA was synthesized using oligo(dT) primer with 3  $\mu$ g of total RNA using M-MLV reverse transcriptase (Life Technologies, CA) according to the manufacturer’s instructions.

Semiquantitative RT-PCR (sqRT-PCR) was performed in a 25  $\mu$ l total reaction volume supplemented with 10–20 ng of first-strand cDNA as the template, 5 pmol concentration of each primer, 5 nmol dNTPs for each, 1 U rTaq DNA polymerase (Takara, Japan), and 1x PCR buffer supplied together with the enzyme. The wheat  $\alpha$ -tubulin gene was used in calibrating cDNA templates. The following thermal cycle profile was observed: 94°C for 3 min; 26–32 cycles of 94°C

for 20 s, 56–62°C (depending on the primer sets) for 25 s, and 72°C for 30 s, and a final extension step of 72°C for 5 min. PCR products were resolved in 2.0% (w/v) agarose gels and visualized with ethidium bromide staining. RT-PCR reactions were independently repeated three times or more to ensure reproducibility.

Quantitative real-time RT-PCR (qRT-PCR) amplifications in 20  $\mu$ L volumes, containing 10  $\mu$ L SYBR<sup>®</sup> Green qPCR Mix (Toyoba), 20 ng template, and 8 pmol of each primer, were performed using the StepOne<sup>™</sup> Real-Time PCR instrument (Applied Biosystems), following the protocol described in [35]. The reactions were conducted in triplicate. The cycle threshold values for each target gene were normalized based on values obtained in corresponding reactions for the wheat  $\alpha$ -tubulin gene. The relative expression was estimated by employing the  $2^{-\Delta\Delta CT}$  method [36].

The RT-PCR primers used in the present study were designed with MacVector 11 (MacVector, NC) and are listed in Table S2. The product length ranged from 150–300 bp.

**2.5. Gene Annotation and Pathway Assignment.** Genes were functionally annotated through homologous search of the NCBI nonredundant (Nr) protein, KEGG [37] (<http://www.kegg.jp/blastkoala>), and Pfam [38] (<http://pfam.xfam.org>) databases. Subcellular locations of the proteins were predicted using the TargetP 1.1 server [39] (<http://www.cbs.dtu.dk/services/TargetP>). Pathway assignments were based on KEGG pathway mapping ([http://www.kegg.jp/kegg/tool/map\\_pathway1.html](http://www.kegg.jp/kegg/tool/map_pathway1.html)) and keyword search of the plant metabolic pathway databases (<http://www.plantcyc.org>).

### 3. Results

**3.1. Identification of Genes Preferentially Expressed in Inflorescence.** Of the 67 cDNA libraries in line with the screening conditions, 29 libraries containing 140,092 sequences were from seed tissues; only three cDNA libraries including 17,732 sequences were from stem tissues (Table S1). The number of ESTs in each library ranged from 1000 to more than 10,000. For identification of the inflorescence development-related genes, these libraries were classified into five types according to the tissues used in library preparation, including seedling-stage leaf and stem, seedling to tillering-stage root, seed (from DPA3 to mature), and inflorescence (including premeiotic anthers, anthers at meiosis, pistil and ovary, immature inflorescence, lemma and palea, spike before flowering, and spike at flowering).

Alignment of the 77,657 wheat EST contigs, downloaded from the PlantGDB-assembled unique transcript (PUT) database, with ESTs of the abovementioned cDNA libraries led to the identification of 335 contigs that matched (95% homology) significantly more ESTs from inflorescence tissues, including nearly one-third with ESTs only from inflorescence tissue libraries. Using these contig sequences as queries in search of the TGAC database [28] with 90% as the homology threshold, 318 nonabundant genomic DNA sequences were obtained, 294 of which yield the expected open reading frames (ORF) with EST

TABLE 1: Chromosome distribution of the identified wheat inflorescence development-related genes.

Gene_ID	Chromosomes			Gene_ID	Chromosomes		
<i>IDG001</i>	1A	1B	1D	<i>IDG031</i>	<i>c</i>	<i>c</i>	3D
<i>IDG002</i>	1A	1B	1D	<i>IDG032</i>	4A	<i>b</i>	5D
<i>IDG003</i>	1A	1B	1D	<i>IDG033</i>	4A	4B	4D
<i>IDG004</i>	1A	1B	1D	<i>IDG034</i>	<i>c</i>	4B	4D
<i>IDG005</i>	1A	1B	1D	<i>IDG035</i>	5A (2)	4B	4D (2)
<i>IDG006</i>	1A	1B	1D	<i>IDG036</i>	5A	5B	5D
<i>IDG007</i>	1A	1B	1D	<i>IDG037</i>	5A	5B	5D
<i>IDG008</i>	1A	1B	1D	<i>IDG038</i>	5A	5B	5D
<i>IDG009</i>	1A	1B	1D	<i>IDG039</i>	5A	5B	5D
<i>IDG010</i>	1A	1B	1D	<i>IDG040</i>	<i>c</i>	5B (2)	5D
<i>IDG011</i>	1A	<i>c<sup>l</sup></i>	<i>a</i>	<i>IDG041</i>	<i>b</i>	5B	5D
<i>IDG012</i>	<i>c</i>	1B	1D	<i>IDG042</i>	6A (8)	6B (3)	6D (3)
<i>IDG013</i>	<i>b</i>	1B	<i>b</i>	<i>IDG043</i>	<i>b</i>	6B	6D (2)
<i>IDG014</i>	2A	2B	2D	<i>IDG044</i>	6A	6B	6D
<i>IDG015</i>	2A	<i>a</i>	2D	<i>IDG045</i>	6A	6B	6D
<i>IDG016</i>	2A	2B	2D	<i>IDG046</i>	6A	<i>c</i>	<i>b</i>
<i>IDG017</i>	2A	2B	<i>b</i>	<i>IDG047</i>	6A	6B	6D
<i>IDG018</i>	<i>a</i>	2B	2D	<i>IDG048</i>	6A	6B	6D
<i>IDG019</i>	3A	3B	3D	<i>IDG049</i>	6A (2)	6B	6D
<i>IDG020</i>	3A	3B	3D	<i>IDG050.1</i>	6A	<i>a</i>	6D
<i>IDG021</i>	3A	3B	3D	<i>IDG050.2</i>	6A	<i>b</i>	6D
<i>IDG022</i>	3A	3B	3D	<i>IDG051</i>	<i>b</i>	6B	6D
<i>IDG023</i>	3A	3B	3D	<i>IDG052</i>	<i>c</i>	6B (2)	<i>c</i>
<i>IDG024</i>	3A	3B	3D	<i>IDG053</i>	<i>a</i>	6B	<i>a</i>
<i>IDG025</i>	3A	3B	3D	<i>IDG054</i>	7A	7B	7D
<i>IDG026</i>	3A	3B	3D	<i>IDG055</i>	7A	7B	7D
<i>IDG027</i>	3A	<i>c</i>	3D	<i>IDG056</i>	7A	7B	7D
<i>IDG022</i>	3A	3B	3D	<i>IDG057</i>	7A	<i>a</i>	7D
<i>IDG029</i>	3A (2)	3B (2)	3D (2)	<i>IDG058</i>	<i>b</i>	7B	<i>b</i>
<i>IDG030</i>	3A	3B	3D	<i>IDG059</i>	<i>c</i>	<i>c</i>	7D

*a*: corresponding gene was not found; *b*: genomic DNA is available but not supported by ESTs; *c*: the gene was not inflorescence predominantly expressed. Copy number is shown in parenthesis.

support in gene prediction with the gene-finding program FGENESH [29]. With 99% homology as the cutoff, 187 genes matched significantly more ESTs from the inflorescence tissues than from any of the other four types of tissues ( $P=0.01$ ). We disregarded 17 of these genes that were neither more abundant in the inflorescence RNA-seq datasets nor in the anther and pistil data sets. Interestingly, most ESTs matching to 14 of the 17 genes came from Ogihara's unpublished cDNA libraries Wh\_FL or Wh\_f, which were constructed with spikelets or spikes at flowering stage. Probably, these genes are expressed at a developmental inflorescence stage not included in the tissues for the RNA sequencing. The remaining 170 genes represented 59 inflorescence development-related nonredundant gene sets, since wheat is an allohexaploid species and the majority of the genes have three orthologous copies (Table 1 and Table S3).

In terms of genomic distribution, the identified genes distributed to chromosomes of homoeologous groups 1, 3, and 6 accounted for 68.2%, those to group 4 chromosomes accounted for only 5.3%. Majority of the nonredundant gene sets have homologs in all three homoeologous chromosomes; however, not all of them showed inflorescence development-related preferential expression, suggesting functional differentiation have occurred among them. Of the 59 nonredundant gene sets, 13 had inflorescence development-related homologs in only two of three subgenomes, eight were solely identified in a single subgenome. Eight nonredundant gene sets had intrachromosomal duplications, of which five were distributed to the homoeologous group 6 chromosomes. A notable example was *IDG042* that had duplications in all three group 6 chromosomes, with a total of 14 copies showing inflorescence development-related preferential expression.

**3.2. Expression Specificity of the Identified Genes.** In mining genes predominantly expressed in inflorescence, we considered all tissues from inflorescence as a whole. It was noted, in the EST analysis, that some genes were solely expressed in spikes at or before anthesis, some were solely in stamen or in pistil, apart from those expressed more abundantly in spikes than in other tissues. To validate the expression patterns of these genes, we tested the significance of expression difference and estimated the expression specificity (RES) of inflorescence (including the stages of Z32, Z39, and Z65), stamen, and pistil relative to the individual vegetative tissues (including kernel) using the RNA-seq data. Based on the expression profiles, we were able to classify the 170 genes into three groups, each with two subgroups (Figure 1 and Table S3). Within each subgroup, the expression profiles were similar between genes, but the relative abundances were not identical even between orthologous genes.

The first group, G1, consisted of 42 genes from 10 gene sets. Overall, they had a low expression level and were expressed more abundantly and specifically in majority of the cases, in the inflorescence as a whole. A few genes, such as *IDG035.2-5A*, *IDG035.1-4D*, *IDG035.2-4D*, *IDG042.1-6A*, *IDG042.3-6A*, *IDG042.4-6A*, and *IDG042.8-6A*, were supported by ESTs but were matched to a negligible number of reads in the RNA-seq datasets. They were classified together with their orthologous or paralogous homologs, since their expression profiles in EST analysis were similar. Genes in subgroup G1-1 also showed enhanced expression, even though less abundantly, in stamen or pistil. Different from those in G1-1, genes in subgroup G1-2 had negligible fragments per kilobase of cDNA model per million mapped reads (FPKM) values from tissues other than the inflorescence.

The second group, G2, consisted of 20 gene sets and 48 genes. This group was basically characterized by expression in one or more of the vegetative tissues and an even higher level of expression in the stamen and/or pistil. *IDG011-1A* was the only exception, which was predominantly expressed in the inflorescence as a whole in spite of a significantly higher expression in the stamen and pistil relative to the vegetative tissues. Genes in subgroup G2-1 all had a significantly higher level of expression in the pistil relative to the vegetative tissues, and except for *IDG044-6B* and *IDG044-6D*, a significantly higher level of expression in the stamen as well. A few genes in this subgroup were expressed more abundantly in the inflorescence (*IDG011-1A*) or in the stamen (*IDG001-1A*, *IDG001-1D*, and *IDG045*). Genes in subgroup G2-2 were expressed more abundantly in the stamen relative to the vegetative tissues, but their expression levels in the pistil were not different from or even lower than those in at least one of the vegetative tissues.

Group G3 was the largest group, including 80 genes from 29 gene sets, characterized by a predominant expression in stamen and a negligible level of expression in the vegetative tissues. The RES values (stamen versus vegetative tissues) were high, ranged from 0.82–1.0. G3-1 consisted of 39 genes, all with a negligible FPKM value from the inflorescence. The remaining genes were different from genes in G3-1 by a significantly enhanced expression in the inflorescence as well

relative to the vegetative tissues, although the expression level was much lower in most cases.

To verify the expression specificity experimentally, sqRT-PCR was performed with tissues of root, node, internode, flag leaf, glume, lemma, palea, lodicule, stamen, pistil, rachis, and kernel 9th day postanthesis, using 27 pairs of primers that corresponds to 64 members of the identified genes (Table S2). All PCR reactions revealed a pattern of predominant expression in at least one of the floral tissues or organs but not in kernels and vegetative organs (Figure 2), which, by and large, were in agreement with results from the EST and RNA-seq data analysis. The qRT-PCR of a selected set of genes, including those with a relatively lower expression level and those expressed in multiple tissues, further confirmed these results (Figure S1).

Generally speaking, the tissue expression profiles between orthologous genes inferred based on the EST data were similar to those from the RNA-seq analysis. But a few exceptions were noted. In most of these cases, a low expression level of certain orthologous members was likely the cause of the discrepancy, for instance, some members of *IDG035* and *IDG042*.

**3.3. Functional Annotation of the Identified Genes.** Forty-nine of the identified gene sets were annotated via homology search and classified into five categories according to their putative biological functions (Table 2). It has to be mentioned that some genes could functionally fall into multiple categories.

The first category included only two gene sets; both encode allergenic proteins. The biological functions of this class of proteins in floral development have not been well characterized. *IDG035* encodes group 3 grass pollen allergens, which have sequence similarity to expansins that promote plant cell wall enlargement and thereby serve as cell wall-loosening agents [40].

The second category included 11 gene sets, most of which belonged to the G2 expression type. In this category, five gene sets code for proteins related to JA, ET, and GA signaling, three for MADS-box transcription factors and three for H2A and H2B proteins. The MADS-box transcription factor proteins encoded by *IDG001* and *IDG021* have 87% similarity. *IDG001* is orthologous to the rice *OsMADS4*. According to the ABCDE model for floral organ identity specification [41], *IDG001* and *IDG021* belong to class B MADS-box genes. *IDG044* encodes an AGL6-like MADS transcription factor and is functionally similar to class E genes [42].

The third category included 28 gene sets, accounting for 57% of the annotated. They were all predominantly or specifically expressed in stamen. Most of them were associated with substance production, transportation, and assembly for anther and pollen development. Of this category, 18 gene sets code for proteins associated with fatty acid and lipid metabolism. Homologs in other plants of most of these genes have been associated with the process of pollen wall development, such as suberin biosynthesis [43, 44], cutin biosynthesis [45–47], pollen sporopollenin biosynthesis [48], and pollen exine formation [49–51]. Other genes in this category have also been associated

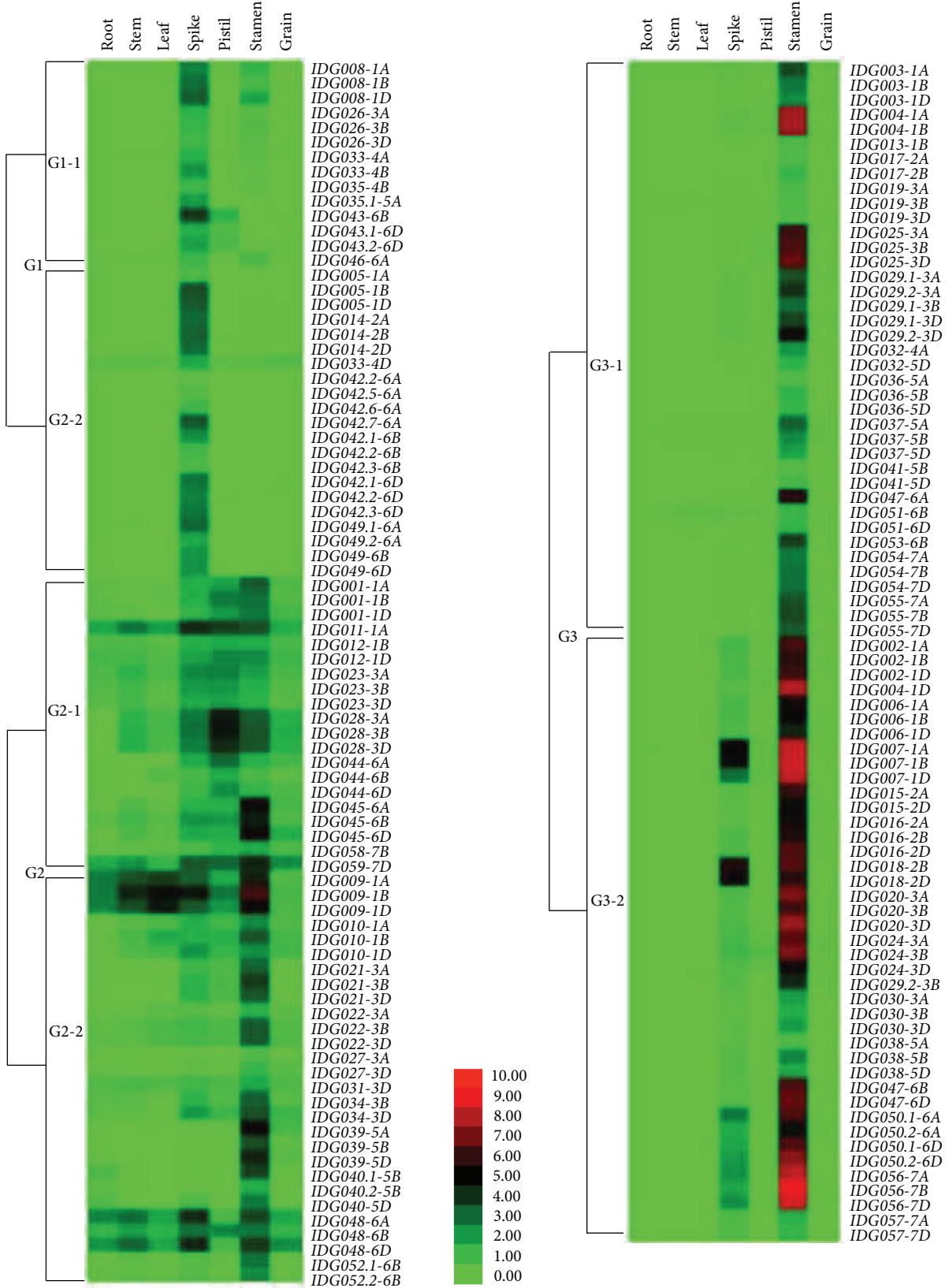


FIGURE 1: Relative expression abundance of the identified genes in different tissues based on RNA-seq data. Apart from seven genes with a negligible number of reads, the remaining 163 genes were divided according to their expression profiles into three groups, each of which had two subgroups.

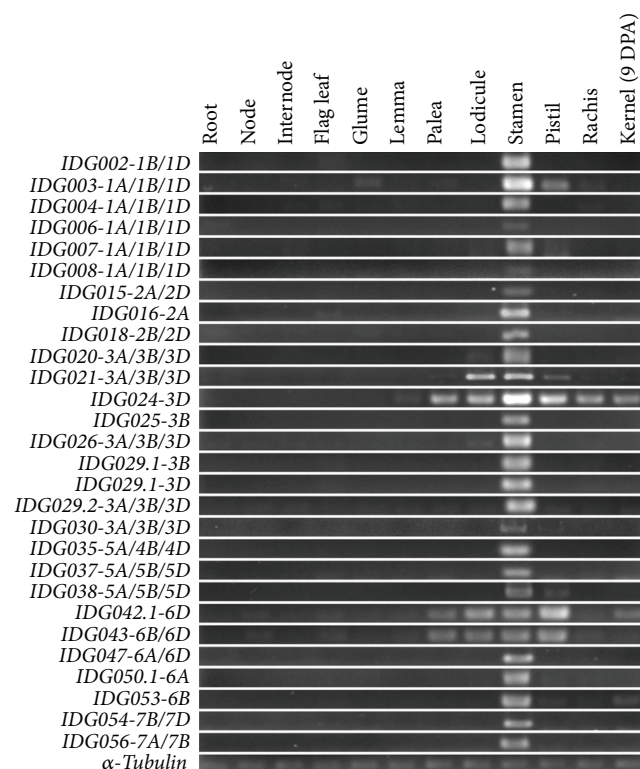


FIGURE 2: Expression of a selected set of identified genes in different tissues. The tissues used included kernel 9th day postanthesis, root, node, internode, flag leaf, glume, lemma, palea, lodicule, stamen, pistil, and rachis at the heading stage of common wheat landrace “Wangshuibai.”

with anther and pollen development. *IDG038* encodes an acidic peroxidase that might participate in the synthesis of phenylpropanoids present in sporopollenin [52]. The products encoded by *IDG007*, *IDG018*, *IDG050*, *IDG055*, and *IDG056* were related to pollen wall formation [53–55]. Both *IDG004* and *IDG025* code for anther-specific RTS-like proteins, required for male fertility and affecting tapetal development [56]. *IDG006* encodes a galactosyltransferase, which is implicated in the accumulation control of glycosylated flavonols in pollen [57]. In *Arabidopsis*, a type II  $\beta$ -(1,3)-galactosyltransferase is required for pollen exine development [58]. *IDG020* codes for a late cornified envelope- (LCE-) like proline-rich protein. LCE proteins are involved in the cornified cell envelope assembly of skins and associated with ROS detoxification [59].

The fourth category only has four gene sets and is associated with pollination and pollen-stigma interactions. For instance, the subtilisin-like protease encoded by *IDG002* was related to anther dehiscence [60]; the products of *IDG008* were related to pollen tube growth [61]; the chemo-cyanins encoded by *IDG040* could be involved in the pollination process [62] and induce pollen tube chemotropism as a diffusible chemotropic factor [63]. *IDG028* codes for a protodermal factor 1-like protein mainly in the pistil. This protein is related to reactive oxygen species (ROS) homeostasis [64]. ROS are involved in pollen tube growth

and rupture [65, 66], implying the role of modulating ROS levels in male reproductive development [67].

The fifth category, the “other” in Table 2, had only four annotated gene sets. Their specific functions in reproductive development still require clarification.

## 4. Discussion

Inflorescence represents a highly specialized plant tissue producing seeds for propagation. Deciphering genes involved in its development is the first step to understand the essence of reproduction and of great importance for seed production manipulation. In this study, we identified 59 nonredundant wheat gene sets that were differentially expressed in wheat inflorescence and encode proteins with diverse functions. Majority of the identified genes were associated with metabolic activities and wall assembly required for the specialized process of pollen maturation and pollination, while few showed predominance or specificity in macrosporogenesis. On one hand, this could be attributed to the fact that much more ESTs from libraries made with anthers were used in the analysis, which had limitation of development stage and tissue-type coverage; on the other hand, it was probably due to the specific structure of pollen grains whose formation requires expression of a specific set of genes or gene network that made the related genes easily recognized through the differential analysis. Our results complemented well with previous studies involved in wheat floral development. The microarray-based transcriptomic analysis of anthers by Crismani et al. was mainly focused on identification of early meiotic genes [15]. In the RNA-seq data comparison of pistillody stamen versus pistil, pistillody stamen versus stamen, and pistil versus stamen, Yang et al. identified 206 genes highly correlated with stamen and pistil development [17]. Among them, however, only a few were functionally annotated as identically as the genes presented in this paper. It was noted that nearly one-third of the identified gene sets in the present study displayed differentiated expression profiles in terms of their subgenome orthologs, implying functional diversification in polyploidy wheat for the inflorescence development.

The whole process of inflorescence development is under regulatory control. A set of MADS transcription factors regulate floral organ identity specification [41, 68, 69]. The three MADS-box genes we identified, one for E-class MADS proteins and two for B-class MADS proteins, differed in their expression profiles (Figure 1), even though both *IDG001* and *IDG021* were mainly expressed in stamen, suggesting they act in concert in determining the anther and pistil identity. The pistillody in alloplasmic wheat was related to expression pattern alteration of class B genes [68]. The identification of genes related to JA, ET, and GA signaling added support for the important roles of JA, ET, and GA signaling cross-talks playing in stamen development [70, 71]. ET signaling is involved in multiple aspects of floral organ development, for instance, nectar secretion, accumulation of stigmatic exudate, and development of the self-incompatible response [72], floral organ

TABLE 2: Functional classifications of the annotated gene sets.

Gene set	Annotation	Molecular or biological function	Reference
1. Allergen			
<i>IDG035</i>	Group 3 grass pollen allergen		[93]
<i>IDG039</i>	Peamaclein-like		
2. Expression regulation			
<i>IDG001</i>	MADS-box transcription factor WM14	Organ identity specification	[68]
<i>IDG003</i>	Gibberellin-regulated protein II-like	Gibberellin signaling	[70]
<i>IDG005</i>	Cysteine protease-like protein	Ethylene signaling	[73]
<i>IDG011</i>	Histone H2A	Chromosome modeling	
<i>IDG012</i>	Histone H2A	Chromosome modeling	
<i>IDG021</i>	MADS-box transcription factor	Organ identity specification	[68]
<i>IDG031</i>	Glyoxysomal fatty acid beta-oxidation multifunctional protein MFP-a	JA signaling	[94]
<i>IDG044</i>	Transcription factor AGL6-like	Organ identity specification	[69]
<i>IDG046</i>	1-aminocyclopropane-1-carboxylate oxidase, EC 1.14.17.4	Ethylene signaling	[72, 95]
<i>IDG048</i>	Histone H2B	Chromosome modeling	
<i>IDG052</i>	1-aminocyclopropane-1-carboxylate oxidase, EC 1.14.17.4	Ethylene signaling	[72, 95]
3. Anther development and pollen wall formation			
<i>IDG004</i>	Anther-specific protein RTS-like	Tapetal development	[56]
<i>IDG006</i>	Galactosyltransferase, EC 2.4.1.-	Cell wall assembly	[81]
<i>IDG007</i>	Protease inhibitor/seed storage/LTP family protein-like	Pollen wall formation	[53]
<i>IDG009</i>	Phosphoethanolamine N-methyltransferase EC 2.1.1.103	Lipid metabolism	[92]
<i>IDG010</i>	Long chain acyl-CoA synthetase, EC 6.2.1.3	Lipid metabolism	[92]
<i>IDG013</i>	Triacylglycerol lipase, EC 3.1.1.3	Lipid metabolism	
<i>IDG017</i>	Cytochrome P450 86B1-like, EC 1.14.15.-	Lipid metabolism	[83]
<i>IDG018</i>	Putative RAFTIN1/BURP domain-containing protein	Pollen wall formation	[53]
<i>IDG019</i>	Cytochrome P450 94A1-like	Lipid metabolism	[45]
<i>IDG020</i>	Late cornified envelope-like proline-rich protein	ROS detoxification	[59]
<i>IDG022</i>	Glycerol-3-phosphate acyltransferase 6, EC 2.3.1.198	Lipid metabolism	[46]
<i>IDG025</i>	Anther-specific protein RTS-like	Tapetal development	[56]
<i>IDG027</i>	Nonspecific phospholipase C2, EC 3.1.4.3	Lipid metabolism	
<i>IDG030</i>	Cytochrome P450 86A2-like	Lipid metabolism	[83]
<i>IDG032</i>	Acyl-[acyl-carrier-protein] desaturase, EC 1.14.19.2	Lipid metabolism	
<i>IDG033</i>	Fatty acyl-CoA reductase 2. EC 1.2.1.42	Lipid metabolism	[48]
<i>IDG036</i>	Cytochrome P450 86B1, fatty acid $\omega$ -hydroxylase, EC 1.14.15.3	Lipid metabolism	[83]
<i>IDG037</i>	Acyl-[acyl-carrier-protein] desaturase, EC 1.14.19.2	Lipid metabolism	
<i>IDG038</i>	Acidic peroxidase, EC 1.11.1.7	Phenylpropanoid biosynthesis	[52]
<i>IDG041</i>	Cytochrome P450 86A1-like, fatty acid $\omega$ -hydroxylase	Lipid metabolism	[83]
<i>IDG045</i>	Fatty aldehyde decarboxylase, EC 4.1.99.5	Lipid metabolism	[50, 51]
<i>IDG050</i>	Nonspecific lipid-transfer protein C6-like	Pollen wall formation	[53]
<i>IDG051</i>	Acyltransferase-like protein, EC 2.3.1.-	Lipid metabolism	
<i>IDG054</i>	Triacylglycerol lipase 2-like, EC 3.1.1.13	Lipid metabolism	
<i>IDG055</i>	Bidirectional sugar cr SWEET	Pollen wall formation	[54, 55]
<i>IDG056</i>	RAFTIN1 protein/BURP domain protein-like	Pollen wall formation	[53]
<i>IDG057</i>	Acyltransferase-like protein, EC 2.3.1.-	Lipid metabolism	
<i>IDG059</i>	Acyl carrier protein	Lipid metabolism	
4. Pollination and pollen-stigma interactions			
<i>IDG002</i>	Subtilisin-like protease	Anther dehiscence	[60]
<i>IDG008</i>	Chalcone synthase-like	Biosynthesis of flavonols	[61]
<i>IDG028</i>	Protodermal factor 1-like	ROS homeostasis	[64]



TABLE 2: Continued.

Gene set	Annotation	Molecular or biological function	Reference
<i>IDG040</i>	Chemocyanin	Pollen tube attraction	[62]
5. Others			
<i>IDG014</i>	Nitrate-induced NOI protein	Plant defense	[96]
<i>IDG034</i>	Nucleoside diphosphate kinase, EC 2.7.4.6	Nucleotide triphosphate generation	
<i>IDG053</i>	Zinc transporter-like	Early reproductive development	
<i>IDG058</i>	Heat shock protein	Thermotolerance	

senescing [73], pollen thermotolerance [74], and timing of anther dehiscence [71]. Cheng et al. demonstrated that GA regulates stamen development through JA signaling [75]. Mutations of genes encoding JA-biosynthetic enzymes result in failure of filament elongation, delayed anther dehiscence, and unviable pollens [70, 76, 77].

A few identified genes were associated with pollen-stigma interactions. All but *IDG028* had transcripts in spikes at meiosis and anthesis stages as well as in anthers but the abundance varied considerably. The subtilase-encoding genes related to anther dehiscence (*IDG002*) were overwhelmingly expressed in anthers, especially at the tetrad stage [60]. In male-sterile lines, their expression was downregulated [77]. Transcripts of pollen allergen-encoding *IDG035* and ACO-encoding *IDG046* were also present in spikes at this stage. Accumulating evidence indicates the involvement of ET signaling in fertilization [78, 79]. Valdiva et al. showed that disruption of a maize group-I allergen affected pollen-pollen competition for access to the ovules [80]. In addition, the galactosyltransferase-encoding *IDG006*, predominantly expressed in stamen, was related to pollen tube elongation [81].

Most of the genes showing anther-specific or predominant expression are related to tapetal and pollen developments (Table 2). Particularly worth mentioning are the gene set-encoding enzymes or proteins participating in lipid metabolism. Lipid metabolism is important to pollen development because the distinct pollen wall structure is mainly made of fatty (lipid) substances produced in the tapetum of anthers [82]. Among the lipid metabolism-related gene sets, five (*IDG017*, *IDG019*, *IDG030*, *IDG036*, and *IDG041*) encode members of 86A, 86B, and 94A subfamilies of cytochrome P450 proteins that are related to fatty acid  $\omega$ -hydroxylation in primary fatty alcohols and suberin monomer biosynthesis for formation of anther cuticle and pollen sporopollenin in monocots and dicots [43–45, 47, 83–86]. These cytochrome P450 proteins might function in different subcellular locations, since *IDG017*, *IDG019*, and *IDG030* have secretory signal peptides, while *IDG036* and *IDG041* do not.

Pathways emerging from the lipid metabolism-related genes included those for alkane and fatty alcohol production and glycerophospholipid metabolism. Genes encoding long-chain acyl-CoA synthetase (*IDG010*) and fatty aldehyde decarboxylase (*CER1*, *IDG045*), the two enzymes involved in the plant alkane-forming pathway [87], were coexpressed in the wheat stamen. This *CER1* gene was

downregulated in pistil or pistillody stamen [17], suggesting its specificity to stamen development. Very long-chain (VLC) alkanes are major components of the tryphine layer covering pollen grains and are needed for proper pollen-pistil signaling and fertility [50]. Mutation of *CER1* in both *Arabidopsis* and rice caused defective pollens [50, 51]. In *Arabidopsis*, the acyl-CoA synthetase gene *ACOS5* was upregulated in the tapetal cells. Its mutation led to failure in pollen production and pollen wall formation [88].

Fatty alcohols are components of surface lipid barriers such as anther cuticle and pollen wall [89]. Fatty acyl-CoA reductase, encoded by *IDG033* in wheat and *Ms2* in *Arabidopsis*, is the key enzyme for the production of fatty alcohols in plastids [48]. Mutation of *Ms2* led to abnormal pollen wall development and reduced pollen fertility [90, 91]. Like *Ms2*, the *IDG033*-encoded proteins have plastidic localization signal peptides. The encoded products of *IDG059*, *IDG032*, and *IDG037*, probably involved in the production of fatty acyl-CoA reductase substrates, all have plastidic localization signals. The diacylglycerol acyltransferase- (DGAT-) like protein encoded by *IDG057* could also carry a plastidic peptide. The *Arabidopsis* *DGAT1* contributes to triacylglycerol biosynthesis and its function loss causes critical defects in normal pollen and embryo development [84]. However, information about its link to the plastidial fatty alcohol pathway is still lacking. The expression profiles of these genes were different, although all expressed in stamen. *IDG059*, *IDG032*, and *IDG037* appeared to be specifically expressed in this organ.

Among the lipid metabolism-related genes, five (*IDG009*, *IDG013*, *IDG022*, *IDG027*, and *IDG054*) encode proteins putatively associated with glycerophospholipid metabolism, in agreement with the findings of Yang et al. [17]. The *IDG009*-encoded phosphoethanolamine *N*-methyltransferase (PEAMT) is the committing enzyme for choline biosynthesis. In *Arabidopsis*, silencing the PEAMT gene resulted in temperature-sensitive male sterility and salt hypersensitivity [92]; knockdown of *GPAT6*, the homolog of *IDG022*, caused defective pollen grains [46]. PEAMT and *GPAT6* also affected pollen tube growth [46, 92]. Moreover, *IDG051* encode proteins predicted with lysophosphatidylethanolamine acyltransferase activities, which probably participates in the phospholipid metabolism in mitochondria.

Only 59 inflorescence development-related nonredundant gene sets were identified in this study. This could be

much less than the actual number of genes differentially expressed in inflorescence tissues. We reasoned that the EST libraries used in gene mining, which had limitations in volume size and representation of tissue and developmental stages, and the strict standard used in gene mining were the main causes. A sampling probability  $\leq 0.0001$  has very likely increase type II error; however, it could minimize false positives, as shown in RNA-seq data analysis and RT-PCR validation, which is beneficial for correct data interpretation.

## 5. Conclusions

In this study, we identified 170 wheat genes for floral identity determination, anther and pollen development, pollen-pistil interaction, and others using the comparative transcriptomics approach. The potential importance of the identified genes to wheat inflorescence development was manifested in the enhanced or specific expression in the floral tissues. We noted that nearly one-third of the gene sets have undergone subgenome differentiation. Of the identified genes, those coding for enzymes or proteins participating in lipid metabolic pathway accounted for the largest category, implying the particularity and important roles of lipid metabolism in wheat reproductive development. This study is useful for understanding the gene network underlying wheat inflorescence morphology and fertility, which eventually will allow us to purposely manipulate fertility in breeding.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

## Acknowledgments

This study was partially supported by Natural Science Foundation of China (31430064 and 30025030), Ministry of Science and Technology of China (2016YFD0101004 and 2016ZX08002003), Jiangsu collaborative innovation initiative for modern crop production (JCIC-MCP), “111” project B08025, and fund from the innovation team program for Jiangsu universities (2014).

## Supplementary Materials

*Supplementary 1.* Figure S1: expression of a selected set of identified genes in different tissues, examined using qRT-PCR. The expression levels were estimated relative to that in stamen. The tissues used included kernel 9th day post-anthesis, root, node, internode, flag leaf, glume, lemma, palea, lodicule, stamen, pistil, and rachis at the heading stage of common wheat landrace “Wangshuibai.”

*Supplementary 2.* Table S1: cDNA libraries used in this study.

*Supplementary 3.* Table S2: primers used in RT-PCR.

*Supplementary 4.* Table S3: sequence, annotation, and expression specificity of 170 identified genes. \* and \*\* significantly different at  $P = 0.05$  and  $0.01$ , respectively.

## References

- [1] J. Kyozuka, “Grass inflorescence: basic structure and diversity,” *Advances in Botanical Research*, vol. 72, pp. 191–219, 2014.
- [2] H. T. Clifford, *Spikelet and Floral Morphology*, Washington, D.C, USA, Smithsonian, 1987.
- [3] L. G. Clark and R. W. Pohl, *Agnes Chase’s First Book of Grasses: The Structure of Grasses Explained for Beginners*, Washington D.C, USA, Smithsonian, 1996.
- [4] G. Theissen and H. Saedler, “Plant biology: floral quartets,” *Nature*, vol. 409, no. 6819, pp. 469–471, 2001.
- [5] D. B. Zhang and Z. Yuan, “Molecular control of grass inflorescence development,” *Annual Review of Plant Biology*, vol. 65, no. 1, pp. 553–578, 2014.
- [6] J. Y. Lee and D. H. Lee, “Use of serial analysis of gene expression technology to reveal changes in gene expression in Arabidopsis pollen undergoing cold stress,” *Plant Physiology*, vol. 132, no. 2, pp. 517–529, 2003.
- [7] T. H. Lee, Y. K. Kim, T. T. Pham et al., “RiceArrayNet: a database for correlating gene expression from transcriptome profiling, and its application to the analysis of coexpressed genes in rice,” *Plant Physiology*, vol. 151, no. 1, pp. 16–33, 2009.
- [8] H. P. Buermans and J. T. den Dunnen, “Next generation sequencing technology: advances and applications,” *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, vol. 1842, no. 10, pp. 1932–1941, 2014.
- [9] J. A. Schrauwen, P. F. de Groot, M. M. van Herpen et al., “Stage-related expression of mRNAs during pollen development in lily and tobacco,” *Planta*, vol. 182, no. 2, pp. 298–304, 1990.
- [10] J. Ma, D. S. Skibbe, J. Fernandes, and V. Walbot, “Male reproductive development: gene expression profiling of maize anther and pollen ontogeny,” *Genome Biology*, vol. 9, no. 12, article R181, 2008.
- [11] P. Deveshwar, W. D. Bovill, R. Sharma, J. A. Able, and S. Kapoor, “Analysis of anther transcriptomes to identify genes contributing to meiosis and male gametophyte development in rice,” *BMC Plant Biology*, vol. 11, no. 1, pp. 78–78, 2011.
- [12] N. Rutley and D. Twell, “A decade of pollen transcriptomics,” *Plant Reproduction*, vol. 28, no. 2, pp. 73–89, 2015.
- [13] P. Chaturvedi, T. Ischebeck, V. Egelhofer, I. Lichtscheidl, and W. Weckwerth, “Cell-specific analysis of the tomato pollen proteome from pollen mother cell to mature pollen provides evidence for developmental priming,” *Journal of Proteome Research*, vol. 12, no. 11, pp. 4892–4903, 2013.
- [14] T. Marcussen, S. R. Sandve, L. Heier et al., “Ancient hybridizations among the ancestral genomes of bread wheat,” *Science*, vol. 345, no. 6194, article 1250092, 2014.
- [15] W. Crismani, U. Baumann, T. Sutton et al., “Microarray expression analysis of meiosis and microsporogenesis in hexaploid bread wheat,” *BMC Genomics*, vol. 7, no. 1, p. 267, 2006.
- [16] S. McIntosh, L. Watson, P. Bundock et al., “SAGE of the developing wheat caryopsis,” *Plant Biotechnology Journal*, vol. 5, no. 1, pp. 69–83, 2007.
- [17] Z. J. Yang, Z. S. Peng, S. H. Wei, M. L. Liao, Y. Yu, and Z. Y. Jang, “Pistillody mutant reveals key insights into stamen and pistil development in wheat (*Triticum aestivum* L.),” *BMC Genomics*, vol. 16, no. 1, p. 211, 2015.

- [18] M. Houde, M. Belcaid, F. Ouellet et al., "Wheat EST resources for functional genomics of abiotic stress," *BMC Genomics*, vol. 7, no. 1, p. 149, 2006.
- [19] N. Z. Ergen and H. Budak, "Sequencing over 13,000 expressed sequence tags from six subtractive cDNA libraries of wild and modern wheats following slow drought stress," *Plant, Cell & Environment*, vol. 32, no. 3, pp. 220–236, 2009.
- [20] A. Manickavelu, K. Kawaura, K. Oishi et al., "Comparative gene expression analysis of susceptible and resistant near-isogenic lines in common wheat infected by *Puccinia triticina*," *DNA Research*, vol. 17, no. 4, pp. 211–222, 2010.
- [21] A. Manickavelu, K. Kawaura, K. Oishi et al., "Comprehensive functional analyses of expressed sequence tags in common wheat (*Triticum aestivum*)," *DNA Research*, vol. 19, no. 2, pp. 165–177, 2012.
- [22] M. Song, W. Xu, Y. Xiang, H. Jia, L. Zhang, and Z. Ma, "Association of jacalin-related lectins with wheat responses to stresses revealed by transcriptional profiling," *Plant Molecular Biology*, vol. 84, no. 1-2, pp. 95–110, 2014.
- [23] O. D. Anderson, N. Huo, and Y. Q. Gu, "The gene space in wheat: the complete -gliadin gene family from the wheat cultivar Chinese Spring," *Functional & Integrative Genomics*, vol. 13, no. 2, pp. 261–273, 2013.
- [24] K. Rikiishi and M. Maekawa, "Seed maturation regulators are related to the control of seed dormancy in wheat (*Triticum aestivum* L.)," *PLoS One*, vol. 9, no. 9, article e107618, 2014.
- [25] M. Domoki, A. Szucs, K. Jager, S. Bottka, B. Barnabas, and A. Feher, "Identification of genes preferentially expressed in wheat egg cells and zygotes," *Plant Cell Reports*, vol. 32, no. 3, pp. 339–348, 2013.
- [26] Z. Y. Chen, X. J. Guo, Z. X. Chen et al., "Genome-wide characterization of developmental stage- and tissue-specific transcription factors in wheat," *BMC Genomics*, vol. 16, no. 1, p. 125, 2015.
- [27] L. N. Ding, H. B. Xu, H. Y. Yi et al., "Resistance to hemibiotrophic *F. graminearum* infection is associated with coordinated and ordered expression of diverse defense signaling pathways," *PLoS One*, vol. 6, no. 4, article e19008, 2011.
- [28] P. A. Wilkinson, M. O. Winfield, G. L. Barker et al., "CerealsDB 3.0: expansion of resources and data integration," *BMC Bioinformatics*, vol. 17, no. 1, p. 256, 2016.
- [29] V. Solovyev, P. Kosarev, I. Seledsov, and D. Vorobyev, "Automatic annotation of eukaryotic genes, pseudogenes and promoters," *Genome Biology*, vol. 7, article S10, Supplement 1, 2006.
- [30] R. Brenchley, M. Spannagl, M. Pfeifer et al., "Analysis of the bread wheat genome using whole-genome shotgun sequencing," *Nature*, vol. 491, no. 7426, pp. 705–710, 2012.
- [31] F. Choulet, A. Alberti, S. Theil et al., "Structural and functional partitioning of bread wheat chromosome 3B," *Science*, vol. 345, no. 6194, article 1249721, 2014.
- [32] S. Lindgreen, "AdapterRemoval: easy cleaning of next-generation sequencing reads," *BMC Research Notes*, vol. 5, no. 1, p. 337, 2012.
- [33] D. Kim, B. Landmead, and S. L. Salzberg, "HISAT: a fast spliced aligner with low memory requirements," *Nature Methods*, vol. 12, no. 4, pp. 357–360, 2015.
- [34] Y. Liao, G. K. Smyth, and W. Shi, "featureCounts: an efficient general purpose program for assigning sequence reads to genomic features," *Bioinformatics*, vol. 30, no. 7, pp. 923–930, 2014.
- [35] M. Song, W. Q. Xu, Y. Xiang, H. Y. Jia, L. X. Zhang, and Z. Q. Ma, "Association of jacalin-related lectins with wheat responses to stresses revealed by transcriptional profiling," *Plant Molecular Biology*, vol. 84, no. 1-2, pp. 95–110, 2014.
- [36] K. J. Livak and T. D. Schmittgen, "Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C_T}$  method," *Methods*, vol. 25, no. 4, pp. 402–408, 2001.
- [37] M. Kanehisa, Y. Sato, and K. Morishima, "BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences," *Journal of Molecular Biology*, vol. 428, no. 4, pp. 726–731, 2016.
- [38] R. D. Finn, P. Coggill, R. Y. Eberhardt et al., "The Pfam protein families database: towards a more sustainable future," *Nucleic Acids Research*, vol. 44, no. D1, pp. D279–D285, 2016.
- [39] O. Emanuelsson, H. Nielsen, S. Brunak, and G. von Heijne, "Predicting subcellular localization of proteins based on their N-terminal amino acid sequence," *Journal of Molecular Biology*, vol. 300, no. 4, pp. 1005–1016, 2000.
- [40] D. J. Cosgrove, P. Bedinger, and D. M. Durachko, "Group I allergens of grass pollen as cell wall-loosening agents," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 12, pp. 6559–6564, 1997.
- [41] L. M. Zahn, J. H. Leebens-Mack, J. M. Arrington et al., "Conservation and divergence in the AGAMOUS subfamily of MADS-box genes: evidence of independent sub- and neofunctionalization events," *Evolution & Development*, vol. 8, no. 1, pp. 30–45, 2006.
- [42] H. F. Li, W. Q. Liang, R. D. Jia et al., "The AGL6-like gene OsMADS6 regulates floral organ and meristem identities in rice," *Cell Research*, vol. 20, no. 3, pp. 299–313, 2010.
- [43] R. Hofer, I. Briesen, M. Beck, F. Pinot, L. Schreiber, and R. Franke, "The *Arabidopsis* cytochrome P450 CYP86A1 encodes a fatty acid  $\omega$ -hydroxylase involved in suberin monomer biosynthesis," *Journal of Experimental Botany*, vol. 59, no. 9, pp. 2347–2360, 2008.
- [44] V. Compagnon, P. Diehl, I. Benveniste et al., "CYP86B1 is required for very long chain  $\omega$ -hydroxyacid and  $\alpha,\omega$ -dicarboxylic acid synthesis in root and seed suberin polyester," *Plant Physiology*, vol. 150, no. 4, pp. 1831–1843, 2009.
- [45] N. Tijet, C. Helvig, F. Pinot et al., "Functional expression in yeast and characterization of a clofibrate-inducible plant cytochrome P-450 (CYP94A1) involved in cutin monomers synthesis," *Biochemical Journal*, vol. 332, no. 2, pp. 583–589, 1998.
- [46] X. C. Li, J. Zhu, J. Yang et al., "Glycerol-3-phosphate acyltransferase 6 (GPAT6) is important for tapetum development in *Arabidopsis* and plays multiple roles in plant fertility," *Molecular Plant*, vol. 5, no. 1, pp. 131–142, 2012.
- [47] S. G. Rupasinghe, H. Duan, and M. A. Schuler, "Molecular definitions of fatty acid hydroxylases in *Arabidopsis thaliana*," *Proteins*, vol. 68, no. 1, pp. 279–293, 2007.
- [48] W. Chen, X. H. Yu, K. Zhang et al., "*Male Sterile2* encodes a plastid-localized fatty acyl carrier protein reductase required for pollen exine development in *Arabidopsis*," *Plant Physiology*, vol. 157, no. 2, pp. 842–853, 2011.
- [49] A. A. Dobritsa, Z. T. Lei, S. Nishikawa et al., "LAP5 and LAP6 encode anther-specific proteins with similarity to calcone synthase essential for pollen exine development in *Arabidopsis thaliana*," *Plant Physiology*, vol. 153, no. 3, pp. 937–955, 2010.
- [50] M. G. Aarts, C. J. Keijzer, W. J. Stiekema, and A. Pereira, "Molecular characterization of the CER1 gene of *Arabidopsis*

- involved in epicuticular wax biosynthesis and pollen fertility,” *The Plant Cell*, vol. 7, no. 12, pp. 2115–2127, 1995.
- [51] K. H. Jung, M. J. Han, D. Y. Lee et al., “Wax-deficient *anther1* is involved in cuticle and wax production in rice anther walls and is required for pollen development,” *The Plant Cell*, vol. 18, no. 11, pp. 3015–3032, 2006.
- [52] A. Skirycz, S. Jozefczuk, M. Stobiecki et al., “Transcription factor AtDOF4;2 affects phenylpropanoid metabolism in *Arabidopsis thaliana*,” *New Phytologist*, vol. 175, no. 3, pp. 425–438, 2007.
- [53] A. M. Wang, Q. Xia, W. S. Xie, R. Datla, and G. Selvaraj, “The classical Ubisch bodies carry a sporophytically produced structural protein (RAFTIN) that is essential for pollen development,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 24, pp. 14487–14492, 2003.
- [54] B. Yang, A. Sugio, and F. F. White, “Os8N3 is a host disease-susceptibility gene for bacterial blight of rice,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 27, pp. 10503–10508, 2006.
- [55] Y. F. Guan, X. Y. Huang, J. Zhu, J. F. Gao, H. X. Zhang, and Z. N. Yang, “RUPTURED POLLEN GRAIN1, a member of the MtN3/saliva gene family, is crucial for exine pattern formation and cell integrity of microspores in *Arabidopsis*,” *Plant Physiology*, vol. 147, no. 2, pp. 852–863, 2008.
- [56] H. Luo, J. Y. Lee, Q. Hu et al., “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 Molecular Biology*, vol. 62, no. 3, pp. 397–408, 2006.
- [57] L. P. Taylor and K. D. Miller, “The use of a photoactivatable kaempferol analogue to probe the role of flavonol 3-O-galactosyltransferase in pollen germination,” *Advances in Experimental Medicine and Biology*, vol. 505, pp. 41–50, 2002.
- [58] T. Suzuki, J. O. Narciso, W. Zeng et al., “KNS4/UPEX1: a type II arabinogalactan  $\beta$ -(1,3)-galactosyltransferase required for pollen exine development,” *Plant Physiology*, vol. 173, no. 1, pp. 183–205, 2017.
- [59] W. P. Vermeij, A. Alia, and C. Backendorf, “ROS quenching potential of the epidermal cornified cell envelope,” *Journal of Investigative Dermatology*, vol. 131, no. 7, pp. 1435–1441, 2011.
- [60] A. A. Taylor, A. Horsch, A. Rzepczyk, C. A. Hasenkampf, and C. D. Riggs, “Maturation and secretion of a serine proteinase is associated with events of late microsporogenesis,” *The Plant Journal*, vol. 12, no. 6, pp. 1261–1271, 1997.
- [61] G. J. vanEldik, W. H. Reijnen, R. K. Ruiter, M. M. A. vanHerpen, J. A. M. Schrauwen, and G. J. Wullems, “Regulation of flavonol biosynthesis during anther and pistil development, and during pollen tube growth in *Solanum tuberosum*,” *The Plant Journal*, vol. 11, no. 1, pp. 105–113, 1997.
- [62] J. Dong, S. T. Kim, and E. M. Lord, “Plantacyanin plays a role in reproduction in *Arabidopsis*,” *Plant Physiology*, vol. 138, no. 2, pp. 778–789, 2005.
- [63] S. Kim, J. C. Mollet, J. Dong, K. L. Zhang, S. Y. Park, and E. M. Lord, “Chemocyanin, a small basic protein from the lily stigma, induces pollen tube chemotropism,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 26, pp. 16125–16130, 2003.
- [64] F. L. Deng, L. L. Tu, J. F. Tan, Y. Li, Y. C. Nie, and X. L. Zhang, “GbpPDF1 is involved in cotton fiber initiation via the core cis-element HDZIP2ATATHB2,” *Plant Physiology*, vol. 158, no. 2, pp. 890–904, 2012.
- [65] M. Potocky, M. A. Jones, R. Bezvoda, N. Smirnov, and V. Zarsky, “Reactive oxygen species produced by NADPH oxidase are involved in pollen tube growth,” *New Phytologist*, vol. 174, no. 4, pp. 742–751, 2007.
- [66] Q. H. Duan, D. Kita, E. A. Johnson et al., “Reactive oxygen species mediate pollen tube rupture to release sperm for fertilization in *Arabidopsis*,” *Nature Communications*, vol. 5, article 3129, 2014.
- [67] L. F. Hu, W. Q. Liang, C. S. Yin et al., “Rice MADS3 regulates ROS homeostasis during late anther development,” *The Plant Cell*, vol. 23, no. 2, pp. 515–533, 2011.
- [68] E. Hama, S. Takumi, Y. Ogihara, and K. Murai, “Pistillody is caused by alterations to the class-B MADS-box gene expression pattern in alloplasmic wheats,” *Planta*, vol. 218, no. 5, pp. 712–720, 2004.
- [69] T. Zhao, Z. F. Ni, Y. Dai, Y. Y. Yao, X. L. Nie, and Q. X. Sun, “Characterization and expression of 42 MADS-box genes in wheat (*Triticum aestivum* L.),” *Molecular Genetics and Genomics*, vol. 276, no. 4, pp. 334–350, 2006.
- [70] Z. Y. Peng, X. Zhou, L. C. Li et al., “Arabidopsis hormone database: a comprehensive genetic and phenotypic information database for plant hormone research in *Arabidopsis*,” *Nucleic Acids Research*, vol. 37, Supplement 1, pp. D975–D982, 2009.
- [71] I. Rieu, M. Wolters-Arts, J. Derksen, C. Mariani, and K. Weterings, “Ethylene regulates the timing of anther dehiscence in tobacco,” *Planta*, vol. 217, no. 1, pp. 131–137, 2003.
- [72] X. Tang, A. Gomes, A. Bhatia, and W. R. Woodson, “Pistil-specific and ethylene-regulated expression of 1-aminocyclopropane-1-carboxylate oxidase genes in petunia flowers,” *The Plant Cell*, vol. 6, no. 9, pp. 1227–1239, 1994.
- [73] M. L. Jones, P. B. Larsen, and W. R. Woodson, “Ethylene-regulated expression of a carnation cysteine proteinase during flower petal senescence,” *Plant Molecular Biology*, vol. 28, no. 3, pp. 505–512, 1995.
- [74] N. Firon, E. Pressman, S. Meir, R. Khoury, and L. Altahan, “Ethylene is involved in maintaining tomato (*Solanum lycopersicum*) pollen quality under heat-stress conditions,” *Aob Plants*, vol. 2012, no. 0, article pls024, 2012.
- [75] H. Cheng, S. S. Song, L. T. Xiao et al., “Gibberellin acts through jasmonate to control the expression of MYB21, MYB24, and MYB57 to promote stamen filament growth in *Arabidopsis*,” *PLoS Genetics*, vol. 5, no. 3, article e1000440, 2009.
- [76] S. Ishiguro, A. Kawai-Oda, J. Ueda, I. Nishida, and K. Okada, “The defective in anther dehiscence1 gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis*,” *The Plant Cell*, vol. 13, no. 10, pp. 2191–2209, 2001.
- [77] P. Lou, J. G. Kang, G. Y. Zhang, G. Bonnema, Z. Y. Fang, and X. W. Wang, “Transcript profiling of a dominant male sterile mutant (*Ms-cd1*) in cabbage during flower bud development,” *Plant Science*, vol. 172, no. 1, pp. 111–119, 2007.
- [78] S. Bhattacharya and I. T. Baldwin, “The post-pollination ethylene burst and the continuation of floral advertisement are harbingers of non-random mate selection in *Nicotiana attenuata*,” *The Plant Journal*, vol. 71, no. 4, pp. 587–601, 2012.
- [79] R. Volz, J. Heydlauff, D. Ripper, L. von Lyncker, and R. Gross-Hardt, “Ethylene signaling is required for synergic

- degeneration and the establishment of a pollen tube block,” *Developmental Cell*, vol. 25, no. 3, pp. 310–316, 2013.
- [80] E. R. Valdivia, Y. Wu, L. C. Li, D. J. Cosgrove, and A. G. Stephenson, “A group-1 grass pollen allergen influences the outcome of pollen competition in maize,” *PLoS One*, vol. 2, no. 1, article e154, 2007.
- [81] S. J. Roy, T. L. Holdaway-Clarke, G. R. Hackett, J. G. Kunkel, E. M. Lord, and P. K. Hepler, “Uncoupling secretion and tip growth in lily pollen tubes: evidence for the role of calcium in exocytosis,” *The Plant Journal*, vol. 19, no. 4, pp. 379–386, 1999.
- [82] S. Blackmore, A. H. Wortley, J. J. Skvarla, and J. R. Rowley, “Pollen wall development in flowering plants,” *New Phytologist*, vol. 174, no. 3, pp. 483–498, 2007.
- [83] D. S. Zhang, W. Q. Liang, Z. Yuan et al., “Tapetum degeneration retardation is critical for aliphatic metabolism and gene regulation during rice pollen development,” *Molecular Plant*, vol. 1, no. 4, pp. 599–610, 2008.
- [84] D. B. Zhang and Z. A. Wilson, “Stamen specification and anther development in rice,” *Chinese Science Bulletin*, vol. 54, no. 14, pp. 2342–2353, 2009.
- [85] D. B. Zhang, X. Luo, and L. Zhu, “Cytological analysis and genetic control of rice anther development,” *Journal of Genetics and Genomics*, vol. 38, no. 9, pp. 379–390, 2011.
- [86] D. S. Zhang, W. Q. Liang, C. S. Yin, J. Zong, F. W. Gu, and D. B. Zhang, “OsC6, encoding a lipid transfer protein, is required for postmeiotic anther development in rice,” *Plant Physiology*, vol. 154, no. 1, pp. 149–162, 2010.
- [87] A. Bernard, F. Domergue, S. Pascal et al., “Reconstitution of plant alkane biosynthesis in yeast demonstrates that *Arabidopsis eceriferum1* and *eciriferum3* are core components of a very-long-chain alkane synthesis complex,” *The Plant Cell*, vol. 24, no. 7, pp. 3106–3118, 2012.
- [88] V. F. Souza, M. S. Pagliarini, C. B. Valle, N. C. Bione, M. U. Menon, and A. B. Mendes-Bonato, “Meiotic behavior of *Brachiaria decumbens* hybrids,” *Genetics and Molecular Research*, vol. 14, no. 4, pp. 12855–12865, 2015.
- [89] O. Rowland and F. Domergue, “Plant fatty acyl reductases: enzymes generating fatty alcohols for protective layers with potential for industrial applications,” *Plant Science*, vol. 193–194, pp. 28–38, 2012.
- [90] M. G. Aarts, R. Hodge, K. Kalantidis et al., “The *Arabidopsis male sterility 2* protein shares similarity with reductases in elongation/condensation complexes,” *The Plant Journal*, vol. 12, no. 3, pp. 615–623, 1997.
- [91] A. A. Dobritsa, S. I. Nishikawa, D. Preuss et al., “LAP3, a novel plant protein required for pollen development, is essential for proper exine formation,” *Sexual Plant Reproduction*, vol. 22, no. 3, pp. 167–177, 2009.
- [92] Z. L. Mou, X. Q. Wang, Z. M. Fu et al., “Silencing of phosphoethanolamine N-methyltransferase results in temperature-sensitive male sterility and salt hypersensitivity in *Arabidopsis*,” *The Plant Cell*, vol. 14, no. 9, pp. 2031–2043, 2002.
- [93] M. A. Zaidi, S. O’Leary, S. B. Wu et al., “A molecular and proteomic investigation of proteins rapidly released from triticale pollen upon hydration,” *Plant Molecular Biology*, vol. 79, no. 1–2, pp. 101–121, 2012.
- [94] A. Baker, I. A. Graham, M. Holdsworth, S. M. Smith, and F. L. Theodoulou, “Chewing the fat:  $\beta$ -oxidation in signalling and development,” *Trends in Plant Science*, vol. 11, no. 3, pp. 124–132, 2006.
- [95] S. F. Yang and N. E. Hoffman, “Ethylene biosynthesis and its regulation in higher plants,” *Annual Review of Plant Physiology*, vol. 35, no. 1, pp. 155–189, 1984.
- [96] A. J. Afzal, J. H. Kim, and D. Mackey, “The role of NOI-domain containing proteins in plant immune signaling,” *BMC Genomics*, vol. 14, no. 1, p. 327, 2013.