The Plant Cell, Vol. 31: 974–992, May 2019, www.plantcell.org © 2019 ASPB.

LARGE-SCALE BIOLOGY

High Temporal-Resolution Transcriptome Landscape of Early Maize Seed Development[OPEN]

Fei Yi,[a,1] Wei Gu,[a,b,1] Jian Chen,[a] Ning Song,[a] Xiang Gao,[a] Xiangbo Zhang,[a] Yingsi Zhou,[a] Xuxu Ma,[a] Weibin Song,[a] Haiming Zhao,[a] Eddi Esteban,[c] Asher Pasha,[c] Nicholas J. Provart,[c] and Jinsheng Lai[a,d,2]

[a] State Key Laboratory of Agrobiotechnology and National Maize Improvement Center, Department of Plant Genetics and Breeding, China Agricultural University, Beijing, 100193, China

[b] China Specialty Maize Research Center (CIMMYT), Crop Breeding and Cultivation Research Institute, Shanghai Academy of Agricultural Sciences, Shanghai, 201403, China

[c] Department of Cell and Systems Biology/Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Ontario M5S 3B2, Canada

[d] Center for Crop Functional Genomics and Molecular Breeding, China Agricultural University, Beijing, 100193, China

ORCID IDs: 0000-0001-5329-8195 (F.Y.); 0000-0003-1340-8403 (W.G.); 0000-0002-1764-8394 (J.C.); 0000-0002-4348-0871 (N.S.); 0000-0003-3882-2598 (X.G.); 0000-0003-0775-0851 (X.Z.); 0000-0002-6541-2116 (Y.Z.); 0000-0003-3992-1811 (X.M.); 0000-0003-3496-2585 (W.S.); 0000-0003-1600-9619 (H.Z.); 0000-0001-9016-9202 (E.E.); 0000-0002-9315-0520 (A.P.); 0000-0001-5551-7232 (N.J.P.); 0000-0001-9202-9641 (J.L.)

The early maize (*Zea mays*) seed undergoes several developmental stages after double fertilization to become fully differentiated within a short period of time, but the genetic control of this highly dynamic and complex developmental process remains largely unknown. Here, we report a high temporal-resolution investigation of transcriptomes using 31 samples collected at an interval of 4 or 6 h within the first six days of seed development. These time-course transcriptomes were clearly separated into four distinct groups corresponding to the stages of double fertilization, coenocyte formation, cellularization, and differentiation. A total of 22,790 expressed genes including 1415 transcription factors (TFs) were detected in early stages of maize seed development. In particular, 1093 genes including 110 TFs were specifically expressed in the seed and displayed high temporal specificity by expressing only in particular period of early seed development. There were 160, 22, 112, and 569 seed-specific genes predominantly expressed in the first 16 h after pollination, coenocyte formation, cellularization, and differentiation stage, respectively. In addition, network analysis predicted 31,256 interactions among 1317 TFs and 14,540 genes. The high temporal-resolution transcriptome atlas reported here provides an important resource for future functional study to unravel the genetic control of seed development.

INTRODUCTION

Maize (*Zea mays*) seed is one of the most important sources of food, feed, and biofuel materials (Godfray et al., 2010), and it serves as an excellent model for research on seed development due to its relatively large size. Maize seed development is initiated in the embryo sac with the fusion of the two pollen sperms with the egg cell and central cells of the female gametophyte to product the progenitors of embryo and endosperm, respectively (Dumas and Mogensen, 1993; Chaudhury et al., 2001). The embryo sac is embedded in the nucellus, which will be gradually degraded after double fertilization. Nucellus degeneration is important for endosperm expansion, and its products are believed to be taken up by endosperm (Russell, 1979; Greenwood et al., 2005). After double fertilization, the zygote undergoes an asymmetric division to form a small apical cell and a large basal cell, which develop into the embryo proper and the suspensor, respectively (Nardmann and Werr, 2009). The embryo proper further forms the mature embryo after the morphogenesis stage and will grow to be the next plant generation (Nardmann and Werr, 2009). The development of endosperm begins with the formation of a coenocyte, which the primary endosperm undergoes several rounds of nuclear divisions but without cytokinesis. The coenocyte then undergoes cellularization and cell differentiation (Lopes and Larkins, 1993; Olsen, 2001; Sabelli and Larkins, 2009; Leroux et al., 2014). After differentiation, the endosperm enlarges significantly through further cell division, cell expansion, and endoreduplication. Different from dicots (in which the endosperm is mostly consumed or absorbed by the developing embryo), maize endosperm serves as a storage tissue to store the proteins and carbohydrates needed for seedling development (Lopes and Larkins, 1993; Berger, 1999; Olsen, 2001; Sabelli and Larkins, 2009).

Understanding the spatial and temporal gene expressional profile along seed development is helpful for the genetic

[1] These authors contributed equally to this work.

[2] Address correspondence to jlai@cau.edu.cn.

The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Jinsheng Lai (jlai@cau.edu.cn).

[OPEN]Articles can be viewed without a subscription.

www.plantcell.org/cgi/doi/10.1105/tpc.18.00961

IN A NUTSHELL

Background: Maize seed is an important source of food, feed and biofuel materials. The early maize seed undergoes several developmental stages after double fertilization to become fully differentiated within a short period of time, but the genetic control of this highly dynamic and complex developmental processes remains largely unknown. Understanding the spatial and temporal gene expressional profile along seed development is useful for unraveling the genetic control of seed development and thus for the genetic improvement of this important crop.

Question: We wanted to know the gene activity dynamic during double fertilization, coenocyte formation, cellularization, and differentiation, four main stages of early maize seed development, especially, which genes are specifically expressed at particular stages of early maize seed development.

Findings: A total of 22,790 expressed genes including 1,415 transcription factors (TFs) were detected in early stages of maize seed development. In particular, 1,093 genes including 110 TFs were specifically expressed in the seed, most of which were newly identified in this study and displayed high temporal specificity by expressing only in particular period of early seed development. There were 160, 22, 112 and 569 seed-specific genes predominantly expressed in the first 16 hours after pollination, coenocyte formation, cellularization and differentiation stage, respectively. In addition, network analysis predicted 31,256 interactions among 1,317 TFs and 14,540 genes. The high-temporal-resolution transcriptome atlas reported here provides an important resource for future functional study to dissect the genetic control of seed development.

Next steps: We plan to select some key genes for CRISPR/Cas9-based gene editing to further explore their function in the genetic control of seed development.

improvement of this important crop. Over the years, several transcriptome profiling studies have been conducted to detect the expressed genes and cellular processes for seed development in Arabidopsis (*Arabidopsis thaliana*; Le et al., 2010; Belmonte et al., 2013), rice (*Oryza sativa*; Xu et al., 2012; Gao et al., 2013), *Tropaeolum majus* (Jensen et al., 2012), and soybean (*Glycine max*; Jones and Vodkin, 2013). In maize, the transcriptome of endosperm was initially characterized using expressional sequence tag sequencing method (Lai et al., 2004). The dynamic of gene expression during seed development was then investigated by a microarray-based approach, which identified 3445 genes with differential expression among samples of six different time points (Liu et al., 2008). The general transcriptome-wide differences between embryo and endosperm had also been analyzed in maize seed 9 days after pollination (DAP) using the RNA sequencing (RNA-seq) method (Lu et al., 2013). Then more detail transcriptome atlas of maize seed development were generated using RNA-seq data from embryo, endosperm, and intact seed sampled at an interval of 2 days from 0 DAP to 38 DAP, which provides an extensive view of transcriptome dynamics over seed development (Chen et al., 2014). To gain the information of spatial distribution of genes in endosperm, a laser-capture microdissection study was reported at 8 DAP, which allowed the identification of a number of compartment specifically expressed genes in the endosperm of this particularly stage (Zhan et al., 2015). Recently, the transcriptomes of isolated mature female and male gametes, 12 and 24 hours after pollination (HAP) zygote, and apical and basal daughter cells were also obtained (Chen et al., 2017).

In maize, the double fertilization events typically finish in the first DAP (Sabelli and Larkins, 2009), with an average of 8 HAP (Chen et al., 2017). The coenocytic stage of maize endosperm usually occurs during 1 to 2 DAP, and then it is followed by a period of cellularization at ~3 to 4 DAP (Sabelli and Larkins, 2009; Leroux et al., 2014). The endosperm cell differentiation starts at ~5 DAP, forming four main cell types: starchy endosperm, aleurone, embryo-surrounding region, and basal endosperm transfer layer (BETL; Olsen, 2001; Sabelli and Larkins, 2009; Leroux et al., 2014). In line with the rapid transition of these developmental stages, large numbers of genes are involved in the key steps of double fertilization, coenocyte formation, cellularization, and differentiation that happen in the first few days of seed development, but these genes may not have been captured in the above-mentioned extensive transcriptome studies. For instance, *embryo sac1* (*ES1*) and *embryo sac4* (*ES4*), two genes encoding secreted peptides and required for micropylar pollen tube guidance and burst, are only expressed in the nucellus during the first few HAP and then show low or even no expression a few hours later (Cordts et al., 2001; Chen et al., 2017). Therefore, it is highly possible that many genes that are important for the early seed development (but are expressed only in a short period of time or in particular developmental stages) have not been identified yet, due to the fact that the previous transcriptome studies did not have sufficient temporal resolution.

Here we report a comprehensive high temporal-resolution investigation of transcriptomes using data for 31 time points, at 4- or 6-h intervals within the first 6 days of maize seed development. This high-density, time-course transcriptome analysis clearly highlighted the timings of double fertilization, coenocyte formation, cellularization, and differentiation in the endosperm. In total, 22,790 genes, including 1415 transcription factors (TFs), were found to be expressed during early maize seed development. These genes were classed into 18 coexpression modules according their expression patterns, which provided further insight into the dynamics of transcriptome reprogramming underlying the developmental and physiological transitions of the four distinct development stages. A total of 1093 genes, including 110 TFs, specifically expressed in seed, were identified; and most of these seed-specific genes had high temporal specificity, being expressed only in a particular period of time within the first six days of maize seed development. TF regulatory network analysis predicted 31,256 interactions among 1317 TFs and 14,540 seed-expressed genes. The high temporal-resolution transcriptomes

presented here provide a valuable resource for the study of seed biology.

RESULTS

The Generation of High Temporal-Resolution Transcriptome Data at Early Stages of Maize Seed Development

To investigate the gene activity dynamic during early maize seed development, we performed RNA-seq for the nucellus (embryo sac included) of inbred line B73 from 0 ~ 144 HAP, with an interval of 4 h (0 to 72 HAP) or 6 h (72 to 144 HAP; Figure 1; Supplemental Figure 1). Two biological replicates, which each were pooled samples from at least three plants, were set up for all 31 time points. Totally, 2.85 billion high-quality reads were generated using the Illumina sequencing platform, and then mapped to the maize B73 reference genome (RefGen_V4; Jiao et al., 2017), using Hisat (Kim et al., 2015). An average ~93% of reads were uniquely mapped (Supplemental Table 1) and only the uniquely mapped reads were further used to calculate the normalized gene expression level as fragments per kilobase of transcript per million mapped reads (FPKM). Comparison of the two biological replicates showed that the expression values between them were highly correlated (average $R^2 = 0.94$). Hence, we took the average FPKM value of the two replicates as the expression level for the sample at each time point. To reduce the influence of transcription noise, we defined a gene as expressed if its FPKM value was ≥1. In total, 22,790 genes including 1415 TFs were found to be expressed in at least one of the 31 samples (Supplemental Data Sets 1 and 2).

To further validate the quality of the gene activity profiles we obtained, we specifically examined the expression patterns of eight genes for which transcript levels were previously reported during early maize seed development. *ZmMCM3*, *ZmMCM6*, *ZmCYC1*, and *ZmCYC3* are genes involved in the cell cycle process and were shown to be induced after fertilization (Sauter et al., 1998; Dresselhaus et al., 1999, 2016). The expression of *ZmMCM3*, *ZmMCM6*, *ZmCYC1*, and *ZmCYC3* are induced in the zygote at 12 and 24 HAP; and *ZmCYC1* and *ZmCYC6* reached highest expression later than that of *ZmMCM3* and *ZmMCM6* (Chen et al., 2017). Here, we found the expression of these four genes began to increase at 8 HAP; *ZmMCM3* and *ZmMCM6* showed the highest expression around 20 HAP, and *ZmCYC1* and *ZmCYC3* showed the highest expression around 32 HAP (Supplemental Figure 2). In addition, *Esr2*, a gene specially expressed in ESR (Bonello et al., 2000), and *Betl10*, a gene related to the differentiation of BETL (Zhan et al., 2015), were expressed after 102 HAP (Supplemental Figure 2), consistent with the idea that endosperm differentiation usually happens at ~4 ~6 DAP (Sabelli and Larkins, 2009). We also found that *ZmSWEET4C*, a hexose transporter gene predominantly expressed in BETL (Sosso et al., 2015), was highly expressed after 102 HAP; and that *ZmYUC1*, an auxin biosynthesis gene (Bernardi et al., 2012; Doll et al., 2017), was rapidly activated after 126 HAP (Supplemental Figure 2), similar to their expression patterns reported previously in Li et al. (2014) and . In summary, the expression dynamics of these genes are in line with previous reports, indicating the high quality and reliability of our data.

High Temporal-Resolution Transcriptomes Can Be Clustered into Four Groups Corresponding to Different Developmental Stages

To gain insight into the transcriptome dynamic of early maize seed development, we performed hierarchical clustering (Figure 2A) and principal component analysis (PCA; Figure 2B) for the 31 time-series samples. In line with the previously reported timing of double fertilization, coenocyte formation, cellularization, and differentiation stages for early maize seed development (Olsen, 2001; Sabelli and Larkins, 2009; Leroux et al., 2014; Chen et al., 2017), these high-density time series transcriptomes can be generally divided into four groups, with each group corresponding to a specific developmental stage (Figure 2C).

Samples from earliest time points (0 to 16 HAP) formed the first cluster and represented the stage around double fertilization

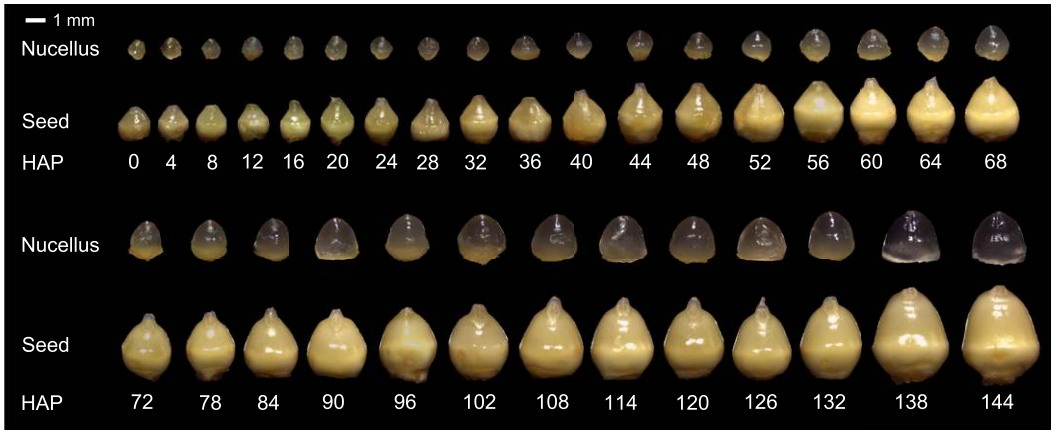

Figure 1. Changes in the Maize Nucellus and Seed from 0 to 144 h After Pollination (HAP).

The nucellus (included embryo sac) samples from 31 different time points were used for transcriptome analysis.

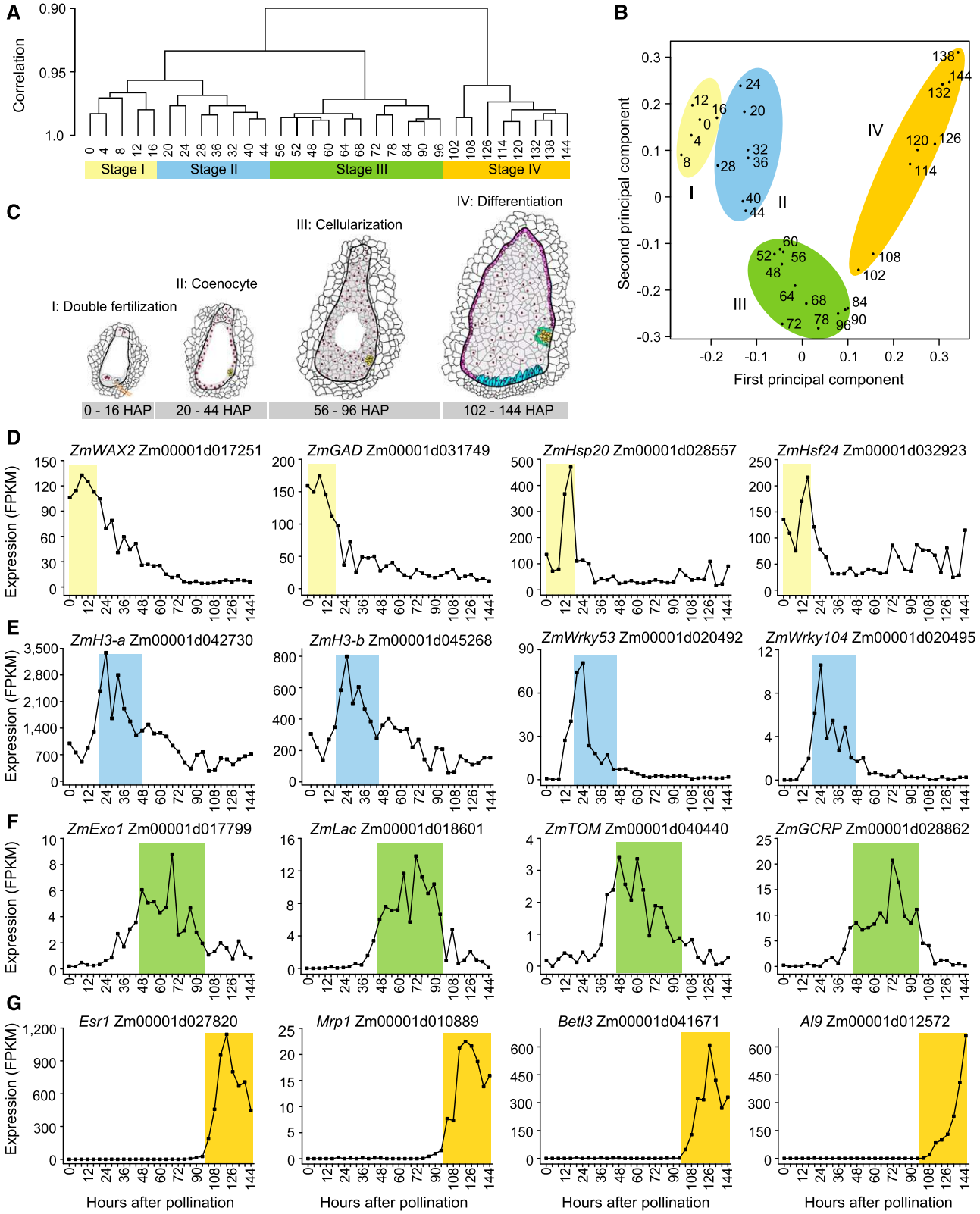

Figure 2. Transcriptome Relationships Among 31 Time Points of Early Maize Seed Development.

(A) Cluster dendrogram showing four distinct development stages: around double fertilization, coenocyte, cellularization, and differentiation.

(Stage I). *WAX2* encodes secreted peptides relating to pollen fertility as reported in Arabidopsis and cucumber (*Cucumis sativus*; Chen et al., 2003; Wang et al., 2015). G*lutamate decarboxylase protein* encodes a non-protein amino acid that plays an important role in pollen tube growth and guidance (Akama and Takaiwa, 2007; Jin et al., 2016). In line with the process of double fertilization, we found both *ZmWAX2* and *Zm* G*lutamate decarboxylase protein* were highly expressed at this stage but were low or even not expressed in later time points (Figure 2D). It was reported that heat shock protein (HSP) and heat shock transcription factor (HSF) are involved in the regulation of reproductive system development, germ cell development, and fertilization in mice (*Mus musculus*) and humans (*Homo sapiens*; Le Masson et al., 2011; Nixon et al., 2017). Here we found that *ZmHSP20* and *ZmHSF24* displayed increased expression after 8 HAP, but rapidly decreased after 16 HAP (Figure 2D), which suggested that *ZmHSP20* and *ZmHSF24* might be important around fertilization in maize.

The samples between time points 20 HAP and 44 HAP formed a second cluster and represented the stage of coenocyte formation in endosperm (Stage II). In this stage, the initial triploid nucleus undergoes several rounds of synchronous division in the absence of cell wall formation and cytokinesis, resulting in the formation of a coenocytic endosperm. As reported in many organisms, canonical H3 genes are expressed during S-stage of the cell cycle and are DNA replication dependent (Ahmad and Henikoff, 2002; Cui et al., 2006; Hamiche and Shuaib, 2013; Otero et al., 2014). According to highly active DNA replication at the coenocytic stage, two canonical H3 genes, *ZmH3-a* (Zm00001d042730) and *ZmH3-b* (Zm00001d045268), showed predominant expression at the coenocytic stage in maize (Figure 2E). *WRKY10* in Arabidopsis is a regulator of seed size and is expressed in the developing endosperm from the two-nuclei stage at ~12 HAP, to endosperm cellularization at ~96 h (Luo et al., 2005). Here we found that its homologous genes in maize, *ZmWRKY53* and *ZmWRKY104*, were mainly expressed at the coenocytic stage (Figure 2E), which suggests that *ZmWRKY53* and *ZmWRKY104* might be important for endosperm proliferation in maize.

Samples between 48 HAP to 96 HAP fell into the third cluster, which corresponds to the cellularization stage (Stage III). *ZmExo1* (Zm00001d017799) encodes an RNA exonuclease. The RNA exonuclease is required for mitotic cell division in *Schizosaccharomyces pombe* (Snee et al., 2016). Collaborative control of cell cycle progression by the RNA exonuclease protein is conserved across species (Snee et al., 2016). *ZmLac* (Zm00001d018601) encodes a laccase that contributes to cell-wall reconstitution in regenerating protoplasts of higher plants (Mayer and Staples, 2002). As reported in rice, the laccase gene *OsLac* could affect grain yield (Zhang et al., 2013). *ZmTOM* (Zm00001d040440) encodes a translocase of the outer mitochondrial membrane (TOM), which can transport mitochondrial precursor proteins (Wiedemann et al., 2003). Previous reports showed that TOM plays an important role in regulation of the cell cycle (Westermann, 2010; Harbauer et al., 2014). *ZmGCRP* (Zm00001d028862) encodes a Gly- and Cys-rich family protein precursor (GCRP). The GCRP proteins play crucial roles in cell to cell signaling and participate in cell division and proliferation in rice (Westermann, 2010; Harbauer et al., 2014). Consistent with the active cell division and cell wall formation that occurs during the cellularization stage, we found these four genes were mainly expressed at this period (Figure 2F).

The fourth cluster was from 102 HAP to 144 HAP, which corresponds to the initial stage of differentiation in the endosperm (Stage IV). *Esr1* is an endosperm-specific gene expressed in a restricted region around embryo and might be involved in the establishment of a physical barrier between embryo and endosperm (Westermann, 2010; Harbauer et al., 2014). *Myeloblastosis* (*MYB*)*-related protein 1* (*MRP1*) and *Betl3* are two BETL-specific genes important for the development and differentiation of BETL (Hueros et al., 1999; Gómez et al., 2009; Zhan et al., 2015). *Al9* is a gene related to aleurone (AL) differentiation (Gómez et al., 2009). We found all these four genes showed a rapidly increased expression at Stage IV (Figure 2G), indicating that this stage is typically by the initiation of endosperm differentiation. In summary, our results demonstrated that our high temporal-resolution transcriptome data are powerful for the stage-specific genes, and that the dynamic transcriptome during the early endosperm development can be separated into four distinct groups corresponding to four different developmental stages.

Gene Expression at Different Developmental Stages of Early Maize Seed

The global hierarchical clustering and PCA analysis graphically display the four main developmental stages of early maize seed. To further provide insights into the functional transitions along early seed development, we clustered all 22,790 expressed genes, including 1415 (6.2%) TFs (Supplemental Data Sets 1 and 2), into 18 coexpression modules using the k-means clustering algorithm, and then performed MapMan annotation to assign genes to functional categories for each module (Figure 3; Supplemental Figure 3)—of which, genes that belong to the first nine modules were mainly expressed at only one of the four developmental stages and represented the particular functions for their corresponding stages (Figure 3A).

Figure 2. (continued).

(B) PCA of the transcriptomes of the 31 time point samples.

(C) Graphic representation of the embryo sac in the four distinct development stages of seed. The pollen tube is shown in orange. Sperm nuclei are shown in dark blue. Polar nuclei and endosperm nuclei are shown in red. The egg cell and embryo cell are shown in yellow. The basal endosperm transfer layer cell, aleurone cell, and embryo-surrounding region cell are shown in light blue, purplish red, and green, respectively.

(D) to **(G)** The marker genes mainly expressed in the stages of around double fertilization **(D)**, coenocyte **(E)**, cellularization **(F),** and differentiation **(G)**. The time points belong to the stage of around double fertilization, coenocyte, cellularization, and differentiation are shown in light yellow, blue, green, and deep yellow, respectively.

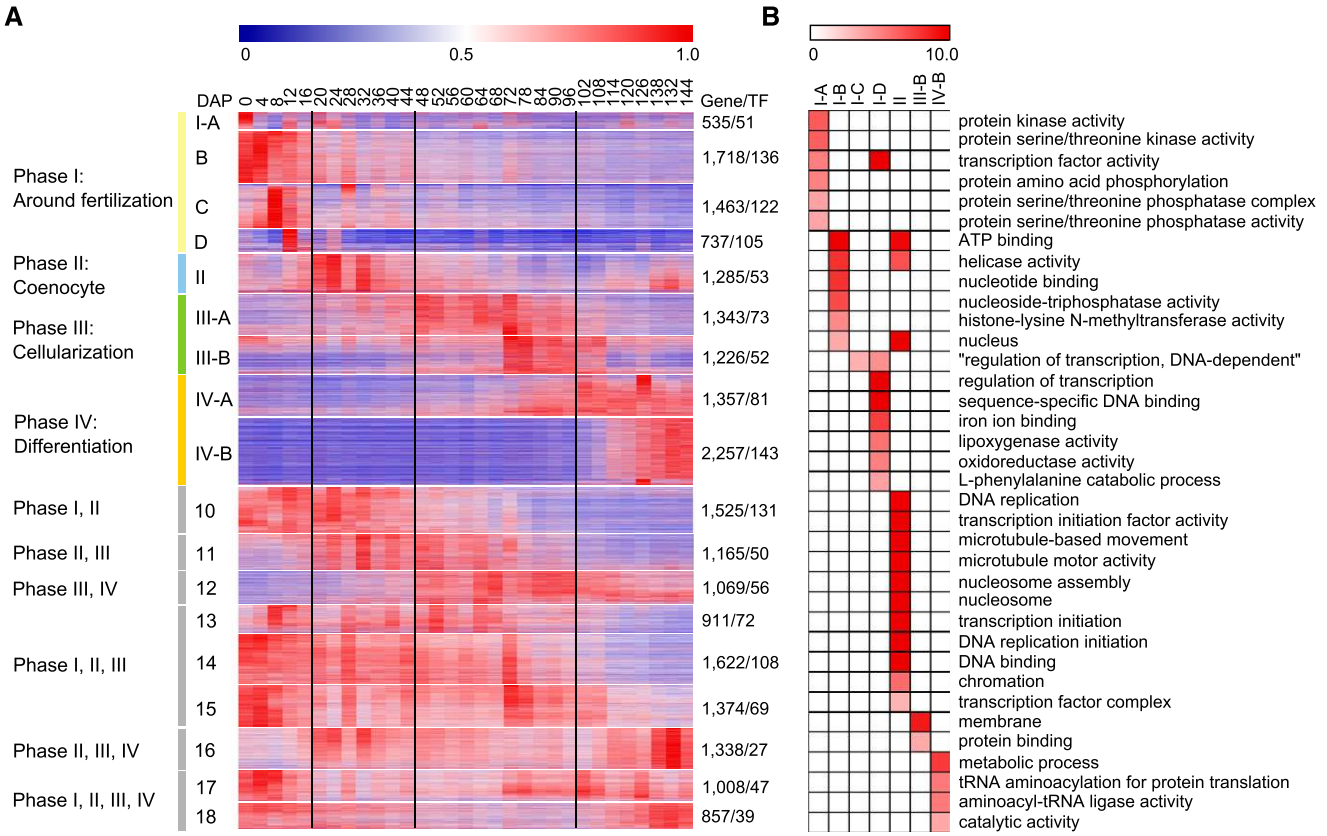

Figure 3. Gene Expression Pattern and Functional Transition Over the Time Course.

(A) Expression patterns of genes in different coexpression modules. For each gene, the FPKM value normalized by the maximum value of all FPKM values of the gene over all time points is shown. The number of genes and TFs in each module are shown on the right.
(B) MapMan functional categories enriched in different coexpression modules. Only significant categories (FDR < 0.05) are displayed.

Genes Expressed around Double Fertilization (Stage I)

The stage around double fertilization (0 to 16 HAP) is best represented by 4453 expressed genes, including 414 TFs, in modules I–A to I–D (Figure 3A; Supplemental Data Set 1). The module I–A (535 genes, 51 TFs) contains a set of genes related to protein serine/threonine kinase activity and amino acid phosphorylation (Figure 3B). These genes might be involved in the initial pollination response, because they mainly expressed at 0 to 4 HAP. The module I–B (1718 genes, 136 TFs) corresponds to genes that were highly expressed at 0 to 12 HAP and then appeared to be low or not expressed at later time points (Figure 3A). As reported previously, fertilization occurs at ~8 HAP on average (Chen et al., 2017). During fertilization, the pollen tube extends to the embryo sac to release sperm to form the zygote (Faure et al., 2003; Luo et al., 2005). Zygotes need energy (ATP) and thus produce energy-rich metabolites for generating ATP (Labarca and Loewus, 1973; Rounds et al., 2011; Obermeyer et al., 2013). Therefore, the genes in module I–B might contribute to the growth of pollen tube, because they were overrepresented by genes involved in ATP binding, helicase activity, nucleotide binding, and nucleoside triphosphatase activity (Figure 3B). For example, *ES1* and *ES4* in module I–B (Supplemental Data Set 1) relate to the pollen tube growth arrest and burst (Cordts et al., 2001).

Ca^{2+} signaling is thought to play important roles in plant growth and development, including key aspects of pollen tube growth and fertilization (Schiøtt et al., 2004; Dresselhaus et al., 2016). In module I–B, we found there were 17 genes involved in the calcium signaling pathway (Supplemental Data Set 1), including Zm00001d031543. Its homolog in Arabidopsis, *Ca^{2+}-ATPASES 9*, functions in the pollen tube plasma membrane and is a key regulator of pollen tube growth and fertilization (Schiøtt et al., 2004).

The genes in module I–C (1463 genes, 122 TFs) were highly expressed at 8 ~ 16 HAP and represented by genes related to transcription regulation (Figure 3B). MYB TFs are involved in controlling various processes such as responses to abiotic and biotic stresses, differentiation, development, metabolism, and defense (Yanhui et al., 2006; Ambawat et al., 2013). We found seven MYB TFs (*MYB3*, *MYB32*, *MYB38*, *MYB81*, *MYB112*, *MYB126*, and *MYB133*) and five MYB-related TFs (*MYBR26*, *MYBR51*, *MYBR78*, *MYBR84*, and *MYBR90*) in module I–C, reflecting the important role of MYB TFs in early seed development. Moreover, nine important plant-specific GARS TFs (*GRAS3*, *GRAS27*, *GRAS34*, *GRAS39*, *GRAS50*, *GRAS61*, *GRAS82*, *GRAS83*, and *GRAS84*) and ten ethylene responsive *APATELA2*

(*AP2*)-*ethylene-responsive element binding protein* (*EREBP*) TFs (*EREB8*, *EREB96*, *EREB117*, *EREB131*, *EREB156*, *EREB158*, *EREB159*, *EREB162*, *EREB192*, and *EREB201*) were also found in module I-C.

Interestingly, we found that in module I–D, there were 737 genes, including 105 TFs, only highly expressed at ~12 HAP (Figure 3A). To further confirm the expression patterns of genes in module I–D, we analyzed the published RNA-seq data of isolated maize gametes, zygotes, and apical and basal daughter cells (Chen et al., 2017). For 512 genes detected in Chen's samples (Chen et al., 2017), ~75% were induced in zygote formation, with high expression at 12 HAP but low expression at 24 HAP (Supplemental Figure 4, Supplemental Data Set 3), indicating the genes in module I–D we identified were indeed transitorily activated after fertilization. The genes in module I–D were enriched in TF activity, transcription regulation, and sequence-specific DNA binding (Figure 3B), suggesting most of which might be key to regulatory function at the beginning of new generation formation. For example, *lipoxygenase-1* (*LOX1*) in module I–D is important for regulation of defense-related signaling molecules and activation of the antioxidative enzyme system (Cho et al., 2012). We also found there were three mitogen-activated protein kinase (MPKs) genes, *MPK1*, *MPKKK11* and *MPKKK18*, in module I–D (Supplemental Data Set 1). It was shown that MPK cascades function as molecular switches in response to spatiotemporal-specific ligand–receptor interactions and the availability of downstream substrates, and are ubiquitous signaling modules in eukaryotes (Widmann et al., 1999; MAPK Group, 2002; Xu and Zhang, 2015).

Genes Expressed during the Coenocyte Formation Stage (Stage II)

The stage of coenocyte formation (20 to 44 HAP) is best represented by 1285 genes, including 53 TFs, in module II (Figure 3A, Supplemental Data Set 1). Consistent with the active chromatin formation and nuclear division that occurs at the coenocytic stage, module II was overrepresented by genes related to DNA replication, transcription initiation factor activity, microtubule-based movement, microtubule motor activity, nucleosome assembly, nucleosome transcription initiation, DNA replication initiation, and DNA binding (Figure 3B). Histones, the major protein components of chromatin, are highly conserved among eukaryotes (Ingouff and Berger, 2010). Based on the reported histone sequences in Arabidopsis (Okada et al., 2005; Ingouff and Berger, 2010; Yelagandula, 2014; Kawashima et al., 2015), we identified a total of 79 histone genes in the maize genome, and we found 66 of these were expressed in our data, including 6 *H1*, 4 *H2A*, 6 *H2A.W*, 3 *H2A.X*, 5 *H2A.Z*, 13 *H2B*, 12 *H3*, 5 *H3.3*, and 12 *H4* (Supplemental Data Set 4**).** Of these 66 expressed histone genes, 71% (47) belonged to module II, including 3 *H1*, 2 *H2A*, 5 *H2A.W*, 1 *H2A.X*, 1 *H2A.Z*, 12 *H2B*, 12 *H3*, and 11 *H4* (Figure 4; Supplemental Data Set 4). Notably, all 12 expressed *H3* were in module II, with predominant expression during the coenocyte stage; and all 5 expressed *H3.3* did not show up in module II. This result was consistent with the reports that canonical H3 deposition is coupled to DNA synthesis during replication and repair, which is extremely activated for coenocyte formation; whereas H3.3 is deposited independently of replication (Ahmad and Henikoff, 2002; Cui et al., 2006). A number of previous works show that different histone H2A variants have distinct functions in diverse biological processes (Talbert and Henikoff, 2010, 2014; Weber and Henikoff, 2014; Kawashima et al., 2015). H2A.W is specific to seed-bearing plants and predominantly localizes in heterochromatin to promote heterochromatin condensation (Yelagandula et al., 2014). We found five of six expressed *H2A.W* genes were in module II. By contrast, for 5 expressed *H2A.Z*, only one *H2A.Z* (Zm00001d027760) was in module II, and the remaining 4 *H2A.Z* were distributed in four different modules (I–C,11,16, and 18; Supplemental Data Set 4). The high variation of expression patterns for different *H2A.Z* genes is in line with the diverse functions of H2A.Z variants, including DNA repair, apparently contradictory roles in gene activation and silencing, nucleosome turnover, heterochromatin, boundary element, and chromatin fiber formation (Zlatanova and Thakar, 2008; Altaf et al., 2009; Dai et al., 2017; Domaschenz et al., 2017).

Genes Expressed during the Cellularization Stage (Stage III)

The stage of endosperm cellularization (48 to 96 HAP) is best represented by 2569 expressed genes, including 125 TFs, in modules III-A and III-B (Figure 3A;Supplemental Data Set 1). The genes in module III–A (1343 genes, 73 TFs) were highly expressed during the entire cellularization stage (48 to 96 HAP; Figure 3A). We found there were 14 auxin pathway genes (Hagen and Guilfoyle, 2002; Yue et al., 2015; Chen et al., 2017) in module III–A, including three ATP Binding Cassette Subfamily B members (ZmABCBs: *ZmABCB11*, *ZmABCB30*, and *ZmABCB35*) representing potential auxin transporter genes, three auxin-responsive factors (ZmARFs: *ZmARF2*, *ZmARF7*, and *ZmARF15*), four genes encoding proteins that interact with ARF regulators (Zm indole-3-acetic acids [IAAs]: *ZmIAA2*, *ZmIAA7*, *ZmIAA15*, and *ZmIAA23*), and four auxin response genes (ZmSAURs—small auxin up RNAs: *ZmSAUR4*, *ZmSAUR22*, *ZmSAUR31*, and *ZmSAUR56*). These results reflected the importance of auxin transporter and response genes in endosperm cellularization. The genes in module III–B (1226 genes, 52 TFs) were mainly expressed at the late cellularization stage (72 to 96 HAP). Genes related to membrane and protein binding were enriched in III–B, which might be involved in the formation of cell membrane during cellularization (Figure 3B). For example, syntaxins are membrane proteins involved in vesicle trafficking and release of neurotransmitters (Burgess et al., 1997; Besserer et al., 2012; Jung et al., 2012). In maize, the syntaxin protein SYP121 could selectively regulate plasma membrane aquaporin trafficking (Besserer et al., 2012). Here we found four syntaxin genes—which we named *ZmSYP121a* (Zm00001d020187), *ZmSYP121b* (Zm00001d041716), *ZmSYP121c* (Zm00001d042018), and *ZmSYP121d* (Zm00001d048147)—were expressed in stage III–B, suggesting they might be involved in cell membrane formation during cellularization.

Genes Expressed during the Differentiation Stage (Stage IV)

The stage of initial endosperm differentiation (102 ~ 144 HAP) is best represented by 3614 genes, including 224 TFs, in module IV–A and IV–B (Figure 3A; Supplemental Data Set 1). The genes in

module IV–A (1357 genes, 81 TFs) were highly expressed from ~102 HAP (Figure 3A), while genes in module IV–B (2257 genes, 143 TFs) were highly expressed from ~114 HAP. These genes might be mainly involved in differentiation. Many genes specifically expressed in endosperm subregions were included in these two modules, including two AL marker genes, *VPP1* (IV–A; Wisniewski and Rogowsky, 2004) and *Al9* (IV–B; Gómez et al., 2009); four ES-specific genes, *Esr1* (IV–B), *Esr2* (IV–B), *Esr3* (IV–-B), and *Esr6* (IV–B; Opsahl-Ferstad et al., 1997; Bonello et al., 2000; Magnard et al., 2000; Balandín et al., 2005; Zhan et al., 2015;

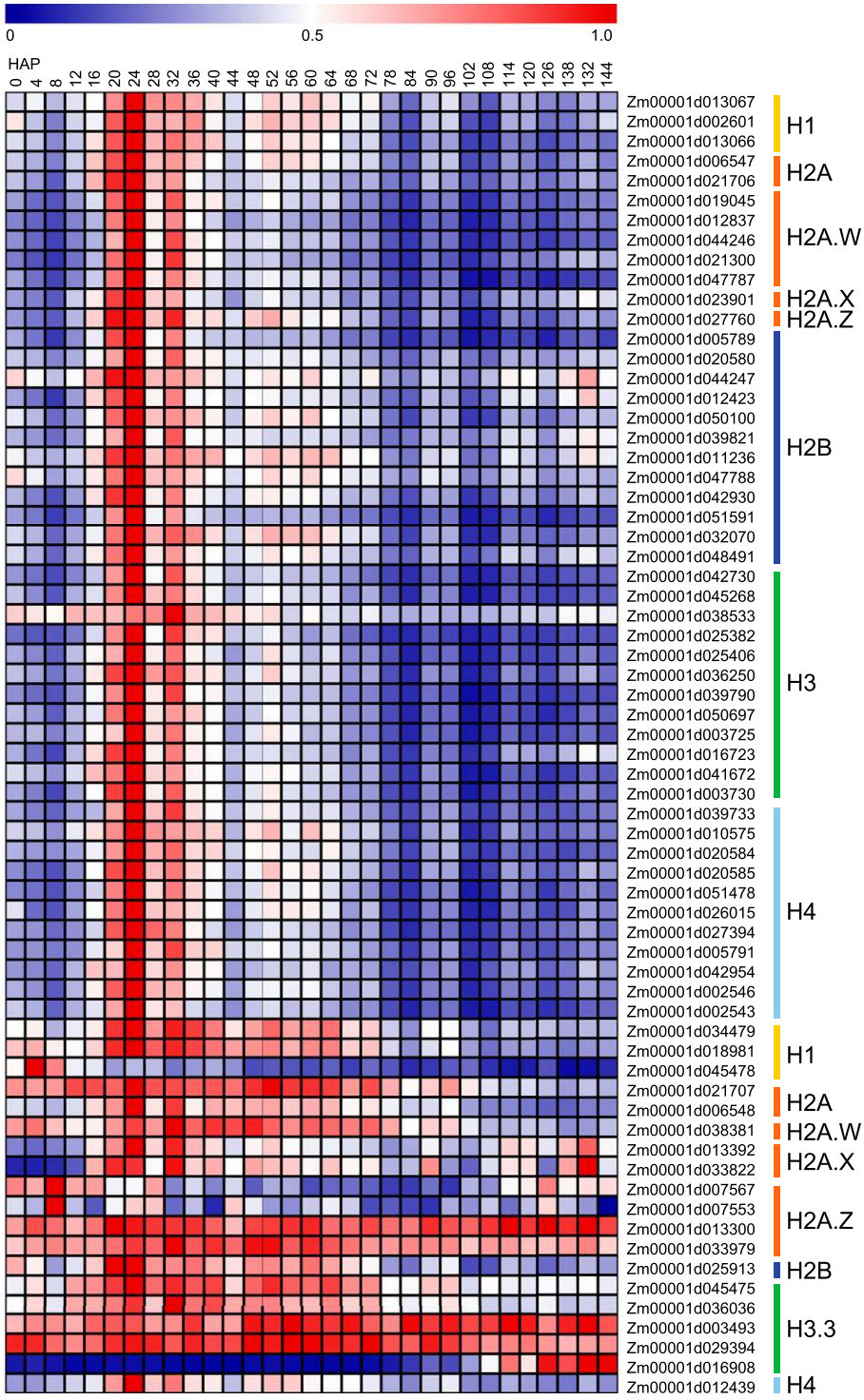

Figure 4. Expression Pattern of Major Histone Protein Genes in 31 Time Points Samples.

Todorow et al., 2018); and 8 BETL-specific genes, *Ebe2* (IV–A), *MYBR33* (IV–B), *MRP1* (IV–B), *Betl-3*, *9*, *10* (IV–B), *Bap2* (IV–B), and *Mn1* (IV–B; Cheng et al., 1996; Hueros et al., 1999; Serna et al., 2001; Magnard et al., 2003; Gómez et al., 2009; Zhan et al., 2015).

Genes Expressed During More than One of the Four Stages

We found that a total of 10,869 genes, including 599 TFs, in modules of M10 ~ M18 were expressed at more than one of the four stages (Figure 3A), indicating that there are some common functional processes in different stages. For example, genes related to microtubule-associated complexes were enriched in M10 (Supplemental Figure 3). They displayed continuous expression at fertilization and coenocytic stages (0 to 44 HAP), which is consistent with the report that microtubule associated complexes are involved in fertilization, mitosis, and cell division (Schatten et al., 1985). In line with the observation that photosynthetic genes are expressed during the development of seed in Arabidopsis (Schmid et al., 2005), here we found genes related to chlorophyll biosynthetic processes were overrepresented in M10 (Supplemental Figure 3), such as the *Chlorophyll synthase G1* (Hunter et al., 2018) and magnesium chelatase gene *Oil yellow 1* (Sawers et al., 2006), both required for chlorophyll biosynthesis. In addition, we found *Geranylgeranyl hydrogenase 1* (Owens et al., 2014), the homolog of rice *Geranylgeranyl reductase* for chlorophyll synthesis (Wang et al., 2014), was also included in M10. These findings implied that the photosynthesis system might start to be established in early development seed in maize. Maize is generally considered as a chilling sensitive species (Miedema, 1982). We found genes related to homoiothermy, ice binding, and response to freezing are enriched in M13, with continuous expression at fertilization, coenocytic, and cellularization stages (0 to 96 HAP; Figure 3A; Supplemental Figure 3). This result suggests that these cold-response genes might function to stabilize membranes against freeze-induced injury and help seeds to develop under suboptimal temperature conditions.

In addition, as shown in M16, 1338 genes including 27 TFs were activated after double fertilization, with continuous expression at stages II to IV. These genes were enriched in many basic functional categories processes, including structural constituent of ribosome, cytoplasm, translation, intracellular, translational elongation, small ribosomal subunit, GTPase activity, RNA binding, protein catabolic process, and vesicle-mediated transport.

Seed-Specific Genes Including TFs and Their Target Genes Involved in the Four Stages of Early Maize Seed Development

The high temporal-resolution transcriptome profiling data generated here provided us with a good opportunity to identify genes specifically expressed in particular stages of early seed development, which is highly informative for inferring gene function and understanding the genetic control of developmental transition. Combined with 23 published non-seed transcriptome data sets, including root, shoot, shoot apical meristem, leaf, cob, tassel, and immature ear (Jia et al., 2009; Wang et al., 2009; Li et al., 2010; Davidson et al., 2011; Bolduc et al., 2012), we identified a total of 1093 genes specifically expressed in seed (Supplemental Figure 5; Supplemental Data Set 5). Of these, 654 genes were not found in our previous transcriptomic study in maize seed (Chen et al., 2014) and in a recent study based on the extensive transcriptomes of 79 B73 tissues (Hoopes et al., 2019). This result is likely because for the first 6 DAP of maize seed development, the intact seed samples of only five and three time points were analyzed by Chen et al. (2014) and Hoopes et al. (2019), respectively. However, the transcriptome data of 31 time points generated in this study enabled us to identify additional early seed-specific genes, especially those specifically expressed in a short period of time. The seed-specific genes identified here accounted for ~6. 20% of all expressed genes detected. By contrast, there were 1415 TFs detected as expression in our data, 10.1% (110) of which are seed-specific TFs (Supplemental Data Set 5). The seed-specific genes and TFs were significantly enriched in the stage of differentiation (after 102 HAP; Supplemental Figure 6), reflecting the important role of seed-specific genes and TFs for the generation of new tissue or cell types during differentiation.

We inferred the gene regulatory network (GRN) that connects TFs with their potential target genes using the method reported previously in Faith et al. (2007) and (Xiong et al., 2017). The GRN is scale free, based on the frequency distribution of the degree of the nodes (Supplemental Figure 7). In total, 14,540 target genes (including 1317 TFs) were included in our newly constructed GRN, with a total of 31,256 interactions (Supplemental Data Set 7). Next, we focused on the network communities connected with seed-specific TFs. In total, 81 seed-specific TFs and 1483 potential target genes were identified, with a total of 2000 interactions. As expected, seed-specific genes were more likely to be regulated by seed-specific TFs as compared with non-seed-specific genes. Of the 513 seed-specific genes in the GRN, 40% (205) displayed a total of 407 interactions with 57 seed-specific TFs. By contrast, for 13,675 non-seed–specific genes in the GRN, only 9.3% (1278) were interacted with 79 seed-specific TFs, resulting in a total of 1593 interactions.

Next, combined with the transcriptome data generated previously in Chen et al. (2014), we further explored the expression patterns of 1093 seed-specific genes that we identified in the later development stage of embryo and endosperm (Figure 5; Table 1).

Specific Genes Expressed around Double Fertilization (Stage I)

The 160 seed-specific genes, including 18 TFs, in group I were mainly expressed around double fertilization (0 to 16 HAP; Figure 5; Supplemental Data Set 5); and in this group, genes related to transcription factor activity, enzyme inhibitor activity, sequence-specific DNA binding, pectin esterase activity, and hydrolase activity are overrepresented (Supplemental Figure 8). The high expression of genes related to enzyme inhibitor activity at fertilization stage might be associated with the protection of female reproductive cells from a variety of biotic stresses, including enzymes from nucellus or synergid lysate, contents of burst pollen tubes, and pathogen attack (McInnis et al., 2006). For example, *ES1* and *ES4*, encoding peptides with structural homology to defensins (proteinase inhibitors) in group I, are expressed in the embryo sac and were suppressed after fertilization (Cordts et al., 2001).

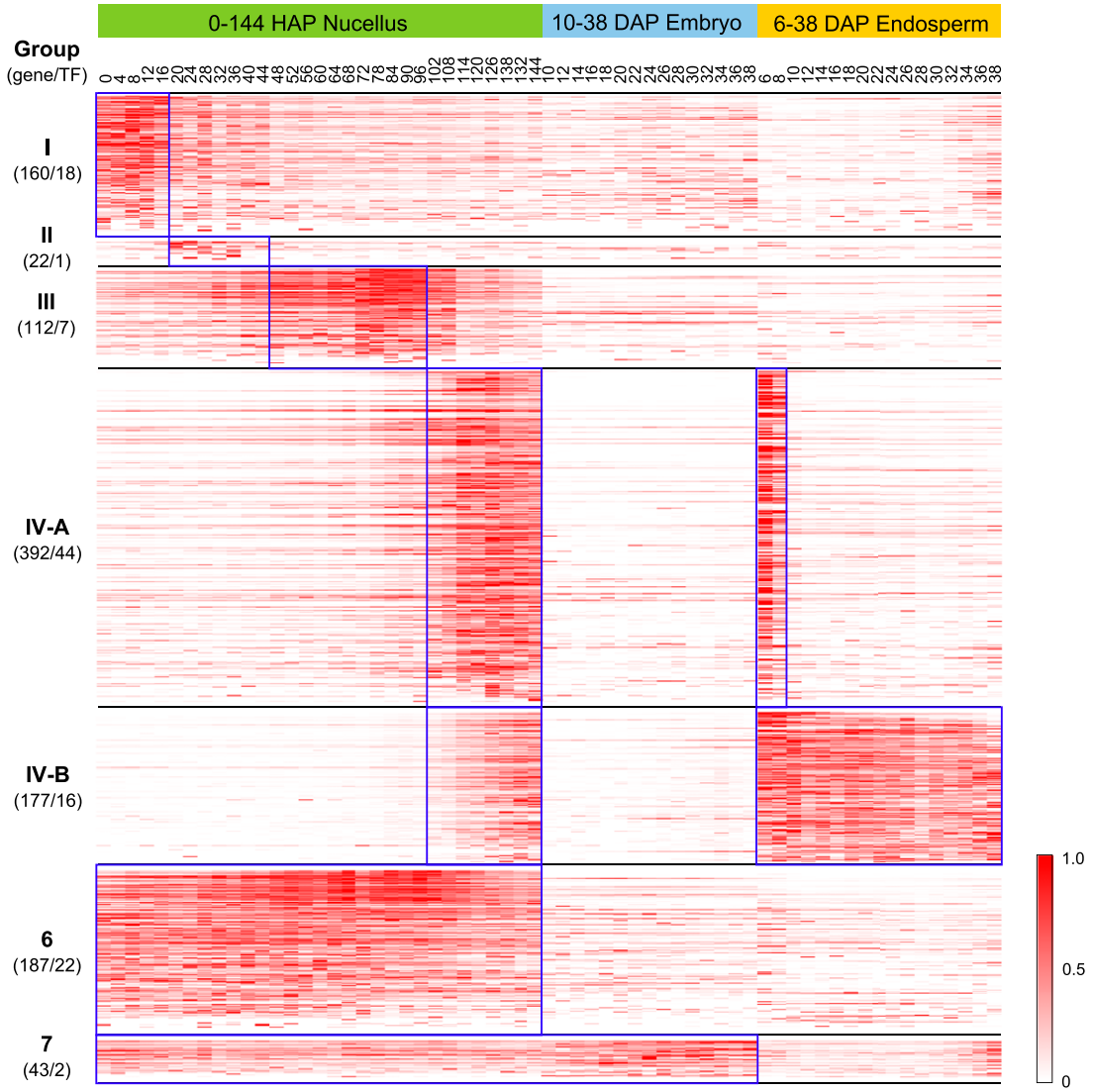

Figure 5. Expression Patterns of Seed-Specific Genes.

Analysis of the expression patterns of seed-specific based on the RNA-seq data of nucellus generated in this study, and the RNA-seq data of embryo and endosperm generated previously (Chen et al., 2014). For each gene is shown the FPKM value normalized by the maximum value of all FPKM values of the gene over all the samples used for analysis. The number of genes and TFs in each group are shown on the left.

We found that 13 seed-specific TFs in group I were included in the GRN, with a total of 148 interactions. The top 5 seed-specific TFs with the most connections in group I are *HSF24*, *HSF20*, *EREB117*, *Basic Leucine Zipper 29* (*BZIP*29), and *GRAS61* (Supplemental Figure 9; Supplemental Data Set 8). *HSF24* and *HSF20* were predicted to interact with 32 and 24 genes, respectively. Notably, *HSF24* and *HSF20* were also highly expressed and ranked as the first and tenth highly expressed seed-specific genes in stage I (0 to 16 HAP; Supplemental Data Set 5). In mice, HSF1 is the major regulator for heat shock transcriptional response, and HSF1-deficient mice exhibited complex phenotypes, including developmental defects and complete female infertility (Le Masson et al., 2011). Here, our results suggested that HSF24 and HSF20 might be the major regulators of heat shock transcriptional response and might play important roles around double fertilization in maize seed. The APATELA2 family TF gene *EREB117* was predicted to interact with 20 genes. Its homolog, *WRINKLED1* (*WRI1*) in Arabidopsis, is involved in the control of storage compound biosynthesis, the mutant of which has a wrinkled seed phenotype (Hanano et al., 2018). *BZIP29* was predicted to interact with 13 genes, including three known genes (*EMP10*, *RH4*, and *Pco090181a*) and two known TFs (*GRAS61* and *C2H2*). Of these, the *EMP10* encodes a mitochondrial PPR protein and is required for embryogenesis and endosperm development in maize (Cai et al., 2017b).

Table 1. The Number of Coexpressed Genes and Seed-Specific Genes Detected in Different Development Stages of Early Maize Seed

Developmental Stage	No. of Genes/TFs	No. of Specific Genes/TFs
Around double fertilization (0 ~ 16 HAP)	4,453/414	160/18
Coenocyte (20 ~ 44 HAP)	1,285/53	22/1
Cellularization (48 ~ 96 HAP)	2,569/125	112/7
Differentiation (102 ~ 144 HAP)	3,614/224	569/60
Other[a]	10,869/599	230/24
Total	22,790/1,415	1,093/110

[a]The genes were expressed at more than one of the four stages.

GRAS are plant-specific TFs and play important roles in many processes such as signal transduction, stress responses, and meristem maintenance (Bolle, 2004; Mayrose et al., 2006; Zhang et al., 2018). At present, only a few GRAS proteins have been characterized in maize, such as ZmGRAS20, which was specifically expressed in endosperm and involved in regulating starch biosynthesis (Cai et al., 2017a). Our results showed that *GRAS61* was mainly expressed around double fertilization and was predicted to interact with 11 genes, including *ES4*, an embryo-sac–specific gene playing an important role in fertilization process (Cordts et al., 2001).

Specific Genes Expressed during Coenocyte Formation Stage (Stage II)

Group II contains 22 seed-specific genes mainly expressed during coenocyte formation (20 to 44 HAP; Figure 5; Supplemental Data Set 5). *BURP domain-containing protein-RD22-like9* (*BURP9*) was included in this group. The BURP domain-containing gene family is a large plant-specific gene family, yet their functions are very poorly understood, especially in maize. *BURP9* was reported to respond to ABA and cold (Gan et al., 2011). Here, we found *BURP9* was mainly expressed during coenocyte formation, which will be helpful for further understanding its function in maize.

ZmCoenocyte4 (ZmCoe4), the only seed-specific TF with predominant expression during coenocyte formation, encodes a WRKY family TF (Supplemental Figure 9; Supplemental Data Set 8). *ZmCoe4* was predicted to interact with 16 genes, including *15 kD zein protein*, *pathogenesis-related protein3*, *glucan endo-1*, *3-beta-glucosidase homolog1*, *chitinase A1*, *umc2348*, and *wound-induced protein1* (Supplemental Data Set 8).

Specific Genes Expressed during the Cellularization Stage (Stage III)

Group III represents 112 seed-specific genes, including 7 TFs, predominantly expressed at the cellularization stage (48 to 96 HAP; Figure 5; Supplemental Data Set 5). *Zea AGAMOUS homolog 2* (*ZAG2*) was included in this group. *ZAG2*, which is homologous to the Arabidopsis floral homeotic gene *AGAMOUS*, is expressed in developing ovules and the inner carpel faces, and it might be important for maize flower development (Schmidt et al., 1993). We found *ZAG2* was highly expressed after pollination, especially at the cellularization stage (48 to 96 HAP), which is consistent with the report in *Orchis italica* that *OitaAG* mRNA levels were high in columns and ovaries (Salemme et al., 2013), particularly after pollination. However, up to now no function has yet been determined for ZAG2.

The top five seed-specific TFs with the most connections in group III are *MYB131*, *MYB16*, *BZIP109*, *ZAG2*, and *BZIP114* (Figure 6A; Supplemental Figure 9; Supplemental Data Set 8). *MYB131*, *MYB16*, and *BZIP109* were predicted to interact with 109, 83, and 51 genes, respectively; some of their potential target genes are overlapped (Figure 6A). There were 18 genes including three known TFs (*MYB8*, *Knox1*, and *GGATA12*) that were predicted to be regulated by MYB16 and BZIP109. There were 17 genes including four previous characterized genes (*Sumo1a*, *Cdpk13*, *Bm1*, and *AY107053*) that were predicted to be regulated by *MYB16* and *MYB131*. In addition, Zm00001d008178, which has a homolog in rice that is an multidrug-resistant-like ABC transporter gene, was predicted to be regulated simultaneously by *MYB16*, *BZIP109*, and *MYB131*. In total, 206 genes, of which 89% (182 genes) were mainly expressed in the cellularization stage, were predicted to interact with *MYB131*, *MYB16*, and/or *BZIP109*, including nine known TFs (*MYB8*, *MYB130*, *MYB23*, *HB20*, *HB64*, *HB84*, *DOF42*, *Knox1*, and *GGATA12*), one unreported ERF TF (Zm00001d016535), and 18 seed-specific genes. The closely related community formed by these genes might be important for cellularization. *ZAG2* was predicted to interact with 50 genes, including *6-phosphogluconate dehydrogenase1*, *adenine phosphoribosyl transferase1*, *aldehyde dehydrogenase2*, and *SBP-domain protein4*. *BZIP11* was also predicted to interact with 50 genes, including *cytochrome c reductase1*, *rotten ear3*, and *Homeobox-tf 28* (Supplemental Figure 9; Supplemental Data Set 8).

Specific Genes Expressed during Differentiation Stage (Stage IV)

About half of seed-specific genes (52%, 569/1093) identified were highly expressed during endosperm differentiation (102 ~ 144 HAP), and these can be further divided into two sub-groups (Figure 5; Supplemental Data Set 5) by combining their expression patterns during the whole development stage of seed (Chen et al., 2014). There are 392 genes including 44 TFs in group IV–A, which displayed high expression in the 6 and 8 DAP endosperm but low or no expression in later endosperm (Figure 5). Of these genes in group IV–A, 51.8% (203/392; Supplemental Data Set 6) were identified as subregion-specific genes in seed according to the RNA-seq results of laser-capture microdissection samples collected at 8 DAP (Zhan et al., 2015). About 77% (157/203) of these are specifically expressed in defined compartments of

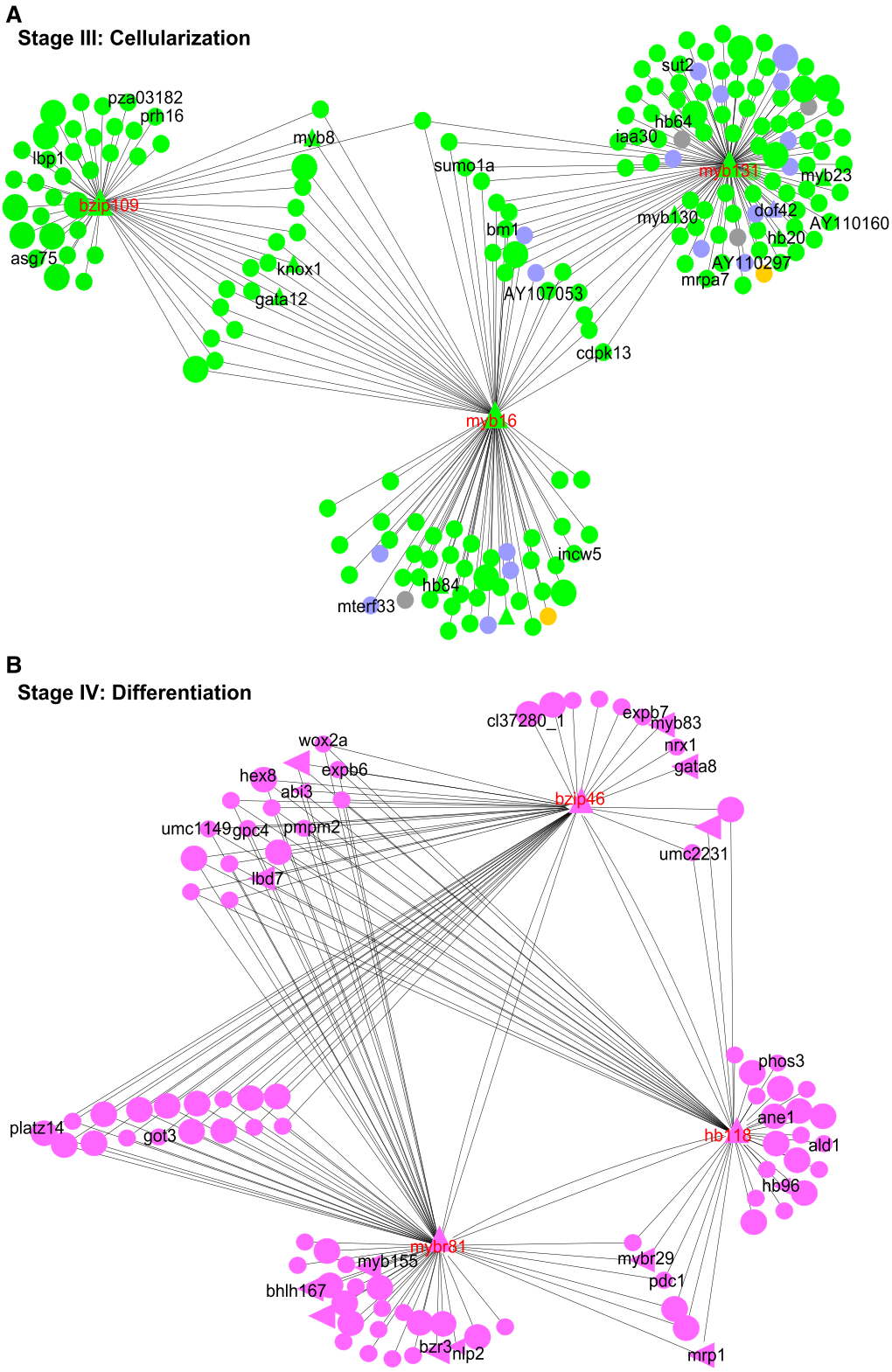

Figure 6. Network Hubs Regulating Genes in Different Seed Development Stages.

(A) Network hubs (*MYB131*, *MYB16,* and *BZIP109*) regulating genes in cellularization stage.

endosperm, including well-known ESR-specific genes (*Esr1*, *Esr2*, *Esr3*, *Esr6*, *Meg14*, and *Male Sterile8*), AL-specific genes (*Al9*, *WOX2b*, *Cadtfr12, Cadtfr14*, and *Sbt1*), and BETL-specific genes (*Bap2*, *Betl-3*, *9*, *10*, *Ebe2*, *Meg6*, *Meg13*, *MYBR81*, and *MRP1*; Supplemental Data Set 6). These results suggest that the genes in group IV–A might play an important role in endosperm cell types differentiation.

Interestingly, we found defense-response–related genes were overrepresented in group IV–A. For example, *Esr6* in IV–A is a defensin gene specifically expressed in ESR that plays a protective role (Balandín et al., 2005). Some of the BETL-specific genes in IV–A, including *Ebe*, *Betl3*, and *Bap2* in IV–A were also suggested to help defend against pathogen entry into the developing seed (Magnard et al., 2003; Barrero et al., 2006). Exploring the function of these seed-specific defense-response–related genes might be useful for understanding the establishment of the defense response mechanism in new differentiated tissues in early endosperm development.

The top five seed-specific TFs with the most connections in group IV are *ARF17*, *MYBR81*, *MYB80*, *BZIP46*, and *HB118*, which were predicted to interact with 73, 67, 53, 48, and 48 genes, respectively (Figure 6B; Supplemental Figure 9; Supplemental Data Set 8). These five TFs are all in group IV–A. As shown in Figure 6B, we found *MYBR81*, *HB118*, and *BZIP46* and their interacting genes formed a closely related community. There were 17 genes, including 8 previously characterized genes (*UMC1149*, *IBD7*, *GPC4*, *HEX8*, *ABI3*, *WOX2*a, *PMPM2*, and *EXBP6*) that were predicted to be simultaneously regulated by *MYBR81*, *HB118*, and *BZIP46*.

In total, 97 genes were predicted to interact with *MYBR81*, *HB118*, and *BZIP46*, which were all mainly expressed in the differentiation stage. About half of these genes (48%, 47/97) were seed-specific genes, including 11 seed-specific TFs (*LBD7*, *BZR3*, *MYBR29*, *MRP1*, *BHLH167*, *GATA8*, *MYB155*, *MYB83*, *WOX2a*, and two unreported TFs). Notably, *MYBR81* and *HB118* are two BETL-specific expression TFs (Zhan et al., 2015). Moreover, we found 13 of their interacting genes were identified as BETL-expressed genes, including five annotated genes (*HB96*, *MYBR29*, *MRP1*, *EXPB7*, and *HEX8*)—of which *MRP1*, an important BETL regulator (Gómez et al., 2009), and *MYBR29* were predicted to be simultaneously regulated by *MYBR81* and *HB118*. In summary, these results suggest the important role of TFs such as *MYBR81*, *BZIP46*, and *HB118* in endosperm differentiation, especially in BETL differentiation.

Compared with genes in group IV–A, genes in group IV–B (177 genes including 16 TFs) were continuously expressed from initial differentiation to endosperm maturation (Figure 5; Supplemental Data Set 5), which suggested that genes in group IV–B might be mainly involved in specific biological processes (such as grain filling) in well-differentiated endosperm that happen over a prolonged period of time. For example, *ZmSWEET4C* in IV–B encodes a hexose transporter that transfers cell wall invertase-derived hexoses into or across the BETL, a key step in seed filling. Indeed, seed filling is defective for the mutants of both *ZmSWEET4c* and its rice ortholog *OsSWEET4*, with a strong empty pericarp phenotype (Sosso et al., 2015). *Floury3* (*FL3*) in IV–B is a maternally expressed imprinted gene, which encodes a plant AT-rich sequence and zinc binding family protein (Li et al., 2017). *FL3* can modulate the biogenesis of tRNAs and 5S rRNA through interactions with RNA polymerase III transcription machinery, which may underlie endosperm development and storage reserve filling (Li et al., 2017). The semidominant negative mutant of *fl3* exhibits severe defects in the endosperm (Li et al., 2017). *Trp aminotransferase related1* (*ZmTar1*) and *ZmYUCCA1* (*ZmYUC1*), two genes involved in different Trp-dependent pathways of IAA biosynthesis, were also found in IV–B. *ZmTar1* is an endosperm-specific paternally expressed imprinted gene critical for the indole-3-pyruvic acid branch (Chourey et al., 2010; Zhang et al., 2011). *ZmYUC1* is an endosperm-specific IAA biosynthesis protein gene critical for the tryptamine branch (Chourey et al., 2010). In the *zmyuc1* mutant, free IAA level is reduced and has ~40% less dry mass as compared with the wild type (Bernardi et al., 2012). Both ZmTar1 and ZmYUC1 are important for highly complex homeostatic control of IAA levels in maize endosperm (Chourey et al., 2010).

There were 13 annotated TFs (*BHLH167*, *MYB155*, *NRP1*, *WOX2a*, *MYB83*, *GATA8*, *DOF36*, *MYB127*, *NAC130*, *SCL1*, *EREB137*, *EREB167*, and *GATA33*) and 3 unknown TFs (Zm00001d006319, Zm00001d040952, and Zm00001d017899) identified in IV–B. The 2 TFs with the most connections in group IV–B are *BHLH167* and *MYB155*, which were both predicted to interact with 40 genes. The *BHLH167* was very recently reported as *Opaque11* (*O11*; Feng et al., 2018). O11 is a major regulator of maize endosperm metabolism, development, and stress response, which thus regulates nutrient metabolism by directly regulating carbohydrate metabolic enzymes and the upstream regulators, including O2 and prolamin box-binding factor (Feng et al., 2018) . The starch and protein accumulation were decreased in *o11*, a classic seed mutant with a small and opaque endosperm (Nelson, 1981). The *MYB155* is reported to be highly expressed in the maize endosperm and involved in the process of starch biosynthesis (Xiao et al., 2017). Collectively, these identified seed-specific genes and TFs in group IV–B might be critical for the grain-filling process, according to the function of the known genes and TFs in this group.

Specific Genes Expressed during More than One of the Four Stages

In total, 230 seed-specific genes, including 24 TFs, were expressed at more than one of the four stages, and they can be

Figure 6. (continued).

(B) Network hubs (*BZIP46*, *MYBR81,* and *HB118*) regulating genes in differentiation stage. Color codes indicate that the gene displayed with the peak expression in the following corresponding stages. Yellow is around double fertilization. Gray is coenocyte. Light green is cellularization. Pink is differentiation. Light blue indicates the genes expressed at more than one of the four stages. Genes are shown as small circles. Seed-specific genes are shown as big circles. Non-seed–specific TFs are shown as small triangles. Seed-specific TFs are shown as big triangles.

further divided into two groups (Figure 5; Supplemental Data Set 5). Genes in group 6 (187 genes, 22 TFs) displayed high expression at 0 to 144 HAP with low or even no expression at the other later stages of seed development (Figure 5). The genes in group 6 were related to cell wall organization, sexual reproduction, and aspartic-type endopeptidase activity. Compared with seed-specific genes in group 6, the genes in group 7 (43 genes including 2 TFs) were found to be expressed at later stages of seed development. The overall expression levels of genes in group 7 in embryo are obviously higher than that in endosperm, implying many of these genes might be mainly involved in the specific biological processes occurring in the embryo. For example, two TFs in group 7, *viviparous-1* (*Vp1*) and *WRI1 transcription factor2*, are reported to play important roles in embryo development. *Vp1* is reported to be highly expressed in maize embryo and controls the anthocyanin pathway by regulating *colored aleurone1* (*C1*). The embryo of *vp1* mutant displays reduced sensitivity to the hormone abscisic acid, resulting in precocious germination, and blocked anthocyanin synthesis in aleurone and embryo tissues (McCarty et al., 1989). *WRI1 transcription factor2* is a key regulator of seed oil biosynthesis in maize. It showed a strong transcriptional induction during the early filling stage of the embryo in maize and could complement the reduced fatty acid content of Arabidopsis *wri1-4* seed (Pouvreau et al., 2011). In summary, the seed-specific genes identified here will be useful to understand the specific biological processes occurring during seed development, especially for the early stages.

DISCUSSION

In this study, we constructed a high temporal-resolution dynamic transcriptome landscape of early maize seed development by sampling 31 time points from 0 to 144 HAP at intervals of 4 h (0 to 72 HAP) or 6 h (72 to 144 HAP). Our tissue samples for transcriptomic analysis contained the nucellus and embryo sac. The nucellus will be degraded gradually after double fertilization (Russell, 1979; Greenwood et al., 2005), while the embryo sac is the area of the initiation of embryo and endosperm development (Dumas and Mogensen, 1993; Chaudhury et al., 2001). As shown by previous morphological observation in Leroux et al. (2014))), the major part of our samples is the nucellus, which should make the largest contribution to the transcriptome data that we generated. However, as the transcriptome reprograming is extremely active in the embryo sac during early seed development (early embryo and endosperm), even though nucellus tissue constitutes a big part of our samples, the dynamic we observed could mostly reflect the activity of earlier endosperm and embryo, particularly for the endosperm tissues (which enlarged much more than the embryo in the earlier stages).

The dynamic transcriptome data provided here clearly demonstrated the four key development stages within the early seed, including the stage around double fertilization, coenocyte formation, as well as cellularization and differentiation of endosperm, which the occurrence times revealed here are consistent with those reported previously (Olsen, 2001; Sabelli and Larkins, 2009; Leroux et al., 2014; Chen et al., 2017). We found there are 4453, 1285, 2569, and 3614 genes mainly expressed at the stages of around double fertilization, coenocyte formation, cellularization, and differentiation, respectively (Table 1), during the early development of maize seed. This large collection of genes provides a rich resource for future functional studies, which will greatly enhance our understanding of the genetic control of early seed development. In particular, we detected 1093 seed-specific genes, including 110 TFs, which will no doubt be the targets of future functional genomics studies. For example, through the GRN analysis, our results suggested that the seed-specific TFs *MYB131*, *MYB16*, and *BZIP109* might be critical for endosperm cellularization, and *MYBR81*, *BZIP46*, and *HB118* might play a key role in endosperm differentiation, which together with their predicted target genes formed a closely related community at cellularization and differentiation stages, respectively. Nevertheless, the exact roles of these seed-specific genes in early seed development remain to be determined.

In summary, our data set provides a high temporal-resolution atlas of gene expression during early maize seed development. These data provide a much-needed high-resolution gene expression profiling during all the stages of early seed development. The seed-specific genes (stage-specific genes) and particularly the TF-target genes GRN uncovered here provide a solid foundation in the future for the identification of key players involved in determining each specific cell type of early seed development.

METHODS

Plant Material Collection and RNA Sequencing

The maize (*Zea mays*) inbred line B73 was grown in the field in May of 2016 in Beijing, China, and it was pollinated in July. All the individual plants were self-pollinated at the same time. The nucellus (embryo sac included) was collected by manual dissection, frozen immediately in liquid nitrogen, and stored at −80°C before processing. Two biological replicates were set up for each of the time points. Each replicate was obtained by pooling samples from at least three plants.

Total RNA was extracted using TRIzol reagent (Invitrogen). RNA-seq libraries were constructed according to the manufacturer's protocol of the Vazyme mRNA-seq library preparation kit (Vazyme) and were sequenced to generate 150-nucleotide paired-end reads on an HiSeq platform (Illumina).

Read Mapping and Analysis

The B73 reference genome (RefGen_v4; Jiao et al., 2017) was downloaded from http://ensembl.gramene.org/Zea_mays/Info/Index. After removing low-quality reads using the SolexaQA (V2.5) software (Cox et al., 2010), Illumina sequencing reads were mapped to the B73 reference genome using Hisat2-2.0.4 (Kim et al., 2015) with default settings for parameters. The bam files of uniquely mapped reads were used as inputs for the Cufflinks (V2.2.0) software (Ghosh and Chan, 2016), and FPKM values were calculated to measure the expression levels of genes. We calculated the Pearson correlation coefficient between biological replicates with the normalized expression levels of $\log_2$ (FPKM value +1).

Hierarchical clustering was performed by the MeV (V4.9) software https://sourceforge.net/projects/mev-tm4/files/ with the HCL method. PCA was performed using the prcomp function in R software (R Team, 2013) with default settings to facilitate graphical interpretation of relatedness among 31 different time points samples. The transformed and normalized gene expression values with $\log_2$ (FPKM +1) were used for hierarchical clustering, and the z-scores of the genes were used for the analysis of PCA.

Gene Coexpression and Functional Enrichment Analysis

The MeV (V4.9) software with the *k*-means method was used for coexpression analysis for 31 different time points samples. The normalized expression values of genes were calculated by dividing their expression level at different time points with their maximum observed FPKM. The figure of merit (Yeung et al., 2001) was used to determine the optimal cluster number. Functional category enrichment for each coexpression module was evaluated with the MapMan (v3.6.0) functional annotation (Thimm et al., 2004). Before conducting the MapMan annotation, we chose the longest protein of each gene as a representative protein and ran the Mercator with default settings. Fisher's exact test was used to examine whether the functional categories were overrepresented for a given module. Resulting *P* values were adjusted to *Q* values by the Benjamini–Hochberg correction, and a false discovery rate of 5% was applied.

Identification of Seed-Specific Gene Expression

For identification of seed-specific genes we used 31 different time points seed samples collected here and 23 non-seed RNA-seq data (Jia et al., 2009; Wang et al., 2009; Li et al., 2010; Davidson et al., 2011; Bolduc et al., 2012) downloaded from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). The method we described previously in Chen et al. (2014) was used. Briefly, the expression levels across all of the samples were normalized using $\log_2$ (FPKM +0.01). Then we calculated z-scores of the given gene in different seed samples compared with the non-seed samples using the normalized expression level. The gene was determined to be seed specifically expressed if it had a z-score above 3 in at least one of the seed samples. Then, combining in the transcriptome data that we generated previously (Chen et al., 2014), we further explored the expression patterns of seed-specific genes we identified in later stage of embryo and endosperm development by performing coexpression analysis using the MeV (V4.9) software.

The subregion-specific genes mentioned in this article were identified based on their compartment specificity scores in different subregions of seed reported previously in Zhan et al. (2015). A gene was defined as a subregion-specific gene if its compartment specificity score was larger than 0.5.

GRN Inference

We used the context likelihood of relatedness algorithm method (Faith et al., 2007) to construct a TF-related gene regulatory network. Mutual information (MI) for calculating the expression similarity between the expression levels of TF and gene pairs were calculated by R software (entropy package; R Team, 2013). The context likelihood of relatedness algorithm calculated regulation strength by comparing the MI between a TF and its gene pairs to the background network distribution of MI for all TFs and gene pairs that included one of the TFs and its target. The final formula is $f(Z_i, Z_j) = \mathrm{SQRT}(Z_i^2 + Z_j^2)$, where Z_i^2 is the z-score between gene i and its background genes, and where Z_j is the z-score between gene j and its background genes (Faith et al., 2007). Finally, we set $f(Z_i, Z_j)$ above 4.5 to identify the tightly regulated relationship between all pairs of genes and TFs.

Accession Numbers

The generated raw reads have been uploaded to NCBI's SRA database and are available under the accession number PRJNA505095. RNA-seq data as FPKM values is available through the eFP Browser engine (http://bar.utoronto.ca/efp_maize/cgi-bin/efpWeb.cgi?dataSource=Early_Seed), which "paints" the expression data onto images representing the samples used to generate the RNA-seq data.

Sequence data for the genes mentioned in this article can be obtained from the literature based on the gene list in Supplemental Table 2.

Supplemental Data

Supplemental Figure 1. Sketch of the sampled region.

Supplemental Figure 2. Validation of RNA-seq data with known genes.

Supplemental Figure 3. MapMan functional categories enriched in different coexpression modules of nucellus.

Supplemental Figure 4. Heat map of the expression patterns of genes in module I–D.

Supplemental Figure 5. Expression patterns of seed-specific genes in all nucellus and non-seed samples.

Supplemental Figure 6. Percentage of genes, tissue-specific genes, TFs, and tissue-specific TFs detected in each time point in the nucellus.

Supplemental Figure 7. Frequency distribution of the degree of nodes in the gene regulatory network.

Supplemental Figure 8. MapMan functional categories enriched in different coexpression modules of specific genes.

Supplemental Figure 9. Network hubs regulating genes in different stages.

Supplemental Table 1. Summary of RNA-Seq read mapping results.

Supplemental Table 2. Accession numbers of genes mentioned.

Supplemental Data Set 1. Expression level of genes in different samples.

Supplemental Data Set 2. List of TFs expressed in nucellus samples.

Supplemental Data Set 3. List of genes in module I–D expressed in gametes, zygotes, and early two-celled pro-embryo cells in maize.

Supplemental Data Set 4. Expression level of histone protein genes in each time point of nucellus.

Supplemental Data Set 5. Summary of seed-specific genes.

Supplemental Data Set 6. Detail information of genes in cluster IV–A and IV–B.

Supplemental Data Set 7. All 31,256 edges of the GRN.

Supplemental Data Set 8. Detail information of the top five highly connected network-specific TFs in corresponding stages I to IV.

ACKNOWLEDGMENTS

This work was supported by grants from National Key Research & Development Program (2016YFD0101803), National Natural Science Foundation of China (91735301, 31421005, 31701432) and 948 project (2016-X33), and China Postdoctoral Science Foundation (2017M620074).

AUTHOR CONTRIBUTIONS

J.L., F.Y., and W.G. designed the experiments. F.Y, W.G, N.S., X.G, X.M, H.Z., and W.S. performed the experiments. F.Y., X.Z., Y.Z., and J.C. analyzed the data. E.E., A.P., and N.P. contributed to the RNA-seq data accessibility via the eFP Browser engine. F.Y., W.G., and J.L. wrote the article.

Received December 18, 2018; revised March 6, 2019; accepted March 25, 2019; published March 26, 2019.

REFERENCES

Ahmad, K., and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Mol. Cell **9:** 1191–1200.

Akama, K., and Takaiwa, F. (2007). C-terminal extension of rice glutamate decarboxylase (OsGAD2) functions as an autoinhibitory domain and overexpression of a truncated mutant results in the accumulation of extremely high levels of GABA in plant cells. J. Exp. Bot. **58:** 2699–2707.

Altaf, M., Auger, A., Covic, M., and Côté, J. (2009). Connection between histone H2A variants and chromatin remodeling complexes. Biochem. Cell Biol. **87:** 35–50.

Ambawat, S., Sharma, P., Yadav, N.R., and Yadav, R.C. (2013). MYB transcription factor genes as regulators for plant responses: an overview. Physiol. Mol. Biol. Plants **19:** 307–321.

Balandín, M., Royo, J., Gómez, E., Muniz, L.M., Molina, A., and Hueros, G. (2005). A protective role for the embryo surrounding region of the maize endosperm, as evidenced by the characterisation of *ZmESR-6*, a defensin gene specifically expressed in this region. Plant Mol. Biol. **58:** 269–282.

Barrero, C., Muñiz, L.M., Gómez, E., Hueros, G., and Royo, J. (2006). Molecular dissection of the interaction between the transcriptional activator ZmMRP-1 and the promoter of BETL-1. Plant Mol. Biol. **62:** 655–668.

Belmonte, M.F., et al. (2013). Comprehensive developmental profiles of gene activity in regions and subregions of the *Arabidopsis* seed. Proc. Natl. Acad. Sci. USA **110:** E435–E444.

Berger, F. (1999). Endosperm development. Curr. Opin. Plant Biol. **2:** 28–32.

Bernardi, J., Lanubile, A., Li, Q.B., Kumar, D., Kladnik, A., Cook, S.D., Ross, J.J., Marocco, A., and Chourey, P.S. (2012). Impaired auxin biosynthesis in the *defective endosperm18* mutant is due to mutational loss of expression in the *ZmYuc1* gene encoding endosperm-specific YUCCA1 protein in maize. Plant Physiol. **160:** 1318–1328.

Besserer, A., Burnotte, E., Bienert, G.P., Chevalier, A.S., Errachid, A., Grefen, C., Blatt, M.R., and Chaumont, F. (2012). Selective regulation of maize plasma membrane aquaporin trafficking and activity by the SNARE SYP121. Plant Cell **24:** 3463–3481.

Bolduc, N., Yilmaz, A., Mejia-Guerra, M.K., Morohashi, K., O'Connor, D., Grotewold, E., and Hake, S. (2012). Unraveling the KNOTTED1 regulatory network in maize meristems. Genes Dev. **26:** 1685–1690.

Bolle, C. (2004). The role of GRAS proteins in plant signal transduction and development. Planta **218:** 683–692.

Bonello, J.F., Opsahl-Ferstad, H.G., Perez, P., Dumas, C., and Rogowsky, P.M. (2000). Esr genes show different levels of expression in the same region of maize endosperm. Gene **246:** 219–227.

Burgess, R.W., Deitcher, D.L., and Schwarz, T.L. (1997). The synaptic protein syntaxin1 is required for cellularization of Drosophila embryos. J. Cell Biol. **138:** 861–875.

Cai, H., Chen, Y., Zhang, M., Cai, R., Cheng, B., Ma, Q., and Zhao, Y. (2017a). A novel GRAS transcription factor, *ZmGRAS20*, regulates starch biosynthesis in rice endosperm. Physiol. Mol. Biol. Plants **23:** 143–154.

Cai, M., Li, S., Sun, F., Sun, Q., Zhao, H., Ren, X., Zhao, Y., Tan, B.C., Zhang, Z., and Qiu, F. (2017b). *Emp10* encodes a mitochondrial PPR protein that affects the cis-splicing of *nad2* intron 1 and seed development in maize. Plant J. **91:** 132–144.

Chaudhury, A.M., Koltunow, A., Payne, T., Luo, M., Tucker, M.R., Dennis, E.S., and Peacock, W.J. (2001). Control of early seed development. Annu. Rev. Cell Dev. Biol. **17:** 677–699.

Chen, J., Zeng, B., Zhang, M., Xie, S., Wang, G., Hauck, A., and Lai, J. (2014). Dynamic transcriptome landscape of maize embryo and endosperm development. Plant Physiol. **166:** 252–264.

Chen, J., Strieder, N., Krohn, N.G., Cyprys, P., Sprunck, S., Engelmann, J.C., and Dresselhaus, T. (2017). Zygotic genome activation occurs shortly after fertilization in maize. Plant Cell **29:** 2106–2125.

Chen, X., Goodwin, S.M., Boroff, V.L., Liu, X., and Jenks, M.A. (2003). Cloning and characterization of the *WAX2* gene of Arabidopsis involved in cuticle membrane and wax production. Plant Cell **15:** 1170–1185.

Cheng, W.H., Taliercio, E.W., and Chourey, P.S. (1996). The *Miniature1* seed locus of maize encodes a cell wall invertase required for normal development of endosperm and maternal cells in the pedicel. Plant Cell **8:** 971–983.

Cho, K., Kim, Y.C., Woo, J.C., Rakwal, R., Agrawal, G.K., Yoeun, S., and Han, O. (2012). Transgenic expression of dual positional maize lipoxygenase-1 leads to the regulation of defense-related signaling molecules and activation of the antioxidative enzyme system in rice. Plant Sci. **185-186:** 238–245.

Chourey, P.S., Li, Q.B., and Kumar, D. (2010). Sugar-hormone cross-talk in seed development: Two redundant pathways of IAA biosynthesis are regulated differentially in the invertase-deficient *miniature1* (*mn1*) seed mutant in maize. Mol. Plant **3:** 1026–1036.

Cordts, S., Bantin, J., Wittich, P.E., Kranz, E., Lörz, H., and Dresselhaus, T. (2001). *ZmES* genes encode peptides with structural homology to defensins and are specifically expressed in the female gametophyte of maize. Plant J. **25:** 103–114.

Cox, M.P., Peterson, D.A., and Biggs, P.J. (2010). SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics **11:** 485.

Cui, B., Liu, Y., and Gorovsky, M.A. (2006). Deposition and function of histone H3 variants in *Tetrahymena thermophila*. Mol. Cell. Biol. **26:** 7719–7730.

Dai, X., Bai, Y., Zhao, L., Dou, X., Liu, Y., Wang, L., Li, Y., Li, W., Hui, Y., Huang, X., Wang, Z., and Qin, Y. (2017). H2A.Z represses gene expression by modulating promoter nucleosome structure and enhancer histone modifications in *Arabidopsis*. Mol. Plant **10:** 1274–1292.

Davidson, R.M., Hansey, C.N., Gowda, M., Childs, K.L., Lin, H., Vaillancourt, B., Sekhon, R.S., de Leon, N., Kaeppler, S.M., Jiang, N., and Robin Buell, C. (2011). Utility of RNA sequencing for analysis of maize reproductive transcriptomes. Plant Genome **3:** 191–203.

Doll, N.M., Depège-Fargeix, N., Rogowsky, P.M., and Widiez, T. (2017). Signaling in early maize kernel development. Mol. Plant **10:** 375–388.

Domaschenz, R., Kurscheid, S., Nekrasov, M., Han, S., and Tremethick, D.J. (2017). The histone variant H2A.Z is a master regulator of the epithelial-mesenchymal transition. Cell Reports **21:** 943–952.

Dresselhaus, T., Cordts, S., Heuer, S., Sauter, M., Lörz, H., and Kranz, E. (1999). Novel ribosomal genes from maize are differentially expressed in the zygotic and somatic cell cycles. Mol. Gen. Genet. **261:** 416–427.

Dresselhaus, T., Sprunck, S., and Wessel, G.M. (2016). Fertilization mechanisms in flowering plants. Curr. Biol. **26:** R125–R139.

Dumas, C., and Mogensen, H.L. (1993). Gametes and fertilization: Maize as a model system for experimental embryogenesis in flowering plants. Plant Cell **5:** 1337–1348.

Faith, J.J., Hayete, B., Thaden, J.T., Mogno, I., Wierzbowski, J., Cottarel, G., Kasif, S., Collins, J.J., and Gardner, T.S. (2007). Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. PLoS Biol. **5:** e8.

Faure, J.E., Rusche, M.L., Thomas, A., Keim, P., Dumas, C., Mogensen, H.L., Rougier, M., and Chaboud, A. (2003). Double fertilization in maize: The two male gametes from a pollen grain have the ability to fuse with egg cells. Plant J. **33:** 1051–1062.

Feng, F., Qi, W., Lv, Y., Yan, S., Xu, L., Yang, W., Yuan, Y., Chen, Y., Zhao, H., and Song, R. (2018). OPAQUE11 is a central hub of the regulatory network for maize endosperm development and nutrient metabolism. Plant Cell **30:** 375–396.

Gan, D., Jiang, H., Zhang, J., Zhao, Y., Zhu, S., and Cheng, B. (2011). Genome-wide analysis of BURP domain-containing genes in maize and sorghum. Mol. Biol. Rep. **38:** 4553–4563.

Gao, Y., Xu, H., Shen, Y., and Wang, J. (2013). Transcriptomic analysis of rice (*Oryza sativa*) endosperm using the RNA-Seq technique. Plant Mol. Biol. **81:** 363–378.

Ghosh, S., and Chan, C.K. (2016). Analysis of RNA-Seq data using TopHat and Cufflinks. Methods Mol. Biol. **1374:** 339–361.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., and Toulmin, C. (2010). Food security: The challenge of feeding 9 billion people. Science **327:** 812–818.

Gómez, E., Royo, J., Muñiz, L.M., Sellam, O., Paul, W., Gerentes, D., Barrero, C., López, M., Perez, P., and Hueros, G. (2009). The maize transcription factor Myb-Related Protein-1 is a key regulator of the differentiation of transfer cells. Plant Cell **21:** 2022–2035.

Greenwood, J.S., Helm, M., and Gietl, C. (2005). Ricinosomes and endosperm transfer cell structure in programmed cell death of the nucellus during *Ricinus* seed development. Proc. Natl. Acad. Sci. USA **102:** 2238–2243.

Hagen, G., and Guilfoyle, T. (2002). Auxin-responsive gene expression: Genes, promoters and regulatory factors. Plant Mol. Biol. **49:** 373–385.

Hamiche, A., and Shuaib, M. (2013). Chaperoning the histone H3 family. Biochim. Biophys. Acta **1819:** 230–237.

Hanano, A., Almousally, I., Shaban, M., and Murphy, D.J. (2018). Arabidopsis plants exposed to dioxin result in a WRINKLED seed phenotype due to 20S proteasomal degradation of WRI1. J. Exp. Bot. **69:** 1781–1794.

Harbauer, A.B., Opalińska, M., Gerbeth, C., Herman, J.S., Rao, S., Schönfisch, B., Guiard, B., Schmidt, O., Pfanner, N., and Meisinger, C. (2014). Mitochondria. Cell cycle-dependent regulation of mitochondrial preprotein translocase. Science **346:** 1109–1113.

Hoopes, G.M., Hamilton, J.P., Wood, J.C., Esteban, E., Pasha, A., Vaillancourt, B., Provart, N.J., and Buell, C.R. (2019). An updated gene atlas for maize reveals organ-specific and stress-induced genes. Plant J. **97:** 1154–1167.

Hueros, G., Royo, J., Maitz, M., Salamini, F., and Thompson, R.D. (1999). Evidence for factors regulating transfer cell-specific expression in maize endosperm. Plant Mol. Biol. **41:** 403–414.

Hunter, C.T., Saunders, J.W., Magallanes-Lundback, M., Christensen, S.A., Willett, D., Stinard, P.S., Li, Q.B., Lee, K., DellaPenna, D., and Koch, K.E. (2018). Maize *w3* disrupts *homogentisate solanesyl transferase* (*ZmHst*) and reveals a plastoquinone-9 independent path for phytoene desaturation and tocopherol accumulation in kernels. Plant Journal **93:** 799–813.

Ingouff, M., and Berger, F. (2010). Histone3 variants in plants. Chromosoma **119:** 27–33.

Jensen, J.K., Schultink, A., Keegstra, K., Wilkerson, C.G., and Pauly, M. (2012). RNA-Seq analysis of developing nasturtium seeds (*Tropaeolum majus*): Identification and characterization of an additional galactosyltransferase involved in xyloglucan biosynthesis. Mol. Plant **5:** 984–992.

Jia, Y., Lisch, D.R., Ohtsu, K., Scanlon, M.J., Nettleton, D., and Schnable, P.S. (2009). Loss of RNA-dependent RNA polymerase 2 (RDR2) function causes widespread and unexpected changes in the expression of transposons, genes, and 24-nt small RNAs. PLoS Genet. **5:** e1000737.

Jiao, Y., et al. (2017). Improved maize reference genome with single-molecule technologies. Nature **546:** 524–527.

Jin, J., et al. (2016) GAD1 encodes a secreted peptide that regulates grain number, grain length, and awn development in rice domestication. Plant Cell **28:** 2453–2463.

Jones, S.I., and Vodkin, L.O. (2013). Using RNA-Seq to profile soybean seed development from fertilization to maturity. PLoS One **8:** e59270.

Jung, J.J., Inamdar, S.M., Tiwari, A., and Choudhury, A. (2012). Regulation of intracellular membrane trafficking and cell dynamics by syntaxin-6. Biosci. Rep. **32:** 383–391.

Kawashima, T., Lorković, Z.J., Nishihama, R., Ishizaki, K., Axelsson, E., Yelagandula, R., Kohchi, T., and Berger, F. (2015). Diversification of histone H2A variants during plant evolution. Trends Plant Sci. **20:** 419–425.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: A fast spliced aligner with low memory requirements. Nat. Methods **12:** 357–360.

Labarca, C., and Loewus, F. (1973). The nutritional role of pistil exudate in pollen tube wall formation in *Lilium longiflorum*: II. Production and utilization of exudate from stigma and stylar canal. Plant Physiol. **52:** 87–92.

Lai, J., Dey, N., Kim, C.S., Bharti, A.K., Rudd, S., Mayer, K.F., Larkins, B.A., Becraft, P., and Messing, J. (2004). Characterization of the maize endosperm transcriptome and its comparison to the rice genome. Genome Res. **14** (10A): 1932–1937.

Le, B.H., et al. (2010). Global analysis of gene activity during *Arabidopsis* seed development and identification of seed-specific transcription factors. Proc. Natl. Acad. Sci. USA **107:** 8063–8070.

Le Masson, F., Razak, Z., Kaigo, M., Audouard, C., Charry, C., Cooke, H., Westwood, J.T., and Christians, E.S. (2011). Identification of heat shock factor 1 molecular and cellular targets during embryonic and adult female meiosis. Mol. Cell. Biol. **31:** 3410–3423.

Leroux, B.M., Goodyke, A.J., Schumacher, K.I., Abbott, C.P., Clore, A.M., Yadegari, R., Larkins, B.A., and Dannenhoffer, J.M. (2014). Maize early endosperm growth and development: From fertilization through cell type differentiation. Am. J. Bot. **101:** 1259–1274.

Li, G., et al. (2014). Temporal patterns of gene expression in developing maize endosperm identified through transcriptome sequencing. Proc. Natl. Acad. Sci. USA **111:** 7582–7587.

Li, P., et al. (2010). The developmental dynamics of the maize leaf transcriptome. Nat. Genet. **42:** 1060–1067.

Li, Q., Wang, J., Ye, J., Zheng, X., Xiang, X., Li, C., Fu, M., Wang, Q., Zhang, Z., and Wu, Y. (2017). The maize imprinted gene *Floury3* encodes a PLATZ protein required for tRNA and 5S rRNA transcription through interaction with RNA polymerase III. Plant Cell **29:** 2661–2675.

Liu, X., Fu, J., Gu, D., Liu, W., Liu, T., Peng, Y., Wang, J., and Wang, G. (2008). Genome-wide analysis of gene expression profiles during the kernel development of maize (*Zea mays* L.). Genomics **91:** 378–387.

Lopes, M.A., and Larkins, B.A. (1993). Endosperm origin, development, and function. Plant Cell **5:** 1383–1399.

Lu, X., Chen, D., Shu, D., Zhang, Z., Wang, W., Klukas, C., Chen, L.L., Fan, Y., Chen, M., and Zhang, C. (2013). The differential transcription network between embryo and endosperm in the early developing maize seed. Plant Physiol. **162:** 440–455.

Luo, M., Dennis, E.S., Berger, F., Peacock, W.J., and Chaudhury, A. (2005). *MINISEED3* (*MINI3*), a *WRKY* family gene, and *HAIKU2* (*IKU2*), a leucine-rich repeat (*LRR*) *KINASE* gene, are regulators of seed size in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **102:** 17531–17536.

Magnard, J.L., Le Deunff, E., Domenech, J., Rogowsky, P.M., Testillano, P.S., Rougier, M., Risueño, M.C., Vergne, P., and Dumas, C. (2000). Genes normally expressed in the endosperm are expressed at early stages of microspore embryogenesis in maize. Plant Mol. Biol. **44:** 559–574.

Magnard, J.L., Lehouque, G., Massonneau, A., Frangne, N., Heckel, T., Gutierrez-Marcos, J.F., Perez, P., Dumas, C., and Rogowsky, P.M. (2003). *ZmEBE* genes show a novel, continuous expression pattern in the central cell before fertilization and in specific domains of the resulting endosperm after fertilization. Plant Mol. Biol. **53:** 821–836.

MAPK Group. (2002). Mitogen-activated protein kinase cascades in plants: A new nomenclature. Trends Plant Sci. **7:** 301–308.

Mayer, A.M., and Staples, R.C. (2002). Laccase: New functions for an old enzyme. Phytochemistry **60:** 551–565.

Mayrose, M., Ekengren, S.K., Melech-Bonfil, S., Martin, G.B., and Sessa, G. (2006). A novel link between tomato GRAS genes, plant disease resistance and mechanical stress response. Mol. Plant Pathol. **7:** 593–604.

McCarty, D.R., Carson, C.B., Stinard, P.S., and Robertson, D.S. (1989). Molecular analysis of *viviparous-1*: An abscisic acid-insensitive mutant of maize. Plant Cell **1:** 523–532.

McInnis, S.M., Desikan, R., Hancock, J.T., and Hiscock, S.J. (2006). Production of reactive oxygen species and reactive nitrogen species by angiosperm stigmas and pollen: Potential signaling crosstalk? New Phytol. **172:** 221–228.

Miedema, P. (1982). The effects of low temperature on *Zea mays*. Adv. Agron. **35:** 93–128.

Nardmann, J., and Werr, W. (2009). Patterning of the maize embryo and the perspective of evolutionary developmental biology. In Handbook of Maize: Its Biology.J.L. Bennetzenand S.C. Hake, eds (New York: Springer), pp. 105–119.

Nelson, O. (1981). The mutations opaque-9 through opaque-13. Maize Genetics Cooperation Newsletter, **55:** 68.

Nixon, B., Bromfield, E.G., Cui, J., and De Iuliis, G.N. (2017). Heat Shock Protein A2 (HSPA2): Regulatory roles in germ cell development and sperm function. Adv. Anat. Embryol. Cell Biol. **222:** 67–93.

Obermeyer, G., Fragner, L., Lang, V., and Weckwerth, W. (2013). Dynamic adaption of metabolic pathways during germination and growth of lily pollen tubes after inhibition of the electron transport chain. Plant Physiol. **162:** 1822–1833.

Okada, T., Endo, M., Singh, M.B., and Bhalla, P.L. (2005). Analysis of the histone H3 gene family in Arabidopsis and identification of the male-gamete-specific variant *AtMGH3*. Plant J. **44:** 557–568.

Olsen, O.-A. (2001). Endosperm development: Cellularization and cell fate specification. Annual Review of Plant Physiology and Plant Molecular Biology. **52:** 233–267.

Opsahl-Ferstad, H.G., Le Deunff, E., Dumas, C., and Rogowsky, P.M. (1997). *ZmEsr*, a novel endosperm-specific gene expressed in a restricted region around the maize embryo. Plant J. **12:** 235–246.

Otero, S., Desvoyes, B., and Gutierrez, C. (2014). Histone H3 dynamics in plant cell cycle and development. Cytogenet. Genome Res. **143:** 114–124.

Owens, B.F., et al. (2014). A foundation for provitamin A biofortification of maize: Genome-wide association and genomic prediction models of carotenoid levels. Genetics **198:** 1699–1716.

Pouvreau, B., Baud, S., Vernoud, V., Morin, V., Py, C., Gendrot, G., Pichon, J.P., Rouster, J., Paul, W., and Rogowsky, P.M. (2011). Duplicate maize *Wrinkled1* transcription factors activate target genes involved in seed oil biosynthesis. Plant Physiol. **156:** 674–686.

Rounds, C.M., Winship, L.J., and Hepler, P.K. (2011). Pollen tube energetics: Respiration, fermentation and the race to the ovule. AoB Plants **2011:** plr019.

R Team. (2013). A language and environment for statistical computing. RA Lang. Environ. Stat. Comput. **55:** 275–286.

Russell, S.D. (1979). Fine structure of megagametophyte development in *Zea mays*. Can. J. Bot. **57:** 1093–1110.

Sabelli, P.A., and Larkins, B.A. (2009). The development of endosperm in grasses. Plant Physiol. **149:** 14–26.

Salemme, M., Sica, M., Gaudio, L., and Aceto, S. (2013). The *OitaAG* and *OitaSTK* genes of the orchid *Orchis italica*: A comparative analysis with other C- and D-class MADS-box genes. Mol. Biol. Rep. **40:** 3523–3535.

Sauter, M., von Wiegen, P., Lörz, H., and Kranz, E. (1998). Cell cycle regulatory genes from maize are differentially controlled during fertilization and first embryonic cell division. Sex. Plant Reprod. **11:** 41–48.

Sawers, R.J.H., Viney, J., Farmer, P.R., Bussey, R.R., Olsefski, G., Anufrikova, K., Hunter, C.N., and Brutnell, T.P. (2006). The maize *Oil yellow1* (*Oy1*) gene encodes the I subunit of magnesium chelatase. Plant Mol. Biol. **60:** 95–106.

Schatten, G., Simerly, C., and Schatten, H. (1985). Microtubule configurations during fertilization, mitosis, and early development in the mouse and the requirement for egg microtubule-mediated motility during mammalian fertilization. Proc. Natl. Acad. Sci. USA **82:** 4152–4156.

Schiøtt, M., Romanowsky, S.M., Baekgaard, L., Jakobsen, M.K., Palmgren, M.G., and Harper, J.F. (2004). A plant plasma membrane Ca^{2+} pump is required for normal pollen tube growth and fertilization. Proc. Natl. Acad. Sci. USA **101:** 9502–9507.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D., and Lohmann, J.U. (2005). A gene expression map of *Arabidopsis thaliana* development. Nat. Genet. **37:** 501–506.

Schmidt, R.J., Veit, B., Mandel, M.A., Mena, M., Hake, S., and Yanofsky, M.F. (1993). Identification and molecular characterization of *ZAG1*, the maize homolog of the Arabidopsis floral homeotic gene AGAMOUS. Plant Cell **5:** 729–737.

Serna, A., Maitz, M., O'Connell, T., Santandrea, G., Thevissen, K., Tienens, K., Hueros, G., Faleri, C., Cai, G., Lottspeich, F., and Thompson, R.D. (2001). Maize endosperm secretes a novel antifungal protein into adjacent maternal tissue. Plant J. **25:** 687–698.

Snee, M.J., Wilson, W.C., Zhu, Y., Chen, S.Y., Wilson, B.A., Kseib, C., O'Neal, J., Mahajan, N., Tomasson, M.H., Arur, S., and Skeath, J.B. (2016). Collaborative control of cell cycle progression by the RNA exonuclease dis3 and Ras is conserved across species. Genetics **203:** 749–762.

Sosso, D., et al. (2015). Seed filling in domesticated maize and rice depends on SWEET-mediated hexose transport. Nat. Genet. **47:** 1489–1493.

Talbert, P.B., and Henikoff, S. (2010). Histone variants—Ancient wrap artists of the epigenome. Nat. Rev. Mol. Cell Biol. **11:** 264–275.

Talbert, P.B., and Henikoff, S. (2014). Environmental responses mediated by histone variants. Trends Cell Biol. **24:** 642–650.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L.A., Rhee, S.Y., and Stitt, M. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J. **37:** 914–939.

Todorow, V., Rahmeh, M., Hofmann, S., Kirn, V., Mahner, S., Jeschke, U., and von Schönfeldt, V. (2018). Correction to: Promotor analysis of ESR1 in endometrial cancer cell lines, endometrial and endometriotic tissue. Arch. Gynecol. Obstet. **298:** 457.

Wang, P., Li, C., Wang, Y., Huang, R., Sun, C., Xu, Z., Zhu, J., Gao, X., Deng, X., and Wang, P. (2014). Identification of a Geranylgeranyl reductase gene for chlorophyll synthesis in rice. Springerplus **3:** 201.

Wang, W., Liu, X., Gai, X., Ren, J., Liu, X., Cai, Y., Wang, Q., and Ren, H. (2015). *Cucumis sativus* L. *WAX2* plays a pivotal role in wax biosynthesis, influencing pollen fertility and plant biotic and abiotic stress responses. Plant Cell Physiol. **56:** 1339–1354.

Wang, X., Elling, A.A., Li, X., Li, N., Peng, Z., He, G., Sun, H., Qi, Y., Liu, X.S., and Deng, X.W. (2009). Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize. Plant Cell **21:** 1053–1069.

Weber, C.M., and Henikoff, S. (2014). Histone variants: Dynamic punctuation in transcription. Genes Dev. **28:** 672–682.

Westermann, B. (2010). Mitochondrial fusion and fission in cell life and death. Nat. Rev. Mol. Cell Biol. **11:** 872–884.

Widmann, C., Gibson, S., Jarpe, M.B., and Johnson, G.L. (1999). Mitogen-activated protein kinase: Conservation of a three-kinase module from yeast to human. Physiol. Rev. **79:** 143–180.

Wiedemann, N., Kozjak, V., Chacinska, A., Schönfisch, B., Rospert, S., Ryan, M.T., Pfanner, N., and Meisinger, C. (2003). Machinery for protein sorting and assembly in the mitochondrial outer membrane. Nature **424:** 565–571.

Wisniewski, J.P., and Rogowsky, P.M. (2004). Vacuolar H^+-translocating inorganic pyrophosphatase (Vpp1) marks partial aleurone cell fate in cereal endosperm development. Plant Mol. Biol. **56:** 325–337.

Xiao, Q., et al. (2017). ZmMYB14 is an important transcription factor involved in the regulation of the activity of the *ZmBT*1 promoter in starch biosynthesis in maize. FEBS J. **284:** 3079–3099.

Xiong, W., Wang, C., Zhang, X., Yang, Q., Shao, R., Lai, J., and Du, C. (2017). Highly interwoven communities of a gene regulatory network unveil topologically important genes for maize seed development. Plant J. **92:** 1143–1156.

Xu, J., and Zhang, S. (2015). Mitogen-activated protein kinase cascades in signaling plant growth and development. Trends Plant Sci. **20:** 56–64.

Xu, H., Gao, Y., and Wang, J. (2012). Transcriptomic analysis of rice (*Oryza sativa*) developing embryos using the RNA-Seq technique. PLoS One **7:** e30646.

Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., Zhiqiang, L., Yunfei, Z., Xiaoxiao, W., Xiaoming, Q., Yunping, S., and Li, Z., et al. (2006) The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol. Biol. **60:** 107–124.

Yelagandula, R., et al. (2014). The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in *Arabidopsis*. Cell **158:** 98–109.

Yeung, K.Y., Haynor, D.R., and Ruzzo, W.L. (2001). Validating clustering for gene expression data. Bioinformatics **17:** 309–318.

Yue, R., Tie, S., Sun, T., Zhang, L., Yang, Y., Qi, J., Yan, S., Han, X., Wang, H., and Shen, C. (2015). Genome-wide identification and expression profiling analysis of *ZmPIN, ZmPILS, ZmLAX* and *ZmABCB* auxin transporter gene families in maize (*Zea mays* L.) under various abiotic stresses. PLoS One **10:** e0118751.

Zhan, J., Thakare, D., Ma, C., Lloyd, A., Nixon, N.M., Arakaki, A.M., Burnett, W.J., Logan, K.O., Wang, D., Wang, X., Drews, G.N., and Yadegari, R. (2015). RNA sequencing of laser-capture microdissected compartments of the maize kernel identifies regulatory modules associated with endosperm cell differentiation. Plant Cell **27:** 513–531.

Zhang, Y.C., et al. (2013). Overexpression of microRNA OsmiR397 improves rice yield by increasing grain size and promoting panicle branching. Nat. Biotechnol. **31:** 848–852.

Zhang, B., Liu, J., Yang, Z.E., Chen, E.Y., Zhang, C.J., Zhang, X.Y., and Li, F.G. (2018). Genome-wide analysis of GRAS transcription factor gene family in *Gossypium hirsutum* L. BMC Genomics **19:** 348.

Zhang, M., Zhao, H., Xie, S., Chen, J., Xu, Y., Wang, K., Zhao, H., Guan, H., Hu, X., Jiao, Y., Song, W., and Lai, J. (2011). Extensive, clustered parental imprinting of protein-coding and noncoding RNAs in developing maize endosperm. Proc. Natl. Acad. Sci. USA **108:** 20042–20047.

Zlatanova, J., and Thakar, A. (2008). H2A.Z: View from the top. Structure **16:** 166–179.