Zygotic Genome Activation Occurs Shortly After Fertilization in Maize

Junyi Chen²*, Nicholas Strieder², Nadia G. Krohn¹*3, Philipp Cyprys¹, Stefanie Sprunck¹, Julia C. Engelmann², and Thomas Dresselhaus¹*

¹Cell Biology and Plant Biochemistry, Biochemie-Zentrum Regensburg, University of Regensburg, 93053 Regensburg, Germany.
²Institute of Functional Genomics, University of Regensburg, 93053 Regensburg, Germany.
³Present address: Department of Agriculture, Regional Campus of Umuarama, State University of Maringa, 87507-190, Umuarama, PR, Brazil.
*Corresponding author: thomas.dresselhaus@ur.de

Short title: Zygotic gene activation in maize

One-sentence summary: Transcription profiles generated from maize gametes and zygotes at different stages reveal a highly dynamic zygotic genome activation pattern, providing insights into early embryo development.

The author 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: Thomas Dresselhaus (thomas.dresselhaus@ur.de)

ABSTRACT

The formation of a zygote via the fusion of an egg and sperm cell and its subsequent asymmetric division (ACD) herald the start of the plant’s life cycle. Zygotic genome activation (ZGA) is thought to occur gradually, with the initial steps of zygote and embryo development being primarily maternally controlled, and subsequent steps being governed by the zygotic genome. Here, using maize (Zea mays) as a model plant system, we determined the timing of zygote development and generated RNA-Seq transcriptome profiles of gametes, zygotes, and apical and basal daughter cells. ZGA occurs shortly after fertilization and involves about 10% of the genome being activated in a highly dynamic pattern. In particular, genes encoding transcriptional regulators of various families are activated shortly after fertilization. Further analyses suggested that chromatin assembly is strongly modified after fertilization, that the egg cell is primed to activate the translational machinery, and that hormones likely play a minor role in the initial steps of early embryo development in maize. Our findings provide important insights into gamete and zygote activity in plants, and our RNA-Seq transcriptome profiles represent a comprehensive, unique RNA-Seq dataset that can be used by the research community.

INTRODUCTION

The life cycles of animals and plants begin with the formation of a zygote, containing the cytoplasm from two gametes, a large egg cell and a small sperm cell. This single cell then develops via embryogenesis into an entire organism consisting of hundreds of different cell types. In contrast to most animal systems, the flowering plant zygote
divides asymmetrically into daughter cells of completely different cell fates. While the small, cytoplasm-rich apical daughter cell further develops into the embryo proper, the highly vacuolated basal cell gives rise to the suspensor, which delivers nutrients to the embryo and positions the embryo proper in the surrounding endosperm tissue of the developing seed. Little is known about the establishment of zygote polarity and the gene regulatory network that leads to asymmetric cell division and cell fate determination in both daughter cells (Zhao et al., 2017). Due to the unequal distribution of their cytoplasm, it is generally thought that maternal factors contributed by the egg regulate zygote and early embryo development. The maternal-to-zygotic transition (MZT) depends on both zygotic genome activation (ZGA) and the degradation of maternal components. In animals, ZGA occurs after the first cell cycle in mammals and as late as the sixth to eighth round of the cell cycle in insects, fish, and amphibians (Schier, 2007; Lee et al., 2014). The time point at which ZGA occurs in plants has long been debated. Currently, the zygote is thought to be in a relatively quiescent transcriptional state, and ZGA is thought to occur gradually rather than as an all-or-none process initiated before the first division (Baroux and Grossniklaus, 2015; Zhao and Sun, 2015). However, analyses of a few candidate genes have indicated that ZGA in flowering plants occurs before zygotic division (Zhao et al., 2017).

To characterize the onset of ZGA at the whole-genome level, it is necessary to determine the transcriptome profiles of both gametes and to identify de novo generated transcripts from the zygotic genome. In the current study, we established methods to manually isolate living male and female gametes, zygotes at different stages, and their daughter cells using maize (Zea mays) as a model flowering plant. We then generated RNA-Seq data from these cells and investigated the transcriptome dynamics after fertilization. We compared the transcriptomes of maize and Oryza sativa (rice) gametes and explored how the cell cycle, chromatin, and auxin pathways are regulated after fertilization. Finally, we identified transcription factor (TF) and receptor-like kinase genes associated with the various cell types and zygotic stages and found that ZGA in maize occurs shortly after fertilization, displaying a highly dynamic pattern.
RESULTS AND DISCUSSION

Manual cell isolation, RNA-Seq, validation, and data quality

Maize plants were grown in indoor walk-in growth rooms to reduce seasonal variability in zygote development. We manually isolated cells of the inbred line B73 at 2-h intervals over a 2.5-day period as described in Methods. We used DAPI staining as a gross indicator to investigate cell cycle stages. As shown in Figure 1A, sperm cells consist mainly of the nucleus, containing highly condensed chromatin. Egg cell chromatin appeared less condensed than sperm cell chromatin and stained very weakly with DAPI. The strongest DAPI staining was detected in zygotes at 24 hours after pollination (HAP), indicating that S-phase was complete. Early anaphase occurred at ~30 HAP, and telophase began at 35 HAP. Cytokinesis was visible at 43 HAP, and asymmetric cell division (ACD) was completed by 48 HAP. The small cytoplasm-rich apical and vacuolated basal cells could be manually separated at ~52 HAP.

The protocol used to generate RNA-Seq data from only a few plant cells is described in Methods. Three biological replicates were prepared from approximately 1,000 sperm cells each and 13–20 cells each for the other stages (Table 1, see also Supplemental Table 1 and Supplemental Figures 1 and 2 for samples, sequencing details, and library quality). A gene expression master list, containing the median expression values of each cell type for the 39,469 annotated maize genes, protein annotations, and homologs from Arabidopsis thaliana and rice is provided in Supplemental Data Set 1.

For validation and initial analysis of the dynamics of gene expression patterns obtained from the RNA-Seq data, we focused on the transcript levels of 12 genes that were previously shown (by single-cell RT-PCR) to be highly or differentially expressed during fertilization and zygote development in maize (Engel et al., 2003). In agreement with previous results, the gamete-expressed membrane protein genes ZmGEX1 (SP vs. all: log₂FC >4.0*) and ZmGEX2 (SP vs. all: log₂FC >7.4*) were highly and specifically expressed in sperm cells, as were Zmsp271 (SP vs. all: log₂FC >3.8*) and Zmsp041 (SP vs. all: log₂FC >5.2*), which were identified in the same screen (Figure 1I, Supplemental Data Sets 1–3). ZmA1 (EC vs. SP: log₂FC =8.7*) and ZmES1-4 (EC vs. AC/BC: log₂FC >2.9 to 9.7*, Zy24 vs. AC/BC: log₂FC >2.4 to 8.7*), encoding secreted peptides required for micropylar pollen tube
guidance and pollen tube burst, respectively, were highly expressed in egg cells and
synergids and were significantly downregulated after fertilization (Cordts et al., 2001;
Marton et al., 2005; Amien et al., 2010) (Supplemental Data Set 3). The cell cycle
genes ZmMCM3, ZmMCM6, ZmCycB1;2, and ZmCycB2;1 were previously shown to
be induced after fertilization (Sauter et al., 1998; Dresselhaus et al., 1999b; Dresselhaus et al., 2006). Expression of ZmMCM3 (Zy12 vs. EC: log$_2$FC =2.7*, AC
vs. Zy24: log$_2$FC =1.8*) and ZmMCM6 (Zy12 vs. EC: log$_2$FC =2.0*, AC vs. Zy24:
log$_2$FC =2.7*), marking the onset of DNA replication during S-phase (Maiorano et al., 2006), peaked in the zygote at 12 HAP, as well as after the first asymmetric zygote
division in the apical cell, which divides more rapidly than the basal cell. The cell
cycle regulatory genes ZmCycB2;1 (Zy24 vs. Zy12 log$_2$FC =3.6*) and ZmCycB1;2
(Zy24 vs. Zy12 log$_2$FC =5.0*), which mark the G2/M-transition (Maiorano et al., 2006), were strongly induced at 24 HAP. In contrast to ZmCycB1;2 (AC/BC vs. Zy12
log$_2$FC >1.9*), the expression levels of ZmCycB2;1 (AC/BC vs. Zy12 log$_2$FC >5.5*)
were also high in apical and basal cells after zygote division (Sauter et al., 1998). In
summary, these dynamic changes in gene expression (Figure 1B) are in perfect
agreement with previous reports, which together with strong correlation between
biological replicates (Supplemental Figure 2) assures the high quality and reliability
of our data.

Contamination of transcriptomes by RNA from maternal tissues has recently
been discussed as a serious issue that can result in poor reproducibility and
misinterpretation of data sets (Schon and Nodine, 2017). We therefore investigated
the presence of transcripts derived from genes expressed in maternal nucellus tissue
surrounding embryo sacs (Chettoor et al., 2014) to evaluate the possibility of
contamination. None of the nucellus-expressed genes, including GRMZM2G570791
(alpha subunit of DNA-directed RNA polymerase), GRMZM2G125823 (heparanase-
like protein), GRMZM2G099420 (Cinnamoyl CoA reductase), and
GRMZM5G803276 and GRMZM2G336859 (encoding unknown proteins), were
detected in any of our data sets. These results indicate that our data sets are free of
maternal RNA contamination and that the two washing steps were sufficient for
removing maternal RNA from the burst maternal nucellus cells.

Comparison of transcriptomic data from maize and rice gametes
A comprehensive comparison of gene expression activity after fertilization has not been reported yet for any plant species, and the present study thus represents the first report of global gene expression patterns in gametes, zygotes, and daughter cells. Therefore, we restricted our comparisons to the transcriptomes of maize and rice gametes (egg and sperm cells). It was not possible to include the transcriptomes of Arabidopsis thaliana gametes in the comparison, as RNA-Seq data were not available, and the available microarray data (Borges et al., 2008; Wuest et al., 2010) could not be accurately normalized to allow us to draw conclusions and lacked information for thousands of genes. In addition, each gamete in the data set was measured in a different experiment.

We used published RNA-Seq data from rice sperm and egg cells (Anderson et al., 2013) and initially identified the rice homologs using public databases, i.e., EnsemblPlants and RiceAnnotationGenomeProject, which combine data from many species to identify putative orthologs. If the identity of the homologs/orthologs was unclear or unknown due to a lack of sequence information, we did not include them in the comparison. To compare transcription patterns in rice versus maize gametes, the gene expression values were binned into 200 expression level categories using the 99th percentile per species as the highest category (see also Supplemental Data Set 4). We selected the 80 most strongly expressed genes (TOP80 genes) in maize sperm and egg cells and compared their expression levels with those of the respective genes in rice (Figure 2). A summary of the TOP30 genes of all maize cell types, including their annotation, is provided in Supplemental Data Set 5.

Many of the predicted rice homologs/orthologs displayed similar, strong expression patterns. The proportion of genes with high expression levels was greater in egg cells than in sperm cells (Figure 2). The observation that many predicted rice homologs/orthologs displayed weaker expression patterns in rice than in maize might be due to our methods (as we summarized rice ortholog data using the median), the difficulty in predicting true orthologs within larger gene families, and/or the lack of common controls to normalize these two studies. Tightly controlled parallel RNA extraction from cells of different species and the identification of an appropriate control cell type common to both plant species may improve interspecies comparisons. However, our comparison pointed to some similarities and general findings regarding genome activity. Among the TOP80 genes expressed in maize sperm cells, 10% encode histones and high-mobility group (HMG) proteins (Figure
This finding might explain the strongly condensed, compact sperm cell chromatin (see also Figure 1A). Indeed, a similar observation was reported for rice sperm cells (Russell et al., 2012b; Anderson et al., 2013). Notably, no chromatin gene is among the TOP80 genes in maize egg cells. This finding correlates well with the less condensed chromatin in these cells and the difficulty in staining egg cells with DAPI (Figure 1A). However, in strong contrast to sperm cells, 20% of the TOP80 genes in maize egg cells encode proteins of the translational machinery, most of which also displayed similar expression patterns in rice egg cells. These include two genes encoding the translation initiation factor IF5A (Dresselhaus et al., 1999a), a gene encoding the translation initiation factor SUI1 (Cui et al., 1998), two genes encoding the translation elongation factor EF1A (Budkevich et al., 2002), and many ribosomal protein genes. These observations indicate that sperm cells are translationally inactive, whereas egg cells are either highly active or well prepared to strongly enhance translation after being activated during the fertilization process.

Another interesting observation relates to transcripts encoding polymorphic EA1-box proteins and small secreted cysteine-rich proteins (CRPs), which we found in maize egg cells, but not in sperm cells. The corresponding proteins play key roles in female gamete cell identity, as well as pollen tube guidance and burst in maize (for review, see (Dresselhaus et al., 2011; Dresselhaus et al., 2016). However, due to their polymorphic nature, no unambiguous homologs of the individual members were identified in the databases from Ensembl Compara or the Rice Genome Annotation Project. Through manual searches, we identified similar genes in rice (Márton et al., 2005; Uebler et al., 2015), but true orthologs could not be predicted based on protein comparisons, and none of the candidate genes have been functionally tested in rice (Figure 2B). However, this finding indicates that genes involved in processes directly associated with fertilization appear to be polymorphic and specific-specific, thus representing prime candidate genes involved in speciation mechanisms (Rieseberg and Willis, 2007).

Histone variants and chromatin-based gene regulation

Histones are the major protein components of chromatin and, more importantly, dynamic regulators of transcription. To begin to uncover the molecular basis of chromatin remodeling and epigenetic reprogramming in plant gametes and the onset of embryo development, we investigated the expression patterns of histone variants...
and chromatin assembly factor genes in more detail (Figure 3, Supplemental Data Set 3G; important gene families).

As mentioned above, some canonical core histones, including two H3 (7 of 17 genes: SP vs. all: log₂FC>1.5*) genes and at least three H4 (9 of 15 genes: SP vs. all: log₂FC>1.4*) genes, were predominantly and highly expressed in sperm cells (Figure 3B and C), and may contribute to the compactness of the chromatin and suggesting that sperm cells are already prepared for S-phase. Moreover, the most highly expressed gene in sperm, ZmHmgd1 (SP vs. all: log₂FC>3.9*), was expressed at much higher levels in sperm cells than in any other cell type examined (>~6-fold) (Figure 3F). ZmHmgd1 encodes an HMG box protein that plays a role in chromosome condensation (Thomas and Travers, 2001) and is thought to possess a role similar to that of histone H1 in chromatin assembly, as both proteins bend linker DNA at the entry and exit points of the nucleosome. This hypothesis may explain why linker histone H1 genes are expressed at rather low levels in sperm cells, although their DNA is densely packed to ensure chromatin stability during sperm delivery. After fertilization, Hmgd1 appeared to be partially replaced by histone H1, as the expression of the H1 gene was strongly activated and peaked in 24 HAP zygotes (9 of 11 genes: Zy24 vs. EC: log₂FC>1.8*) (Figure 3A), whereas ZmHmgd1 (Zy24 vs. SP: log₂FC= -5.0*) was strongly repressed.

The most abundant H3 gene in maize sperm (GRMZM2G145758, SP vs. all: log₂FC>4.0*) encodes an unusual replication-independent (RI) H3.3-like variant that was also predominantly expressed in anthers and was previously designated as ZmAPH3. Phylogenetic analysis showed that ZmAPH3 is most similar to the Arabidopsis male gamete-specific histone H3 variant gene AtMGH3 (Okada et al., 2005) (Supplemental Figure 3A). However, the horizontal distance and the low number of histone variants and species included in the phylogenetic analysis suggest that they share a rather distant relationship. All other sperm cell-enriched H3 variants belong to the RI H3.3 group (Supplemental Figure 3A) (four orthologous genes, RI H3.3: SP vs. Zy24: log₂FC>1.6*). In Arabidopsis, a limited subset of H3.3 variants referred to as the main HTR (HISTONE THREE-RELATED) protein genes are predominately expressed in the male gamete (Ingouff et al., 2010). Three AtMGH3 homologs and two other RI H3.3 variants are also highly expressed in rice sperm cells, while other core H3 genes exhibit only limited expression in these cells.
These findings suggest that RI H3.3 variants encoded by only a few genes act as major histone H3s in the sperm cells of flowering plants. Two replication-coupled canonical H2As (GRMZM2G151826 and GRMZM2G041381) exhibited high transcript levels in maize sperm cells, but the most highly expressed H2A (GRMZM2G149775, SP vs. all, log$_2$FC>2.7) belongs to the RI H2A.Z class (Supplemental Figure 3B), which is associated with nucleosomes at transcription start sites, especially those also containing H3.3 (Deal and Henikoff, 2011). Therefore, together with histone H3.3 and H2A.Z, which mark active chromatin via RI nucleosome assembly (Ahmad and Henikoff, 2002; Deal and Henikoff, 2011), HMGD1 (encoded by ZmHmgd1), which likely represents the most abundant chromatin architectural protein in sperm that binds to highly accessible regulatory chromatin and active promoters, probably keeps the highly condensed sperm cell chromatin at least partially accessible. This hypothesis is supported by our RNA-Seq data set, which includes transcripts of approximately 11,000 differentially upregulated genes in sperm cells versus zygotes at 12 HAP (Figure 4F), including a subset of TF genes, despite the compact chromatin in sperm cells. Upon fertilization, H3 variants from male and female gametes are actively removed in Arabidopsis (Ingouff et al., 2010), a scenario that likely also occurs in maize, as the expression of the canonical core histone repertoire was activated in conjunction with a dramatic reduction in H2A.Z, H3.3, and Hmgd1 expression and a strong increase in H1 expression (Figure 3A). Thus, it appears that the paternal chromatin was reprogrammed by newly synthesized and entirely different sets of histones in the zygote. This theme may even extend to apical and basal cells, since the list of expressed orthologs of H3 and H4 again shifted to genes different from those expressed in zygotes.

Finally, while the canonical core histones were mostly expressed at very low levels in egg cells (Figure 3), they were induced in zygotes at 12 HAP, indicating the G1 phase of the cell cycle. Expression peaked in zygotes at 24 HAP, suggesting that DNA replication was almost complete (Figure 3). These observations support the hypothesis that the egg cell is arrested and requires activation after gamete fusion. Thus, chromatin-based transcriptional reprogramming in the zygote may represent a key step in MZT and the initiation of the sporophytic generation.

ZGA occurs shortly after fertilization in maize
To identify the global onset of ZGA, it is important to examine the transcriptomes of both gametes, i.e., egg and sperm cells, as well as zygotes at different stages. We therefore investigated zygotes collected at 12 and 24 HAP. At 12 HAP, i.e., only ~4 h after fertilization, all egg cells were fertilized, as indicated by the presence of degenerated receptive synergid cells. These egg cells were considered to represent early zygotes. We chose to investigate zygotes at 24 HAP, as they appeared to be activated at this point, as indicated by the duplication of DNA content (Figure 1A). At later time points after pollination, the zygotes underwent mitosis, but the stages appeared less synchronous than those at earlier time points. We chose genes with an abs(log2FoldChange) >1 and adjusted p-value<0.05 in the respective comparison as genes with a potential biological function and determined whether they were induced or repressed. Based on these criteria, sperm cells formed the most distinct group, expressing 4,090 differentially upregulated genes in all comparisons (Figure 4F). As all of the other cell types are similar, fewer genes were induced in these cells compared to all other cell types; 482 differentially upregulated genes were detected in egg cells in all comparisons, 109 were detected in zygotes at 12 HAP, 31 were detected in zygotes at 24 HAP, and 6 and 8 were detected in apical and basal cells, respectively (Figure 4F, Supplemental Data Set 6).

We compared zygotes at 12 HAP to both sperm and egg cells and identified 3,605 induced genes (9.1% of the maize transcriptome of 39,469 annotated genes; Schnable et al., 2009) over all chromosomes shortly after fertilization (Figure 4B). This high number of activated genes indicates that global ZGA in maize already occurs in the early zygote and not several days after fertilization, as previously reported (Grimanelli et al., 2005). Although de novo transcription has also been observed in Arabidopsis zygotes (e.g., Autran et al., 2011; Nodine and Bartel, 2012; Del Toro-De León et al., 2014), the timing of global ZGA in this species is unclear, and some reported results are currently under debate due to contamination from maternal tissues (Schon and Nodine, 2017).

We compared zygotes at 12 HAP with sperm and egg cells and identified 7,356 and 1,933 induced genes, respectively (Figure 4B). Furthermore, we identified 1,730 differentially upregulated genes in sperm and egg cells versus zygotes at 12 HAP (Figure 4A), which can also be viewed as genes downregulated in zygotes at 12 HAP compared to gametes. These findings also suggest that major rearrangements in the transcriptome occur following fertilization.
To obtain a global overview of transcriptome dynamics during zygote development, we defined various gene expression profiles as capturing not only on and off states, but also the induction or repression of genes during the transition from gametes to zygotes and their descendant cells (Figure 4E, Supplemental Data Set 6). Very specific gene expression profiles were found for egg cells: 326 genes were upregulated in this cell type, whereas 95 genes were downregulated. Shortly after fertilization, 356 and 1,510 genes were transiently induced only in zygotes at 12 HAP and in zygotes at both stages, respectively. Approximately 10% of the genes in each group encode transcriptional regulators. Of the 3,808 genes induced after fertilization, 223 were predominantly expressed in apical cells, whereas 182 were predominantly expressed in basal cells. These results indicate that the expression levels of many genes are higher in the zygote than in its progenitor cells, supporting the notion of an early onset for ZGA. Few genes were transiently repressed after fertilization (137 genes). Together, these findings reveal a highly dynamic transcriptional landscape after fertilization in maize and demonstrate that ZGA occurs shortly after fertilization in this plant. Studies involving pollination with other inbred lines are now needed to distinguish between maternal and paternal transcripts and thus to determine whether both genomes contribute equally to ZGA in this species.

Transcription factor activation schemes in gametes and zygotes

As TFs are major regulators of gene expression, we next examined TF gene expression levels in gametes and early zygotic embryos based on the maize TF database Grassius (www.grassius.org). Across all cell types, 1,478 of 2,630 maize TF genes were expressed in gametes and zygotes (log₂FC>1 and padj<0.05 in at least one comparison). Comparing their transcription levels during early development (Figure 5A, Supplemental Data Set 3) showed that zygotes formed a group distinct from apical and basal cells, and that both groups differed from gametes. We identified 428 TF genes that were induced in zygotes (12 and 24 HAP), 189 of which were strongly activated shortly after fertilization (Zy12 vs. SP and EC, log₂FC>3). We detected 25 TF genes specifically expressed in pro-embryonic cells. Only 23 paternal and 103 maternal TF genes were expressed at similar levels in gametes and zygotes (Supplemental Data Set 3). While we could not distinguish between paternal and maternal mRNAs in our assays and were thus unable to identify the
maternal-to-zygotic transition (MZT) (Baroux and Grossniklaus, 2015), the results suggest that approximately 8.6% (25+103/1,478) of expressed TF genes are parentally transmitted, while 29% (428/1,478) are newly or more intensely transcribed in maize zygotes. These data support the notion that TF genes are activated early in maize zygotes (at 12 HAP; 403 of 428 TF genes were induced in zygotes vs. 25 of 428 at 24 HAP) (Figure 5A).

We then identified important TF classes from the various expression profiles (Supplemental Data Set 6; profiles of TF classes are summarized in Supplemental Data Set 7). In sperm cells, many genes belong to the TF class AT-rich interactive domain (ARID 5/10) proteins. Members of this class have been implicated in sperm cell development in plants and mammals. In Arabidopsis, ARID1 is necessary for the appropriate expression of DUO1, a major TF required for sperm cell formation (Zheng et al., 2014). In mice, the loss of ARID4A combined with ARID4B haploinsufficiency leads to spermatogenic arrest (Wu et al., 2013). In egg cells, we detected high proportions of TFs from the classes FAR1-like (3/15), mTERF (8/30), Sigma70-like (2/9), S1Fa-like (2/2), and GeBP (5/21) compared to the other cell types analyzed. The first four classes are related to plastid development (Zhou et al., 1995; Ouyang et al., 2011; Kleine, 2012; Wei et al., 2012). This finding suggests that a tightly controlled regulatory network controls plastid development during the first steps in the plant life cycle. Members of the GL1 enhancer binding protein (GeBP) class of leucine-zipper TFs have been linked to the cytokinin response in Arabidopsis and are thought to be mainly expressed in vegetative meristems and in the primordia of young leaves (Chevalier et al., 2008). Our data suggest that this class of TFs also plays a role in the transition from egg cells to early zygotes.

As discussed above, at the first zygotic time point (12 HAP), genes for many TF classes were induced (Supplemental Data Set 7). The auxin-responsive ARF TFs (7/38) and the ethylene-responsive AP2-EREBP TFs (34/212) might be important at this stage, as crosstalk between these pathways is thought to be essential during zygote and pro-embryo development (see below). In addition, many genes from the TF classes C3H/CCCH (18/54), Trihelix (11/41), ZIM (10/36), MADS (15/77), NAC (26/134), bZIP (24/128), and Homeobox (26/133) were induced at this time point. MADS-box TFs are associated with reproductive organ development and play a role in gametes and in zygotic embryogenesis (Schreiber et al., 2004; Lehti-Shiu et al., 2005).
Activation of embryo patterning

After ACD, WOX genes encoding homeodomain TFs mark apical and basal cell fate upon zygote division in Arabidopsis (Breuninger et al., 2008). In maize, ZmWOX9A and ZmWOX9B likely represent the homologs of AtWOX8 and AtWOX9, respectively (Salvo et al., 2014). Both ZmWOX9A and ZmWOX9B were induced shortly after fertilization (Zy12 vs. SP/EC: log2FC>5.9* and 7.0*, respectively) and were expressed at higher levels in basal cells than in apical cell, like their counterparts in Arabidopsis (Figure 5B, Supplemental Data Set 3B), indicating that they might play a similar role in early embryonic patterning. However, unlike AtWOX2, which marks apical descendants of the zygote, ZmWOX2A was expressed at very low levels in basal cells, with almost no expression in apical cells. Instead, this gene was expressed much later during seed development in the endosperm (Maize eFP Browser at bar.utoronto.ca). This finding indicates that pattern formation regulated by ZmWOX2A and other genes likely occurs later in maize embryogenesis than in Arabidopsis embryogenesis (Zhao et al., 2017). Notably, ZmWOX13A is the only WOX gene that was transcribed at high levels in sperm cells (SP vs. all: log2FC>4.3*); whether it marks cell identity of the male gamete or represents a zygote activator remains to be investigated. In Arabidopsis, BBM and LEC TF genes are key players in embryogenesis, and the presence of either gene is sufficient to induce competence for embryogenesis (Lotan et al., 1998; Boutilier et al., 2002). Overexpressing a combination of the maize homologs WUS2 and BBM was recently shown to significantly increase embryogenic potential in tissue culture and thus represents a key mechanism to increase the transformation efficiency in maize (Lowe et al., 2016). Notably, the maize homologs ZmBBML1 (Zy12 vs. SP/EC: log2FC>10.5*), ZmBBML2 (Zy12 vs. SP/EC: log2FC>10.2*), and ZmLEC1 (Zy12 vs. SP/EC: log2FC>7.9*) were already induced at 12 HAP (Figure 5B), suggesting that the egg cell quickly acquires embryogenic competence and that the characteristic embryonic transcription program is activated shortly after fertilization. ZmBBML3 was induced in zygotes at 24 HAP (Zy24 vs. Zy12: log2FC=6.5*) and after zygote division (AP/BC vs. Zy24: log2FC=3.5/3.9*), suggesting step-by-step activation of the embryonic program. A comparison of apical and basal cells with zygotes at 24 HAP revealed 2,228 genes induced in both cell types, 832 induced only in apical cells, and 485 induced only in basal cells (Figure 4D). This induction was accompanied by
the downregulation of 2,182 genes in apical and basal cells versus zygotes at 24 HAP (Figure 4C). These data suggest that global rearrangements also occur in the transcriptomes of apical and basal cells compared to their predecessor.

The development of multicellular organisms often involves ACDs to generate daughter cells with different cell fates. Spindle positioning is particularly associated with the generation of symmetric or asymmetric cell fates (Siller and Doe, 2009). The MATH-BTB domain protein ZmMAB1, a component of a CUL3-E3 ubiquitin ligase complex, regulates spindle length during the development of the male and female germline in maize (Juranic et al., 2012). ZmMAB1 may also function like its animal homolog, the key ACD regulator MEL-26, a factor required for embryogenic morphogenesis that regulates the formation of mitotic spindles in the early embryo (Pintard et al., 2003). As shown in Figure 5C, MAB family genes were strongly upregulated (ZmMAB1-3, ZmMAB6 and ZmMAB24, Zy12 vs. SP: log2FC >1.5 to 10.7*) after fertilization in maize. The highest expression levels of these genes were detected in zygotes at 24 HAP (Supplemental Data Set 3A; important gene families), suggesting that they play roles in processes such as spindle positioning during the first asymmetric zygote division (ZmMAB2-3, Zy24 vs. Zy12 log2FC=3.1* and 1.8* respectively). Functional studies of ZmMABs are now needed to investigate this hypothesis.

Cell cycle regulation during zygote development

Since previous reports provide only a glimpse of cell cycle regulation in plant gametes and during zygote development, we investigated the expression patterns of important cell cycle regulator genes. First, we searched for orthologs of Arabidopsis cell cycle genes (Vandepoele et al., 2002) and then included cyclins and other cell cycle-related factors described previously for maize (see Methods), resulting in a list of 89 cell cycle genes (Supplemental Data Set 3). As shown in Figure 6A, hierarchical clustering of these genes from different cell types clearly separated sperm cell genes from a group of genes from egg cells and zygotes at 12 HAP. The expression patterns of the genes from zygotes at 24 HAP were more similar to those of apical and basal cells.

The G1-phase of sperm cells is characterized by high expression levels of Cdc27-like genes (GRMZM2G148626, GRMZM2G005536, SP vs. all: log2FC>5.7*) encoding subunits of the APC complex, which controls CDK degradation in M- and
G1-phase, and relatively low levels of CDK gene expression. Furthermore, retinoblastoma-related protein genes *RBR1* and *RBR2* were highly expressed in sperm cells (*Figure 6A*, SP vs. all: log$_2$FC$>$2.9$^*$). RBR1 and RBR2 mediate G1-phase arrest by inhibiting E2F TFs, which in turn promote DNA replication (Sabelli et al., 2013). Compared to the other cell types, egg cells showed by far the lowest expression levels of cell cycle genes and lacked a typical cell cycle phase-specific gene expression pattern. This finding, together with the results of cell cycle gene expression analysis and the microscopy evidence reported above (*Figure 1*), indicate that the egg cell is in a resting G0 stage rather than in G1 (*Figure 6B*). Thus, the egg cell must be activated and its cell cycle synchronized with the sperm cell cycle stage before karyogamy (fusion of both nuclei) is executed. A typical G1-phase expression pattern was observed in zygotes at 12 HAP (*Figure 6A*), with slightly upregulated expression of E2F TF genes (GRMZM2G060000, GRMZM2G361659, GRMZM2G378665, Zy12 vs. EC, log$_2$FC $>$2.5$^*$) and (especially) their cell cycle target genes (*Supplemental Data Set 3*). The latter include mitochrondosome maintenance genes (*ZmMCM3*-6, Zy12 vs. EC: log$_2$FC $>$2.0$^*$), encoding DNA-replication licensing factors required for replication initiation, and the gene encoding proliferating cell nuclear antigen (PCNA), which acts as a scaffold to recruit proteins involved in DNA replication and repair (Sabelli et al., 2009; Tuteja et al., 2011). Notably, *RBR3*, encoding an activating factor of *MCM2*-7 transcription, was specifically induced in zygotes at 12 HAP (Zy12 vs. EC: log$_2$FC $=$2.2$^*$), while *RBR1* and *RBR2*, two genes encoding repressors of E2F TFs, were repressed in these zygotes compared to sperm cells (Zy12 vs. SP: log$_2$FC $<$-3.8$^*$). This expression pattern coincided with the induction of CDK A genes (Zy12 vs. SP: log$_2$FC $>$ 1.1 to 7.0$^*$) and low levels of expression of CDK B genes, indicating that fertilized egg cells had been activated and zygotes at 12 HAP were in G1-phase. *MCM2*-7 expression levels were lower in zygotes at 24 HAP versus 12 HAP, whereas CDK B2 genes (9 of 14 genes, Zy24 vs. Zy12: log$_2$FC $>$2.8$^*$) and B-type cyclin genes were induced (8 of 9 genes, Zy24 vs. Zy12: log$_2$FC $>$1.9$^*$), especially *ZmCycB1;2*, suggesting that S-phase was completed and the zygotes were prepared for the G2/M-phase transition (Meijer and Murray, 2001). M-phase took place at 27–35 HAP (*Figure 1A*). Notably, G1-phase markers such as *MCM2*-7 were expressed at slightly higher levels in apical versus basal cells, hinting at more rapid cell cycle progression in the apical cell after ACD.
Taken together, our cell cycle analysis by microscopy of DNA staining patterns, transcription data from selected gene sets, and global cell expression analysis of cell cycle genes allowed us to determine the timing of zygote development in maize (Figure 6B). On average, fertilization occurs at ~8 HAP. Sperm cells appear in G1-phase, and egg cells are in a resting G0-phase of the cell cycle. Activated zygotes are in G1 phase at 12 HAP (~4 h after fertilization) and at G1/S-phase at 24 HAP. Mitosis typically occurs at 26–36 HAP, and cytokinesis lasts until 44 HAP. ACD is completed between 44 and 50 HAP, generating both apical and basal cells in G1-phase.

The role of auxin in early embryogenesis in maize

To demonstrate the utility of our data sets, we chose the auxin pathway, which plays a key role in early embryo patterning in Arabidopsis, as an example for analysis. Auxin gradients generated by PIN efflux carrier proteins establish the apical-basal axis upon the first ACD in the Arabidopsis zygote (Friml et al., 2003). The molecular mechanisms that determine axis formation during early embryogenesis in monocots are largely unknown. We therefore analyzed the expression of auxin biosynthesis, transport, and response genes, as shown in Figure 7 (Supplemental Data Set 3D; important gene families). In maize, the earliest ZmPIN1a localization was observed at 6 DAP at the adaxial side of the embryo proper (Chen et al., 2014). To identify potential PIN proteins that function at the earliest stage of embryo patterning, we analyzed the expression of all PIN family genes. Among the 15 PIN genes in maize (Yue et al., 2015), only ZmPIN8 was weakly expressed in the zygote and upregulated after ACD in both apical and basal cells (AC/BC vs. Zy12: log2FC=3.8/3.7*). A number of ZmABCBs, representing potential auxin transporter genes (Yue et al., 2015), were also active in maize gametes and/or zygotes. Auxin biosynthesis genes ZmTAR1 (Zy12 vs. SP/EC: log2FC>1.2*) and ZmYUC3 (Zy12 vs. EC: log2FC=9.4*) were significantly induced after fertilization, with higher expression levels in basal cells versus apical cells (ZmTAR1 BC vs. AC:log2FC=1.4*). Auxin receptor genes ZmABP1, ZmABP4, and ZmABPL were also expressed in egg cells (EC vs. SP: log2FC>1.1*) but not in sperm cells and became transiently activated after fertilization (Zy12 vs. SP: log2FC>1.9*, AC/BC vs. Zy24: log2FC< -1.4*). Of the auxin responsive factor (ARF) genes examined, ZmARF7 had the highest expression level in the egg cell and was almost completely switched off after
fertilization (EC vs. SP/Zy12/Zy24: log2FC>3.1*, AC/BC vs. EC: log2FC< -4.1*).

Other ARF genes, such as ZmARF8, 13, 17, and 28, were expressed at similar levels in all cells except sperm cells. In general, ARF transcript levels were higher in the apical daughter cell of the zygote than in the other daughter cell (Figure 7B: AC/BC vs. Zy24). A few AUX/IAA repressor genes encoding proteins that interact with ARF regulators were activated after fertilization (Figure 7B: Zy12 vs. SP/EC). In particular, ZmIAA17 (Zy12 vs. SP/EC log2FC>4.3) was transiently expressed only shortly after fertilization; ZmIAA17 might be involved in the inactivation of ZmARF17-regulated gene expression patterns. Another highly upregulated gene, ZmSAUR7 (Zy12 vs. SP/EC: log2FC> 5.2*, AC/BC vs. Zy12: log2FC< -2.6*), is one of 79 SAUR (SMALL AUXIN UP RNAs) genes in maize, representing the largest family of auxin response genes (Ren and Gray, 2015). Globally, we found that auxin pathway genes were highly induced in the early zygote at 12 HAP and expressed at decreasing levels from 24 HAP zygotes to apical and basal cells (Figure 7B). By contrast, in Arabidopsis, these genes continue to show a strong auxin response, especially in the apical cell after ACD (for review, see Petrášek and Friml, 2009). Moreover, transcripts encoding homologs of key players in auxin-regulated early embryo patterning in Arabidopsis, such as ARF5 (MP), IAA12 (BDL), PIN1, and PIN7, were absent in zygotes and their daughter cells in maize.

The observation that gametes and early zygotes are equipped with transcripts encoding proteins for auxin biosynthesis, transport, and perception, as well as the identification of strongly regulated auxin response genes (Supplemental Data Set 8), indicates that auxin-regulated early embryo patterning is likely different in maize and other grasses compared to Arabidopsis, providing an entry point for investigating the role of this important hormone during early embryogenesis in grasses. In addition to auxin signaling, we also obtained hints about brassinosteroid and ethylene signaling during early embryo development in maize, which will be investigated in the future.

**Cell signaling during fertilization and early embryogenesis**

Ca²⁺ signaling is thought to play a pivotal role in fertilization by regulating a plethora of cellular processes (Chen et al., 2015). Annexins are a class of Ca²⁺-regulated proteins that link Ca²⁺ signaling to membrane and cytoskeleton dynamics (Gerke et al., 2005). Of the 12 genes encoding annexins in maize (Zhang et al., 2015),
ZmAnn6 and ZmAnn7 exhibited very high, transient expression in egg cells (EC vs. SP: log2FC=8.5/2.1*, EC vs. Zy12: log2FC=1.1/-*) and early zygotes (AC/BC vs. Zy12: log2FC< -4.1/-3.2*), respectively (Figure 8A, Supplemental Data Set 3C; important gene families). Thus, these proteins might function as key players in intra- and intercellular Ca2+ signaling during fertilization and early embryo development.

Signal perception and transduction through cell-surface receptor-like kinases (RLKs) likely play also roles in gamete interaction, fertilization, and early seed development in plants. We detected at least three RLK genes that were preferentially expressed in sperm cells (GRMZM2G011806, GRMZM2G016480, GRMZM2G428554, SP vs. EC/Zy12/Zy24: log2FC>2.3*) (Figure 8B, Supplemental Data Set 3E; important gene families), representing potential players in gamete recognition and/or sperm activation. We did not identify RLK genes that were preferentially expressed in egg cells. In Arabidopsis, SOMATIC EMBRYOGENESIS RECEPTOR KINASE1 (AtSERK1) is expressed in developing ovules and early embryos, and enhances embryonic competence in cell culture (Hecht et al., 2001). While its maize ortholog ZmSERK1 (Salvo et al., 2014) was expressed at very low levels in gametes and zygotes, ZmSERK2 and ZmSERK3 were expressed in zygotes (e.g., ZmSERK3 Zy12 vs. EC: log2FC=3.7*) and daughter cells (Figure 8C), indicating the involvement of similar signaling pathways in embryonic initiation. Moreover, several RLK genes (GRMZM2G038165, GRMZM2G428554, GRMZM2G089461) were upregulated and differentially expressed in apical and basal cells (AC/BC vs. Zy24: log2FC>1.5*), thus representing exciting candidates for future functional studies investigating cellular communication during early embryo development in grasses.

Conclusions

Detailed analysis of global gene expression patterns in plant gametes, zygotes, and manually separated apical and basal cells has allowed the onset of global ZGA in maize to be determined for the first time. The observation that ZGA occurs soon after fertilization, displaying a highly dynamic and partially transient pattern, is surprising and contradicts previous studies using a limited number of genes. These studies indicated that the zygote is in a relatively quiescent transcriptional state, that only a few genes are de novo activated in the zygote, and that ZGA occurs gradually rather than all at once (Baroux and Grossniklaus, 2015; Zhao and Sun, 2015). The striking
differences in the expression patterns of cell cycle regulators between sperm and egg cells coincide with a distinct chromatin state in sperm cells and define a quiescent cell cycle state in egg cells, although egg cells appear to be translationally highly active or well prepared to quickly activate the translational machinery after fertilization. The chromatin state in sperm appears to depend on replication-independent histone assembly and the HMG protein ZmHmgd1, which likely keeps the highly condensed sperm cell chromatin at least partially accessible, as demonstrated by the numerous transcribed genes. In addition, our data allowed us to differentiate between the stages during G1-phase that occur in zygotes at 12 HAP, apical cells, and basal cells, and they suggest a preference for certain CDKs and cyclins during the first two cell cycles in plants. Analysis of the expression levels of TFs, structural regulators, and signaling pathway genes allowed us to identify relevant genes homologous to key, well-known Arabidopsis regulators as well as novel candidate genes, which will serve as a starting point for many future studies.

In summary, our analyses of the genes described above represent only a few examples of how our comprehensive dataset can be used. This gene expression atlas should further accelerate the identification of key players involved in many biological processes, including fertilization, early embryogenesis, and the cell cycle, as well as the translational machinery. In addition, our data set could be used to uncover genes (and their corresponding promoters) for use in future efforts aimed at increasing seed yield and quality in maize and other crops.

METHODS

Plant materials and growth conditions
Maize (Zea mays) inbred line B73 was cultivated in a walk-in plant growth room at 26°C under illumination of 24,000 lux using alternating SON-T Agro and HPI-T Plus bulbs with a 16 h light/8 h dark cycle and a relative humidity of 60%. Flowers at anthesis and pollinated cobs were used to isolate cells for RNA-Seq.

Isolation of cells from male and female gametophytes
Hundreds of maize plants were grown to collect sufficient numbers of manually isolated cells. Each biological replicate consisted of pooled cells from different
plants. Only the middle part of the cob was used for cell isolation from excised ovules. A whole cob was used to isolate cells at a defined developmental stage. Sperm cells were released from maize pollen grains (male gametophytes) by osmotic shock and separated using density gradient centrifugation on a discontinuous Percoll gradient. The detailed protocol is given below. Egg cells were isolated from embryo sacs (female gametophytes) of unpollinated ovules as described (Kranz et al., 1991). Early and late developmental stage zygotes were isolated from ovules at 12 hours after pollination (HAP) and 24 HAP, respectively, as previously described (Cordts et al., 2001). Apical and basal cells were dissected from two-celled proembryos isolated from ovules at between 48-52 HAP following the procedure used for zygotes with some modifications (described in detail below). All cells isolated from ovules were individually collected using a microcapillary and washed twice in mannitol solution (480 mOsmol·kg\(^{-1}\) H\(_2\)O). Cells showing cytoplasmic streaming were individually transferred to 0.5 ml Eppendorf RNA/DNA LoBind microcentrifuge tubes, immediately frozen in liquid nitrogen, and stored at −80°C for mRNA extraction. Three biological replicates (each representing an independent pool of cells) were carried out, each with ~1,000 sperm cells, 20 egg cells, 14 to 15 zygotes at 12 HAP, 16 to 17 zygotes at 24 HAP, 16 apical cells, and 13 to 14 basal cells (Table 1). All three biological replicates of each cell type were used for RNA-Seq and subsequent transcriptome analyses.

**Sperm cell isolation**

Sperm cells were isolated as described (Dupuis et al., 1987) with some modifications. Pollen grains were collected upon shedding, immersed in 550 mOsmol·kg\(^{-1}\) H\(_2\)O mannitol solution (100 mg pollen/ml solution), and incubated on a platform shaker with slow agitation (80 rpm) for 1 h. The resulting lysate was filtered through a 40 μm cell strainer to remove exines and unruptured pollen grains, resulting in a yellowish filtrate containing sperm cells and starch granules. A Percoll gradient was prepared in a 30 ml COREX tube, consisting of 5 ml 30% (v/v) Percoll in 550 mOsmol·kg\(^{-1}\) H\(_2\)O mannitol solution at the bottom, 6 ml 20% (v/v) Percoll in 550 mOsmol·kg\(^{-1}\) H\(_2\)O mannitol solution in the middle, and 6 ml 15% (v/v) Percoll in 550 mOsmol·kg\(^{-1}\) H\(_2\)O mannitol solution at the top. Sperm-containing filtrate (10 ml) was layered on top of the Percoll gradient and centrifuged in a swing-out rotor at
12,000xg for 1 h at 4°C. After centrifugation, distinct white layers were visible in the 15/20% Percoll interphase and the 20/30% Percoll interphase. The 20/30% interphase, which was enriched in sperm cells, was carefully aspirated using a Pasteur pipette and transferred to 15 ml or 50 ml Falcon tubes. At least 10 volumes of fresh 550 mOsmol·kg⁻¹ H₂O mannitol solution were added to the sperm cell-enriched fraction, and the cells were washed by carefully inverting the tube several times. The sperm cells were pelleted by centrifugation at 2,500xg for 15 minutes at 4°C and the supernatant was removed without disturbing the pellet, leaving a volume of approximately 50–100 μl. The pellet was resuspended in the remaining supernatant by careful pipetting, resulting in a solution highly enriched in sperm cells. Cell counting was performed using a Neubauer counting chamber. Isolated sperm cells were used immediately or shock-frozen in liquid nitrogen and stored at -80°C.

**Isolation of apical and basal cells**

To identify the time point of asymmetric cell division (ACD) of the zygote, several cobs were pollinated and analyzed at different intervals after pollination. The first zygotes were analyzed at 24 HAP. Subsequent examinations were performed at 1-h intervals; on average, zygote ACD was observed at ~48 HAP. Apical and basal cells were subsequently separated using cell wall degrading enzyme solution containing 1.5% Driselase (Sigma), 1.5% pectinase (Fluka), 0.5% pectolyase Y23 (Karlan), 1.0% hemicellulase (Sigma), 1.0% cellulase “Onozuka R10” (Serva), and 1.5% maceroenzyme (Karlan) in mannitol solution (480 mOsmol·kg⁻¹ H₂O). The enzyme solution (100 μl) was combined with 1 ml mannitol solution (480 mOsmol·kg⁻¹ H₂O), and ovary sections containing embryo sacs at 48 HAP were incubated in the diluted enzyme solution for 30 min at room temperature, followed by manual dissection of two-celled proembryos. The attachment between the apical and basal cell protoplasts was gently touched with a very thin glass needle to separate both cells. The cells were washed twice in mannitol solution (480 mOsmol·kg⁻¹ H₂O), collected in 0.5 mL Eppendorf RNA/DNA LoBind microcentrifuge tubes, immediately frozen in liquid nitrogen, and stored at -80°C for mRNA isolation.

**RNA extraction, cDNA preparation and purification**
The mRNA was extracted from cell samples using a Dynabeads mRNA DIRECT Micro Kit (Life Technologies). A SMARTer Ultra Low RNA Kit for Illumina Sequencing (Clontech Laboratories) was used to generate first-strand cDNA. Double-stranded cDNA was amplified by LD PCR (15 cycles) and purified via magnetic bead cleanup using an Agencourt AMPure PCR Purification Kit (Beckman Coulter). The quality of the purified cDNA was measured using an Agilent 2100 Bioanalyzer with an Agilent High Sensitivity DNA Kit (Agilent Technologies), frozen in liquid nitrogen, and stored at −80°C.

**Library construction and Illumina sequencing**

Library preparation was carried out as described in the Adapted Nextera Sample Preparation protocol (Clontech Laboratories) for use with the SMARTer Ultra Low RNA Kit for Illumina Sequencing. Input cDNA (5 ng) was tagmented (tagged and fragmented) via the Nextera transposome. The products were purified and amplified via a limited-cycle PCR program to generate multiplexed sequencing libraries. The libraries were quantified using a KAPA SYBR FAST ABI Prism Library Quantification Kit (Kapa Biosystems). Equimolar amounts of each library were pooled and used for cluster generation on the cBot system with Illumina TruSeq PE Cluster v3 and Illumina TruSeq Dual Index Sequencing Primer paired end kits. Sequencing runs were performed on a HiSeq 1000 instrument using the dual indexed 2 x 100 cycles paired end (PE) protocol and TruSeq SBS v3 reagents according to the Illumina HiSeq 1000 System User Guide. Image analysis and base calling resulted in .bcl files, which were converted into .fastq files with CASAVA1.8.2 software. Library preparation and Illumina sequencing runs were performed at the Genomics Core Facility “KFB - Center for Fluorescent Bioanalytics” (www.kfb-regensburg.de). The raw data (.fastq) plus supplemental tables, including count and TPM data for all replicates was uploaded to GEO and is available under accession number GSE98379.

**Bioinformatic and statistical analyses**

(I) **Quality control and alignment:** The quality of sequencing data from the RNA-Seq libraries was assessed with FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Primer contamination from the SMARTer Ultra Low RNA Kit was removed with cutadapt (ver.1.13; Martin,
Size and quality trimming were then performed using Trimmomatic (ver. 0.32 (Am et al., 2014), ILLUMINACLIP: NexteraPE-PE_SMARTer.fa:2:30:10:2:true TRAILING:26 LEADING:26 MINLEN:25). The results from trimming are summarized in Supplemental Table 1 (mapping information). Trimmed reads were aligned to the maize genome (AGPv3, INSEC Assembly GCA_000005005.5, release 23, ftp://ftp.ensemblgenomes.org/pub/plants/). STAR (Dobin et al., 2013) was used to align the mRNA reads. Alignment statistics are summarized in Supplemental Table 1. Duplicate reads were identified with picard MarkDuplicates (http://broadinstitute.github.io/picard). The remaining reads were then assessed on the gene level using featureCounts from the Rsubread R-library (Liao et al., 2014) using the annotation information supplied by Gramene, release 5b+, (http://ftp.gramene.org/maizesequence.org/release5b+/zea_mays.protein_coding.gff). Gene annotation and ortholog information was retrieved from EnsemblPlants (www.plants.ensembl.org) via Biomart (BiomaRt7, release 29, Z. mays genes (AGPv3 (5b)), A. thaliana genes (2010-09-TAIR10) and O. sativa japonica genes (IRGSP-1.0), restricting the gene model to AGPv3 (5b) from EnsemblPlants.

(II) Measurement of mRNA transcript abundance: Transcripts per million (TPM) was used as a measure of mRNA abundance, which takes into account the length of the RNA transcripts and the sequencing depth (Wagner et al., 2012). Gene length was approximated by determining the sum of the exon lengths of the gene model. The TPMs were calculated for each gene in each sample. The median TPM value was calculated from three biological replicates for each cell type. An annotated list of genes with their expression levels at each developmental stage (cell type) and orthologs (based on Ensembl Compara from A. thaliana and O. sativa) was assembled (Supplemental Data Set 1).

(III) Differential expression analysis and Venn diagrams: Count data at the gene level were analyzed with DESeq2 (Love et al., 2014). All 15 cell-type-to-cell-type comparisons were performed and corrected for multiple testing over all genes and cell type comparisons using FDR (Benjamini and Hochberg, 1995) (Supplemental Data Set 2). Genes with a log2FoldChange>1 and FDR <0.05 (“padj” in supplemental tables) were considered significant differentially expressed, and Venn diagrams (R-library: vennerable; https://r-forge.r-project.org/projects/vennerable) were generated for certain comparisons (Figure 4).
Expression profiles and pathway analysis

Genes showing differential expression (log$_2$FoldChange$>1$ and FDR $<0.05$) in at least one comparison were subjected to further analyses. To visualize gene expression values in heatmaps and to compare gene expression profiles during different stages of very early plant development, square root transformed TPM values were utilized. The median expression values were then transformed to standard units to follow the expression values of one gene across different developmental stages in the heatmaps. The dendrogram of the samples is based on Euclidean distances combined with hierarchical clustering with complete linkage. To classify genes into specific expression profiles, Pearson correlation analysis of the gene expression vector (square root transformed median TPM values as above) versus a binary vector encoding the different expression profiles, i.e., expressed only in egg cells and during fertilization (0,1,1,1,0,0)/(SP,EC,ZY12,ZY24,AC,BC) was performed (Supplemental Data Set 1). A stringent cutoff value of $>0.9$ (Pearson correlation coefficient) was used to define a positive correlation with the respective profile.

Comparison of gene expression levels in maize and rice gametes

To compare the expression of the 80 most highly expressed genes from maize sperm and egg cells with their orthologs in rice (Oryza sativa ssp. japonica), data from Ensembl Compara (http://plants.ensembl.org, based on protein information for 53 species) and the Rice Genome Annotation project (RGAP) (http://rapdb.dna.affrc.go.jp/download/archive/RAP_MSU_2016-08-05.txt.gz) and orthologs from RGAP based on OrthoMCL were used (Li et al., 2003; Ouyang et al., 2007; Vilella et al., 2009). While OrthoMCL is based only on a comparison of protein sequences, Ensembl Compara also uses information from phylogenetic trees and therefore interprets sequence similarities based on the evolutionary development of a gene. Rice gene expression data (TPM) were obtained from Anderson et al. (2013). Maize and rice TPMs were square root transformed, and plastid transcripts were removed. Then, for each organism, the data were binned into 200 equally spaced expression categories, with the 99th percentile representing the maximum expression level. The color scale for the heatmaps shown in Figure 2 represents the expression bin, indicating the relative expression level in the organism shown. For each maize gene, the respective rice orthologs in both resources were identified and
compared. If there were common orthologs predicted by both resources, these were chosen, and if not, all predictions were used. From this selected set of orthologs, the most highly expressed ortholog in rice was used for plotting. All orthologs are listed in Supplemental Data Set 4. In the heatmaps (Figure 2), the maximum of the color scale represents the 99th percentile, resulting in the same color for all maize genes shown (because all of the top 80 genes are within the most highly expressed 1% of the genes). Dark red bars in the rice column indicate that the gene was also in the top 1% most highly expressed genes in the rice data, and the lighter color indicates lower expression. Cyan lines represent expression levels as bar charts, with the dotted line indicating the median value of the column.

Transcription factor gene list

A list of potential TFs in maize based on conserved domains at the protein level was retrieved from Grassius (http://grassius.org/tfomecollection.html, Maize_TFome_Bulk_data.txt). A list of median expression values (TPM>1) of all TF genes, including their gene names and TF families, is shown in Supplemental Data Set 7. Important TF classes were identified by analyzing the fraction of expressed genes (TPM>1 in percent). Whenever the fraction of expressed genes of one cell type in one TF class exceeded the fractions in the other cell types, this class was discussed in the main text.

Cell cycle gene list

A list of cell cycle regulators identified in Arabidopsis (Vandepoele et al., 2002) was used to retrieve maize orthologs from the Rice Genome Annotation Project Database (University of Michigan), also comprising a full data set for maize (Kawahara et al., 2013) (http://rice.plantbiology.msu.edu/annotation_pseudo_apk.shtml). Predicted cyclins (Hu et al., 2010) were added to the list, as well as genes reported to be involved in cell-cycle regulation in previous publications (Sauter et al., 1998; Dresselhaus et al., 1999b; Dresselhaus et al., 2006; Buendia-Monreal et al., 2011; Dante et al., 2014). Finally, the DNA replication-indicating genes encoding PCNA and MCM homologous proteins were added to the list manually based on Ensemble plants (www.plants.ensemble.org) (Sabelli et al., 2009). The genes were sorted into
different classes: CDKs and cyclins, followed by additional cell cycle-related factors (Supplemental Data Set 3).

Auxin pathway analysis

Information about auxin pathways was downloaded from (http://www.genome.jp/dbget-bin/www_bget?ath04075). Maize genes in the different categories were obtained from the literature based on the gene list Supplemental Data Set 3D (important gene families). For each category (i.e., SAURs), all significantly (p-adjusted <0.05) differentially expressed genes from the respective comparisons were selected. The median of the log$_2$FoldChanges of these genes was calculated and represented by a color scale ranging from green (-2) to red (2). All log2FoldChanges above or below these values were set to 2 (-2).

Accession Numbers:
The raw data (.fastq) plus supplemental tables, including count and TPM data for all replicates was uploaded to GEO and is available under accession number GSE98379. Gene identifiers are listed in Supplemental Data Set 1.

SUPPLEMENTAL DATA

Supplemental Figure 1. cDNA prepared from maize gametes/zygotes and initial validation based on the presence of selected transcripts.

Supplemental Figure 2. RNA-seq correlation plot and hierarchical clustering of maize gamete and zygote samples.

Supplemental Figure 3. Phylogenetic analysis of protein sequences of histone H3 and H2A variants.

Supplemental Table 1. Summary of NGS runs, alignment process to Ensembl genome (AGPv3, ver: 82.6) and annotation to Ensembl genebuild (AGPv3_5b, ver: 82.6).

Supplemental Data Set 1. Master table: binary identifier for venn diagram category membership. Median gene expression value (TPM, transcripts per million) of three biological replicates, for the indicated cell types, followed by genome coordinates, description, GO and InterPro annotation as well as information on homologous genes from Arabidopsis and rice.
Supplemental Data Set 2. List of differentially expressed genes for all 15 cell type comparisons, with abs(log2FC)>1 and p-adjusted <0.05.

Supplemental Data Set 3. Lists of transcript levels of known genes required for fertilization and zygote development (see Figure 1B), expression profiles of transcription factors, cell cycle genes, and important gene families (see Figures 5, 6, 7, and 8).

Supplemental Data Set 4. Lists of TOP80 maize sperm and egg cell genes compared to their predicted homologs/orthologs in rice gametes (see also Figure 2).

Supplemental Data Set 5. Lists of 30 most highly expressed genes in maize sperm cell, egg cell, zygote 12HAP, zygote 24HAP, apical cell and basal cell, respectively (for detailed annotation see Suppl. Data Set 1).

Supplemental Data Set 6. Lists of gene expression profiles selected for Figure 4E.

Supplemental Data Set 7. List of expressed transcription factor classes (see also Figure 5A and B).

Supplemental Data Set 8. List of genes in the auxin pathway (see also Figure 7).

ACKNOWLEDGMENTS

We thank Thomas Stempfl and Christoph Möhle from the Kompetenzzentrum Fluoreszent Bioanalytik (KFB), a genomics core facility based at the University of Regensburg, for their support in optimizing the procedure used to generate libraries for Illumina sequencing from only a few plant cells. Funding through the Collaborate Research Center SFB960 of the German Research Foundation (DFG) to TD, JCE, and SS is gratefully acknowledged. The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

TD initiated and designed the project. JC, NGK, and PC performed wet lab experiments and interpreted results. JC, NS, SS, and JCE performed bioinformatics analyses. JC, NS, and TD wrote the manuscript with input from all authors.
FIGURE LEGENDS

Figure 1. Time course of zygote development and validation of RNA-Seq data. (A) Sperm cells. (B) Egg cell. (C) Zygote at 24 hours after pollination (HAP). (D) Zygote at 30 HAP at anaphase. (E) Zygote at 35 HAP at telophase. Arrowheads indicate phragmoplast between daughter chromosome sets. (F) Zygote at 43 HAP during asymmetric cell division (ACD). (G) Zygote at 48 HAP. Cytokinesis is completed. Dotted line indicates the cell division plane. (H) Apical (AC) and basal cell (BC) have been separated at 54 HAP. Bright field microscopy images are shown on the left and UV images of DAPI stained cells on the right. Scale bars: 10 µm. (I) Validation of RNA-Seq data with known genes, as indicated. Top row: Genes preferentially expressed in sperm cells; middle row: Genes highly expressed in egg cells and downregulated after fertilization; bottom row: Expression of cell cycle regulators. Transcript levels are shown as TPM (transcripts per million) values (means ± SD) of three biological replicates.

Figure 2. Heatmap showing a comparison of the 80 most highly expressed genes in maize gametes and their predicted orthologs in the corresponding rice gametes. (A) TOP80 most highly expressed genes in maize sperm cells and (B) TOP80 genes in maize egg cells. Note that all maize genes display a TPM>200 and are thus indicated by red bars. For better visualization, rice and maize gene expression values (TPM) were square root transformed and classified into 80 expression categories using the 99th percentile of the data to summarize all higher expression values. Plastid genomes were excluded, as they showed overshooting of expression in the rice data. Non-expressed genes in the rice data and genes lacking a clear homolog are marked by black bars. Orthologous gene information is based on the EnsemblPlants Compara database, The Rice Genome Annotation Project (RGAP), and orthologs from RGAP based on OrthoMCL. Maize genes encoding histones and high-mobility group genes are shaded in purple, proteins involved in translation are in green, and EA1-box proteins and predicted secreted CRPs are in yellow. Proteins were classified according to InterPro.

Figure 3. Expression of major chromatin structure protein genes in maize gametes, zygotes, and daughter cells. (A) Histone H1, (B) histone H3, (C) histone
H4, (D) histone H2A, (E) histone H2B, and (F) high mobility group protein Hmgd1. All significantly expressed genes of the various gene families are presented. See Supplemental Figure 3 for phylogenetic analysis of male-specific and active chromatin-specific variants of histones H3 and H2A. Gene identifiers are listed in Supplemental Data Set 7 (important gene families). Transcript levels in the cells indicated are given as TPM values (means ± SD).

**Figure 4. Gene expression dynamics in maize gametes, zygotes, and two-celled pro-embryo cells.** Genes with a abs(log2FC)>1 and p-adjusted <0.05 were considered to be biologically meaningful. (A) Comparison of the number of genes induced in sperm cells (SP) and egg cells (EC) versus zygotes at 12 HAP (Zy12). (B) Comparisons of genes upregulated in zygotes at 12 HAP versus sperm and egg cells. (C) Number of genes induced in zygotes at 24 HAP compared to apical (AC) and basal (BC) cells. (D) Comparison of genes upregulated in apical and basal cells versus zygotes at 24 HAP (Zy24). (E) Selected gene expression profiles across the specific cell types analyzed (see Supplemental Data Set 6). The Pearson correlation (>0.9) of square root transformed TPM values per gene and binary profile vectors were used to identify genes belonging to a specific profile. The total number of genes including TFs contained in each profile is shown at the bottom. The mean expression level of all genes per specific cell type is plotted in each profile (black line). (F) Table of differentially upregulated genes (log2FC>1, p-adjusted<0.05) of row cell type versus column cell type. The last column show the number of differentially upregulated genes in a row cell type versus all other cell types.

**Figure 5. Expression levels of transcription factor (TF) and MAB genes in gametes, zygotes, and early two-celled pro-embryo cells in maize.** (A) Expression levels of TF genes in gametes and at early developmental stages after fertilization. Genes with abs(log2FC)>1 and p-adjusted <0.05 in at least one cell type comparison were considered. To make the expression levels comparable across multiple cell types, z-scores were calculated from the square root transformed TPM values corresponding to the number of standard deviations between the cell type-specific expression value and the mean expression value of the respective gene. Genes were ordered by TF classes (black/blue bars). Black and blue fonts were used alternatively to distinguish the various classes. See Supplemental Data Sets 2
and 7 for details. (B) Expression analysis of selected maize genes encoding homologs of Arabidopsis early embryogenesis-related TFs as well as (C) MATH-BTB (MAB) family genes involved in ACD.

Figure 6. Gene expression analysis of cell cycle regulators. (A) Major regulators of the maize cell cycle (Supplemental Data Set 2). Transformation of gene expression values as described above. Genes were ordered into protein classes based on data from the literature (see Methods for details). (B) Summary of the time course of pollen tube (PT) growth, fertilization, and zygote development. Cell cycle stages of the selected cells are indicated based on the currently reported expression data.

Figure 7. Analysis of the auxin pathway in gametes, zygotes, and during early embryogenesis. (A) Expression analysis of the most highly expressed auxin biosynthesis, transport, and auxin response-related genes. (B) Auxin pathway analysis using three developmental transitions during zygote development. The median values over the log₂FoldChanges of all differentially expressed genes (p-adjusted <0.05) from the respective comparison are color-coded in green (downregulated) and red (upregulated). All log₂FoldChanges above 2 or below –2 were set to 2/–2 to improve visualization. White boxes: no significant log₂FC found, +u: ubiquitination, blue line: inhibition, dotted gray line: dissociation, red dashed line: expression. Gene pathway based on KEGG analysis. See Supplemental Data Set 2 for genes.

Figure 8. Expression analysis of selected maize genes with putative roles in signaling during gamete interaction and early embryogenesis. (A) Ca²⁺-dependent phospholipid-binding annexin family proteins, (B) sperm-specific receptor-like kinases, and (C) fertilization-regulated receptor-like kinases. Gene identifiers are listed in Supplemental Data Set 2 (important gene families). Transcript levels in the cells indicated are given as TPM values (means ± SD).
Table 1. Summary of samples, NGS runs, alignment to the Ensembl genome (AGPv3, ver. 82.6), and annotation to Ensembl genebuild (AGPv3_5b, ver. 82.6). The number of genes expressed in at least two of three replicates are given. See Supplemental Table 1 for details.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Cell number per replicate</th>
<th>All reads after trimming</th>
<th>Aligned pairs mapped unique</th>
<th>% Mapped reads per trimmed reads</th>
<th>Expressed genes (TPM&gt;1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sperm Cell1</td>
<td>~1,000</td>
<td>68,988,450</td>
<td>26,514,492</td>
<td>77</td>
<td>11,819</td>
</tr>
<tr>
<td>Sperm Cell2</td>
<td>~1,000</td>
<td>102,929,736</td>
<td>38,689,716</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Sperm Cell3</td>
<td>~1,000</td>
<td>86,448,904</td>
<td>33,271,851</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Egg Cell1</td>
<td>20</td>
<td>57,338,078</td>
<td>13,901,359</td>
<td>48</td>
<td>16,026</td>
</tr>
<tr>
<td>Egg Cell2</td>
<td>20</td>
<td>63,488,722</td>
<td>22,978,462</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>Egg Cell3</td>
<td>20</td>
<td>69,199,124</td>
<td>25,147,141</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>Zygote_12HAP1</td>
<td>14</td>
<td>70,632,412</td>
<td>15,761,484</td>
<td>45</td>
<td>19,865</td>
</tr>
<tr>
<td>Zygote_12HAP2</td>
<td>15</td>
<td>76,840,694</td>
<td>19,297,245</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Zygote_12HAP3</td>
<td>15</td>
<td>55,130,854</td>
<td>11,323,005</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>Zygote_24HAP1</td>
<td>16</td>
<td>82,825,752</td>
<td>20,424,191</td>
<td>49</td>
<td>19,171</td>
</tr>
<tr>
<td>Zygote_24HAP2</td>
<td>16</td>
<td>80,911,418</td>
<td>20,173,448</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Zygote_24HAP3</td>
<td>17</td>
<td>87,134,202</td>
<td>19,443,455</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Apical Cell1</td>
<td>16</td>
<td>94,779,064</td>
<td>29,713,124</td>
<td>63</td>
<td>17,747</td>
</tr>
<tr>
<td>Apical Cell2</td>
<td>16</td>
<td>68,636,668</td>
<td>20,927,946</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>Apical Cell3</td>
<td>16</td>
<td>74,074,566</td>
<td>23,422,549</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>Basal Cell1</td>
<td>13</td>
<td>71,456,410</td>
<td>22,440,462</td>
<td>63</td>
<td>18,069</td>
</tr>
<tr>
<td>Basal Cell2</td>
<td>13</td>
<td>68,963,988</td>
<td>21,238,495</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>Basal Cell3</td>
<td>14</td>
<td>67,950,552</td>
<td>20,886,111</td>
<td>61</td>
<td></td>
</tr>
</tbody>
</table>

REFERENCES


COTYLEDON1 is sufficient to induce embryo development in vegetative cells.

Cell 93, 1195-1205.


Schreiber, D.N., Bantin, J., and Dresselhaus, T. (2004). The MADS box transcription factor ZmMADS2 is required for anther and pollen maturation in


Figure 1. Time course of zygote development and validation of RNA-Seq data. (A) Sperm cells. (B) Egg cell. (C) Zygote at 24 hours after pollination (HAP). (D) Zygote at 30 HAP at anaphase. (E) Zygote at 35 HAP at telophase. Arrowheads indicate phragmoplast between daughter chromosome sets. (F) Zygote at 43 HAP during asymmetric cell division (ACD). (G) Zygote at 48 HAP. Cytokinesis is completed. Dotted line indicates the cell division plane. (H) Apical (AC) and basal cell (BC) have been separated at 54 HAP. Bright field microscopy images are shown on the left and UV images of DAPI stained cells on the right. Scale bars: 10 µm. (I) Validation of RNA-Seq data with known genes, as indicated. Top row: Genes preferentially expressed in sperm cells; middle row: Genes highly expressed in egg cells and downregulated after fertilization; bottom row: Expression of cell cycle regulators. Transcript levels are shown as TPM (transcripts per million) values (means ± SD) of three biological replicates.
Figure 2. Heatmap showing a comparison of the 80 most highly expressed genes in maize gametes and their predicted orthologs in the corresponding rice gametes. (A) TOP80 most highly expressed genes in maize sperm cells and (B) TOP80 genes in maize egg cells. Note that all maize genes display a TPM>200 and are thus indicated by red bars. For better visualization, rice and maize gene expression values (TPM) were square root transformed and classified into 80 expression categories using the 0.99th percentile of the data to summarize all higher expression values. Plastid genomes were excluded, as they showed overshooting of expression in the rice data. Non-expressed genes in the rice data and genes lacking a clear homolog are marked by black bars. Orthologous gene information is based on the EnsemblPlants Compara database, The Rice Genome Annotation Project (RGAP), and orthologs from RGAP based on OrthoMCL. Maize genes encoding histones and high-mobility group genes are shaded in purple, proteins involved in translation are in green, and EA1-box proteins and predicted secreted CRPs are in yellow. Proteins were classified according to InterPro.
Figure 3. Expression of major chromatin structure protein genes in maize gametes, zygotes, and daughter cells. (A) Histone H1, (B) histone H3, (C) histone H4, (D) histone H2A, (E) histone H2B, and (F) high mobility group protein Hmgd1. All significantly expressed genes of the various gene families are presented. See Supplemental Figure S3 for phylogenetic analysis of male-specific and active chromatin-specific variants of histones H3 and H2A. Gene identifiers are listed in Supplemental Table S8 (important gene families). Transcript levels in the cells indicated are given as TPM values (means ± SD).
Figure 4. Gene expression dynamics in maize gametes, zygotes, and two-celled pro-embryo cells. Genes with an abs(log2FC)>1 and p-adjusted <0.05 were considered to be biologically meaningful. (A) Comparison of the number of genes induced in sperm cells (SP) and egg cells (EC) versus zygotes at 12 HAP (Zy12). (B) Comparisons of genes upregulated in zygotes at 12 HAP versus sperm and egg cells. (C) Number of genes induced in zygotes at 24 HAP compared to apical (AC) and basal (BC) cells. (D) Comparison of genes upregulated in apical and basal cells versus zygotes at 24 HAP (Zy24). (E) Selected gene expression profiles across the specific cell types analyzed (see Supplemental Table S6). The Pearson correlation (>0.9) of square root transformed TPM values per gene and binary profile vectors were used to identify genes belonging to a specific profile. The total number of genes including TFs contained in each profile is shown at the bottom. The mean expression level of all genes per specific cell type is plotted in each profile (black line). (F) Table of differentially upregulated genes (log2FC>1, p-adjusted<0.05) of row cell type versus column cell type. The last column show the number of differentially upregulated genes in a row cell type versus all other cell types.
Figure 5. Expression levels of transcription factor (TF) and MAB genes in gametes, zygotes, and early two-celled pro-embryo cells in maize. (A) Expression levels of TF genes in gametes and at early developmental stages after fertilization. Genes with abs(log2FC)>1 and p-adjusted <0.05 in at least one cell type comparison were considered. To make the expression levels comparable across multiple cell types, z-scores were calculated from the square root transformed TPM values corresponding to the number of standard deviations between the cell type-specific expression value and the mean expression value of the respective gene. Genes were ordered by TF classes (black/blue bars). Black and blue fonts were used alternatively to distinguish the various classes. See Supplemental Tables S7 and S8 for details. (B) Expression analysis of selected maize genes encoding homologs of Arabidopsis early embryogenesis-related TFs as well as (C) MATH-BTB (MAB) family genes involved in ACD.
Figure 6. Gene expression analysis of cell cycle regulators. (A) Major regulators of the maize cell cycle (Supplemental Table S7). Transformation of gene expression values as described above. Genes were ordered into protein classes based on data from the literature (see Methods for details). (B) Summary of the time course of pollen tube (PT) growth, fertilization, and zygote development. Cell cycle stages of the selected cells are indicated based on the currently reported expression data.
Figure 7. Analysis of the auxin pathway in gametes, zygotes, and during early embryogenesis. (A) Expression analysis of the most highly expressed auxin biosynthesis, transport, and auxin response-related genes. (B) Auxin pathway analysis using three developmental transitions during zygote development. The median values over the log2FoldChanges of all differentially expressed genes (p-adjusted <0.05) from the respective comparison are color-coded in green (downregulated) and red (upregulated). All log2FoldChanges above 2 or below –2 were set to 2/–2 to improve visualization. White boxes: no significant log2FC found, +u: ubiquitination, blue line: inhibition, dotted gray line: dissociation, red dashed line: expression. Gene pathway based on KEGG analysis. See Supplementary Table 7 for genes.
Figure 8. Expression analysis of selected maize genes with putative roles in signaling during gamete interaction and early embryogenesis. (A) Ca\textsuperscript{2+}-dependent phospholipid-binding annexin family proteins, (B) sperm-specific receptor-like kinases, and (C) fertilization-regulated receptor-like kinases. Gene identifiers are listed in Supplemental Table S7 (important gene families). Transcript levels in the cells indicated are given as TPM values (means ± SD).


Zhao, P., and Sun, M.X. (2015). The maternal-to-zygotic transition in higher plants: Available approaches, critical limitations, and


Zygotic Genome Activation Occurs Shortly After Fertilization in Maize
Junyi Chen, Nicholas Strieder, Nadia G. Krohn, Philippe Cyprys, Stefanie Sprunck, Julia C. Engelmann and Thomas Dresselhaus
Plant Cell; originally published online August 16, 2017;
DOI 10.1105/tpc.17.00099

This information is current as of August 29, 2017