Maternal epigenetic pathways control parental contributions to Arabidopsis early embryogenesis.

Daphné Autran, Célia Baroux, Michael T Raissig, Thomas Lenormand, Michael Wittig, Stefan Grob, Andrea Steimer, Matthias Barann, Ulrich C Klostermeier, Olivier Leblanc, et al.

Cell

Maternal Epigenetic Pathways Control Parental Contributions to *Arabidopsis* Early Embryogenesis

Daphné Autran,[1,6] Célia Baroux,[2,6] Michael T. Raissig,[2] Thomas Lenormand,[3] Michael Wittig,[4] Stefan Grob,[2] Andrea Steimer,[2] Matthias Barann,[4] Ulrich C. Klostermeier,[4] Olivier Leblanc,[1] Jean-Philippe Vielle-Calzada,[5] Phillip Rosenstiel,[4] Daniel Grimanelli,[1,*] and Ueli Grossniklaus[2,*]

[1]Diversité, Adaptation et Développement des Plantes, Institut de Recherche pour le Développement, Université de Montpellier, UMR 232, Montpellier 34394, France

[2]Institute of Plant Biology and Zürich-Basel Plant Science Center, University of Zürich, Zürich CH-8008, Switzerland

[3]Centre d'Ecologie Fonctionnelle et Evolutive, CNRS, Université de Montpellier, UMR 5175, Montpellier 34293, France

[4]Institute of Clinical Molecular Biology, Christian Albrechts University Kiel, Kiel 24105, Germany

[5]National Laboratory of Genomics for Biodiversity, CINVESTAV Guanajuato, Irapuato CP36500, Guanajuato, Mexico

[6]These authors contributed equally to this work

*Correspondence: daniel.grimanelli@ird.fr (D.G.), grossnik@botinst.uzh.ch (U.G.)

DOI 10.1016/j.cell.2011.04.014

SUMMARY

Defining the contributions and interactions of paternal and maternal genomes during embryo development is critical to understand the fundamental processes involved in hybrid vigor, hybrid sterility, and reproductive isolation. To determine the parental contributions and their regulation during *Arabidopsis* embryogenesis, we combined deep-sequencing-based RNA profiling and genetic analyses. At the 2–4 cell stage there is a strong, genome-wide dominance of maternal transcripts, although transcripts are contributed by both parental genomes. At the globular stage the relative paternal contribution is higher, largely due to a gradual activation of the paternal genome. We identified two antagonistic maternal pathways that control these parental contributions. Paternal alleles are initially downregulated by the chromatin siRNA pathway, linked to DNA and histone methylation, whereas transcriptional activation requires maternal activity of the histone chaperone complex CAF1. Our results define maternal epigenetic pathways controlling the parental contributions in plant embryos, which are distinct from those regulating genomic imprinting.

INTRODUCTION

In most animal species, the zygote is transcriptionally quiescent, and early embryogenesis is governed by maternal products stored in the oocyte prior to fertilization (Andéol, 1994). Depending on the species, zygotic genome activation (ZGA) takes place after one to several cell divisions. ZGA is a gradual process that relies on large-scale chromatin reprogramming leading to an increasing number of zygotically expressed genes (Tadros and Lipshitz, 2009). Maternal transcripts and proteins inherited from the gametes are progressively degraded, and biparental zygotic transcripts gradually take over the control of development. As a result, parental contributions to the embryonic transcriptome dynamically change during early development, with an initial maternal control that is of variable duration (1 to 15 cell cycles) (Baroux et al., 2008; Tadros and Lipshitz, 2009). Strikingly, such a maternal influence occurs in animals as evolutionarily divergent as insects, amphibians, and mammals. Understanding the parental contributions and the regulation of zygotic genome expression during early embryogenesis is a key question in developmental and evolutionary biology.

In flowering plants, the knowledge about the regulation and dynamics of parental contributions during early embryogenesis remains fragmented, despite its importance in understanding hybrid vigor, hybrid viability, parent-of-origin-dependent interploidy, and nonself pollination (xenia) effects determined by interactions of parental genomes after fertilization (Bushell et al., 2003; Jahnke et al., 2010; Meyer and Scholten, 2007; Pahlavani and Abolhasani, 2006). In flowering plants, double fertilization produces the zygote, which develops into the embryo (Movie S1 available online), and the endosperm, an embryo-nurturing tissue. Both fertilization products develop within maternal integuments, forming the seed. Genetic studies have shown that seed development is under maternal influence (Chaudhury and Berger, 2001), but the composite nature of the seed makes determining the origin of maternal effects complex. Recently, downregulation of RNA Polymerase II (PolII) in the mature *Arabidopsis* egg cell revealed that the embryo developed to the preglobular stage in absence of significant de novo transcription (Pillot et al., 2010b). Thus, de novo transcription of parental genomes is not an absolute requirement for early embryogenesis, but the timing, dynamics, and mechanisms of zygotic genome activation have yet to be elucidated. Reporter and profiling studies on whole seeds from *Arabidopsis* and maize have identified several transcripts with a dominant maternal representation at early stages, whereas paternal transcripts were detected only later (Baroux et al., 2001; Grimanelli et al.,

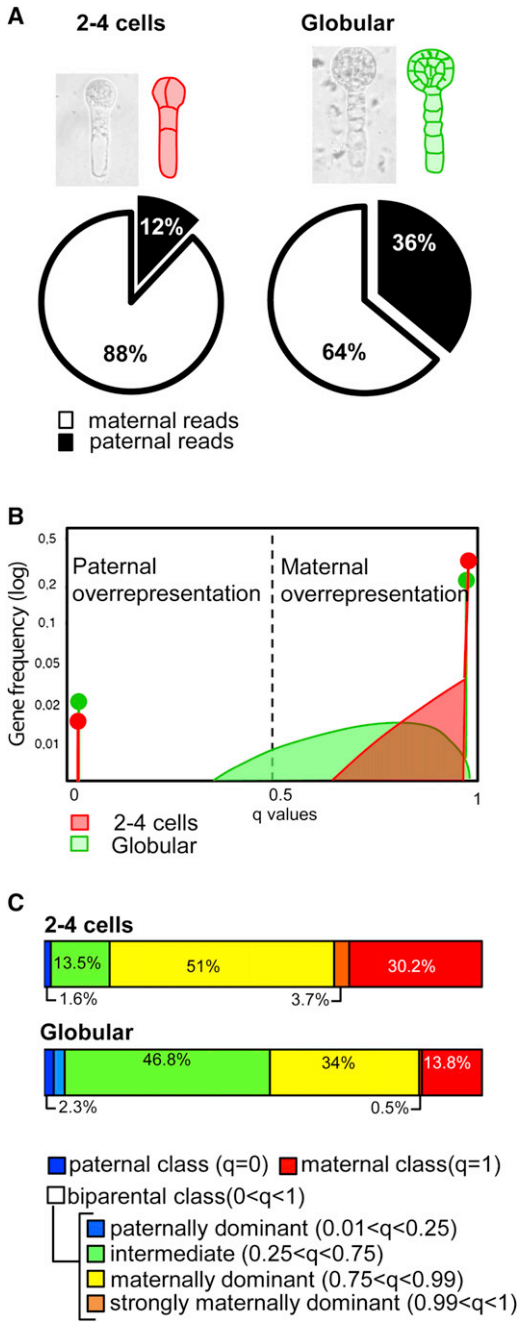


Figure 1. Parental Contributions to the *Arabidopsis* Early Embryo Transcriptome

(A) Distribution of maternal versus paternal reads covering the embryonic transcriptome of isolated embryos. Embryos were derived from a cross between polymorphic parents. Informative SOLiD reads (Table S1) mapping to known SNPs were assigned to the maternal (*Ler*) or paternal (Col) parent.

(B) Informative reads identified 3973 and 3078 genes in the 2–4 cell and globular transcriptomes, respectively. The graph shows a likelihood-based gene distribution according to the proportion of maternal transcripts, q. q = 1 and q = 0 represent genes contributed only maternally or paternally, respectively. Genes with 0 < q < 1 are contributed biparentally. y axis: proportion of genes (log scale) ; x axis:q values along 1% quantiles. The full-colored circles indicate extreme quantiles ($0 \leq q < 0.01$ and $0.99 < q \leq 1$).

(C) Composite diagram representation of the gene distribution as drawn in (B) according to q intervals as labeled. Related to Figure S1, Table S1, and Table S2.

2005; Vielle-Calzada et al., 2000). In contrast, a biparental expression in the zygote and early embryo was shown for certain other genes (Aw et al., 2010; Meyer and Scholten, 2007; Ronceret et al., 2005, 2008; Scholten et al., 2002; Weijers et al., 2001). Furthermore, near-saturation mutagenesis screens identified a plethora of mutations affecting embryo development at or before the globular stage. The majority of these segregate as zygotic recessive traits, indicating biparental contributions to early embryogenesis (Tzafrir et al., 2004). Thus, to date there is no clear understanding of the relative parental contributions to plant embryogenesis nor are the mechanisms regulating the respective contributions known.

We performed allele-specific profiling of the embryonic transcriptome and quantified relative transcript contributions of the paternal and maternal genomes in *Arabidopsis* embryos. We demonstrate a strong, genome-wide dominance of maternal transcripts at early stages, although many transcripts are biparentally represented. This finding reconciles the observation made by several laboratories of both maternal and zygotic effects during early embryogenesis. Using reporter and profiling analyses, we found an increasing contribution of paternal products as embryo development proceeds, with kinetics differing on a gene-by-gene basis. We identified two maternal epigenetic pathways, involving the chromatin siRNA pathway and the histone chaperone complex CAF1, which act antagonistically to regulate the paternal contribution. Importantly, we show that these pathways are distinct from those regulating genomic imprinting.

RESULTS

Analysis of Parental Contributions in Isolated *Arabidopsis* Embryos

To unambiguously define the parental contributions in the early embryo, we profiled the transcriptome of early-stage embryos in an allele-specific manner. We dissected 2–4 cell and globular stage embryos (Figure 1A) derived from a cross between the polymorphic Landsberg *erecta* (*Ler*) and Columbia (Col) accessions. The embryonic transcriptome was sequenced using a SOLiD v3 platform (Figure 1; see Table S1). Reads covering single-nucleotide polymorphisms (SNPs) (Borevitz et al., 2007) were extracted and are referred to hereafter as informative reads with respect to parental origin. Informative reads were used to quantify parental contributions both globally (Figure 1A), and for each gene individually. Genes were considered biparentally expressed at a given stage whenever both parental alleles were identified in the corresponding transcriptome. The classification of genes for which only a single allelic variant was detected was more ambiguous. Transcripts might be detected from only one parent because of uniparental expression or because of low sampling rates, which is of particular concern for genes expressed at low levels. To circumvent this difficulty, we made a probabilistic model describing the distribution of genes according to the proportion of maternal transcripts, q,

which was adjusted to best-fit the observations (Experimental Procedures). From this likelihood-based distribution, we estimated the frequency of genes fitting the biparental class ($0 < q < 1$, both parental alleles were detected), the paternal class ($q = 0$, only the paternal alleles were detected), or the maternal class ($q = 1$, only the maternal alleles were detected). In addition, each gene could be assigned a probability of falling within each class (Table S2). This integrated approach allowed a quantitative and qualitative analysis of parental contributions to the embryonic transcriptome.

The Transcriptome of 2–4 Cell Embryos Is Maternally Dominant despite Significant Contributions from Both Parental Genomes

Strikingly, at the 2–4 cell embryo stage 88.4% of the informative reads (n = 135,142) were of maternal origin (Figure 1A). This was confirmed in an independent biological replicate (Figure S1). The informative reads represented 3973 loci located throughout the genome. The gene distribution drawn for q quantiles showed a strong bias toward maternal overrepresentation (Figure 1B) with 85% of the genes described by $q > 0.75$ (Figure 1C; Table S1). Transcripts of the biparental class ($0 < q < 1$; 68.2% of the identified genes) contributed strongly to this maternal dominance with 54.7% genes described by $0.75 < q < 1$, i.e., maternal overrepresentation (Figure 1C and Table S1). Furthermore, our analysis revealed 30.2% transcripts of the maternal class ($q = 1$) against only 1.6% transcripts of the paternal class ($q = 0$) at the 2–4 cell stage. Thus, in *Arabidopsis* both parental genomes contribute to the early embryonic transcriptome but overall it is clearly dominated by maternal transcripts.

The Paternal Contribution Is Higher at the Globular Stage Concomitant with a Gradual Activation of Paternal Alleles

To determine how parental transcript contributions change during embryogenesis, we extended our allele-specific profiling to embryos at the globular stage. Although maternal dominance was maintained, the paternal contribution increased, as shown by 35.9% paternal reads versus 11.6% at the 2–4 cell stage (Figure 1A and Table S1). These informative reads identified 3078 loci, for which the q distribution remained skewed toward maternal overrepresentation, although to a lesser extent than at the 2–4 cell stage (Figure 1B). The maternal class ($q = 1$) represented 13.8% of the genes (versus 30.2% at the 2–4 cell stage), and the paternal class ($q = 0$) increased marginally to 2.3% (versus 1.6% at the 2–4 cell stage) (Table S1). Concomitantly, the biparental class increased, now comprising 83.9% of genes, with 34.5% showing maternal overrepresentation ($0.75 < q < 1$) (versus 54.7% at the 2–4 cell stage) (Figure 1C). Importantly, the globular stage transcriptome shared 2417 genes (78.5%) with that of the 2–4 cell stage, representing 95% of the reads (Figure 2A). We analyzed the changes of parental contributions among these shared genes by quantifying class transitions (Figure 2B). For instance a transition from the maternal class to the biparental or paternal class indicates de novo activation of the paternal allele and represents 21.5% (515) of the genes (Figure 2C). De novo activation of the maternal allele was less prominent with only 0.8% genes, mostly because few paternally expressed genes were identified at the 2–4 cell stage. This analysis identified loci with a decay of one parental transcript, with 2.5% and 9.7% showing loss of their maternal or paternal transcripts, respectively (Figure 2C). However, most quantitative changes occurred in the biparentally expressed class (54.4%, 1315 genes) where the majority showed a marked increase in the relative paternal contribution (778 genes, Figure 2D). This could result from decay of maternal RNAs, de novo transcription of paternal alleles, or a combination of both.

To refine the timing of activation of paternal alleles, we monitored paternal activity of six marker lines expressing a reporter gene under diverse promoters active during early embryogenesis. The lines reflect genes with diverse cellular functions (Table S3) and showed either early, intermediate, or late paternal activity, respectively (e.g., *RPS5A*, *CYCB1*;1, and *ET1041*), but did not appear in our allele-specific transcriptome due to the absence of a referenced SNP in their sequence. The marker lines clearly showed distinct expression depending on whether they were maternally or paternally inherited (Figure 3A and Figure S2A). We scored the number of F1 embryos showing paternal marker expression at the same developmental stage and in a wild-type maternal background (three to seven biological replicates each, Table S4). For all markers the proportion of stained embryos increased with developmental progression (Figure 3B and Figure S2B). Consistent with our RNA profiling results, paternal expression of the markers displayed gene-specific activation timing (developmental stage at which the first expression was detected) and kinetics (incremental increase in the fraction of progeny showing expression). For instance, in a L*er* maternal background the paternal *ET1041* marker showed only 4% stained embryos at the 2–4 cell stage, whereas the *RPS5A* and *GRP23* markers showed 58% and 67%, respectively (Figure 3B). In contrast, maternally transmitted markers showed consistent expression in essentially all embryos, even at earliest stages (Figures S2A and S2C).

Taken together with the transcriptome study, these findings strongly suggest that paternally inherited alleles—even those that are detectable at a very early stage—are activated gradually after fertilization. This is consistent with the previous observations made for individual genes in *Arabidopsis*, and reconciles earlier, apparently conflicting reports. Whether maternal loci follow similar or different activation kinetics cannot be easily resolved because of the potential importance of maternal carryover. Nevertheless, our observations are reminiscent of the gradual, de novo, expression of zygotic genes reported in animals (Tadros and Lipshitz, 2009).

Maternal KRYPTONITE Activity Controls Paternal Contribution in the Early Embryo

In our marker analyses we observed that expression was influenced by the maternal genotype (i.e., the accession) (Figure S2B). This indicated a maternal control of paternal expression as suggested previously (Ngo et al., 2007). The gradual increase in paternal allele expression might reflect the progressive release of a silencing mechanism. In *Arabidopsis*, silent chromatin is enriched in histone H3 dimethylated at lysine 9 (H3K9me2), a modification principally deposited by the SUVH4 histone methyltransferase KRYPTONITE (KYP) (Jackson et al.,

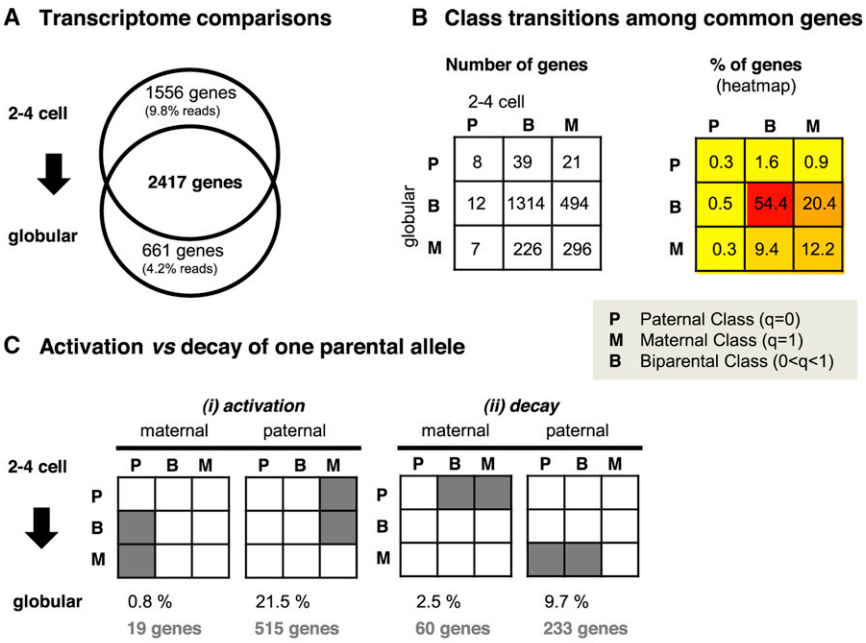


Figure 2. Dynamic Changes of Parental Contributions during Embryo Development

(A) Venn diagrams showing the number of genes identified by informative reads and shared between 2–4 cell and globular embryo transcriptomes. Global coverage for specific genes is indicated in brackets. Common genes are covered by the majority of reads (90%–95%).

(B) The transition tables describe the changes for common genes in the parental class (P, B, M, see legend) that occurred during development (2–4 cell → globular transition).

(C) Subsets of class transitions illustrate dynamic changes in allele representation as indicated (activation/de novo expression and decreased expression/decay of one parental allele) during developmental progression. Note that "decay" may correspond to a decrease in SNP coverage falling below detection threshold rather than an absolute loss of transcript. The % are the sum of the % genes in (B) falling into the gray transitions.

(D) A vast majority of common genes (54.4%) is biparentally represented at both stages. The proportion of maternal transcripts (q) was calculated for 1125 common genes sequenced on both alleles and plotted as indicated (left). The interpretation of relative changes toward higher paternal or maternal representation is shown (right). The inset shows the number of genes with changes in q values > 10% (dashed lines). A total of 245 genes showed no or <10% change. Related to Table S1 and Table S2.

2002). To investigate the possibility that maternal KYP regulates the activity of paternal alleles, we quantified reporter gene expression in seeds resulting from a cross between a maternal *kyp* mutant and wild-type pollen carrying the marker. For all reporters tested, lack of maternal KYP activity significantly increased the proportion of embryos showing early paternal reporter activity (before the 16-cell embryo stage) (Figures 4A and 4B, Figure S2A, and Table S4). We confirmed the maternal effect of the *kyp* mutation on the endogenous *AGP18* locus by allele-specific RT-PCR (Figure 4C). These data strongly suggest a role for maternal KYP activity in repressing the transcription of paternal alleles during early embryogenesis.

To investigate the maternal effect of the *kyp* mutation at the genome-wide level, we performed allele-specific profiling of 2–4 cell embryos dissected from crosses between maternal *kyp* (*Ler*) and paternal wild-type (Col) parents (Table S1). Embryos inheriting maternal *kyp* (kyp^m/KYP^p) showed a strong increase in the proportion of paternal reads (35.9% versus 11.6% in 2–4 cell stage wild-type embryos) (Figure 4D and replicate Figure S1) resulting in a similar paternal contribution as in wild-type embryos at the globular stage (36.1%; Figure 1A). Informative reads in kyp^m/KYP^p embryos identified 3125 genes (Table S1) for which the q distribution was shifted toward a higher paternal representation compared to wild-type embryos at the 2–4 cell stage (Figure 4E). Notably, the proportion of the paternal class increased markedly (10.3% at q = 0, versus 1.6% in wild-type 2–4 cell embryos, Figure 4F) whereas the proportion of genes in the maternal class diminished (Figure 4F). In addition, genes in the biparental class ($0 < q < 1$) showed a higher paternal contribution compared to wild-type embryos at the 2–4 cell stage (shifted distribution Figure 4E; less maternally dominant genes Figure 4F). Thus, the maternal *kyp* mutation affected a large number of genes throughout the genome, resulting in a higher paternal contribution to the early embryonic transcriptome.

Importantly, 2461 of these 3125 genes were common to the wild-type transcriptome of embryos at the same stage and were covered by 94.6% of the informative reads (Figure S3B). This suggests that the maternal *kyp* mutation does not drastically alter the 2–4 cell stage embryonic transcriptome, but instead modifies the relative parental contributions of genes normally expressed at this stage. Interestingly, the maternal *kyp* mutation induced class transitions similar to those induced by developmental progression (Figures S3C–S3E compared to Figures 2B–2D). The changes in the relative contribution of

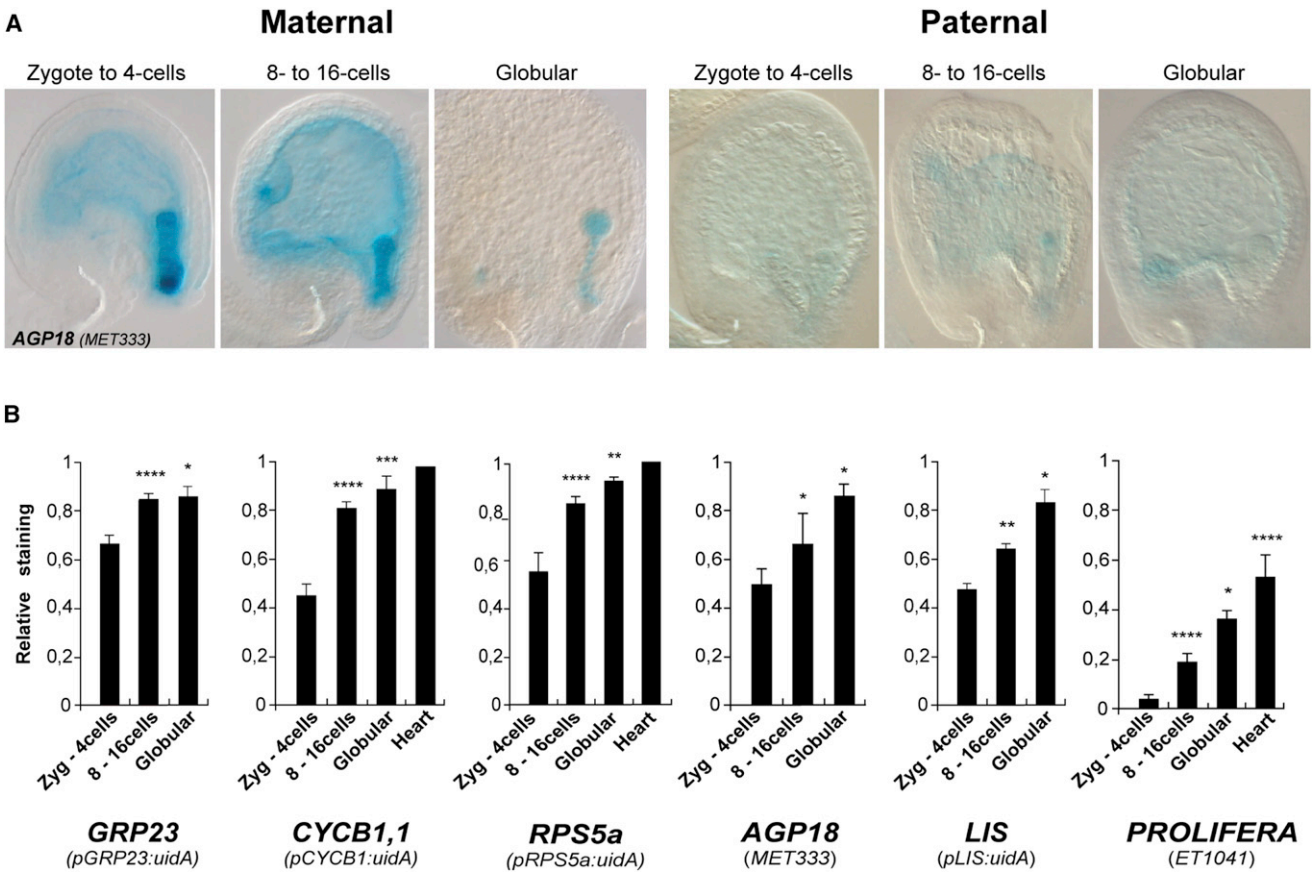


Figure 3. Gradual Activation of Paternal Markers during Early Embryo Development

(A) Representative panel showing differential expression of the marker tested (Table S3 and Table S4), here *MET333* reporting *AGP18* expression, when transmitted maternally (left) or paternally (right), as monitored by histochemical detection of GUS (blue substrate).

(B) Expression of the paternal markers was scored as the proportion of embryos showing GUS staining at a given developmental stage in a wild-type maternal background. Two-tailed Fisher's exact tests were used to assess differences between two consecutive developmental classes: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$. Error bars represent standard error between independent biological replicates.

Related to Figure S2, Table S3, and Table S4.

biparental class genes were highly correlated (Figure 4G), except for a subgroup (158 genes) showing a higher increase of the paternal contribution in kyp^{m}/KYP^{p} 2–4 cell embryos as compared to wild-type globular embryos (Figure 4G). The analysis also identified genes showing a gain or a loss of one parental transcript (Figure S3D). We conservatively estimate that the maternal *kyp* mutation induces an increased paternal and maternal contribution for 1307 and 117 genes, respectively (Figure S3F).

Taken together these results strongly suggest that maternal KYP activity downregulates early transcription of many paternal, and some maternal, alleles of loci throughout the genome. This maternal control may be progressively released during embryogenesis, leading to the gradual activation of the paternal genome.

The Chromatin siRNA Pathway Maternally Controls the Paternal Contribution to the Early Embryo

Because KYP-dependent H3K9me2 in *Arabidopsis* is linked to non-CG DNA methylation, we verified the effects of mutations in *DOMAIN REARRANGED METHYLTRANSFERASE2* (*DRM2*) and *CHROMOMETHYLASE3* (*CMT3*), two genes that control non-CG methylation (Feng et al., 2010). Similar to *kyp*, both mutations inherited maternally resulted in precocious activation of the paternally inherited markers (Figure 5A and Figures S4A and S4B). By contrast, maternal mutations in either *METHYLTRANSFERASE1* (*MET1*) or *DECREASED DNA METHYLATION1* (*DDM1*), involved in the maintenance of CG DNA methylation (Feng et al., 2010), had no consistent effect (Figure S4C). Thus, non-CG but not CG DNA methylation participates in the transcriptional repression of paternal marker genes.

In *Arabidopsis*, DNA and histone methylation can be mediated by small-interfering RNAs (siRNAs) via the chromatin siRNA pathway, also known as the RNA-directed DNA methylation (RdDM) pathway (Brodersen and Voinnet, 2006). To determine whether the RdDM pathway plays a role in the maternal repression of paternal alleles, we tested mutations in the following RdDM components: *NRPD1a* (PolIV), *NRPD1b* (PolV/*NRPE1*), *DCL3*, *RDR2*, and *AGO4* (Table S3). When inherited maternally, all these mutations allowed an earlier and stronger detection of paternally transmitted markers (Figure 5A). By contrast, the *dcl2-1* mutant, which affects a distinct siRNA-dependent

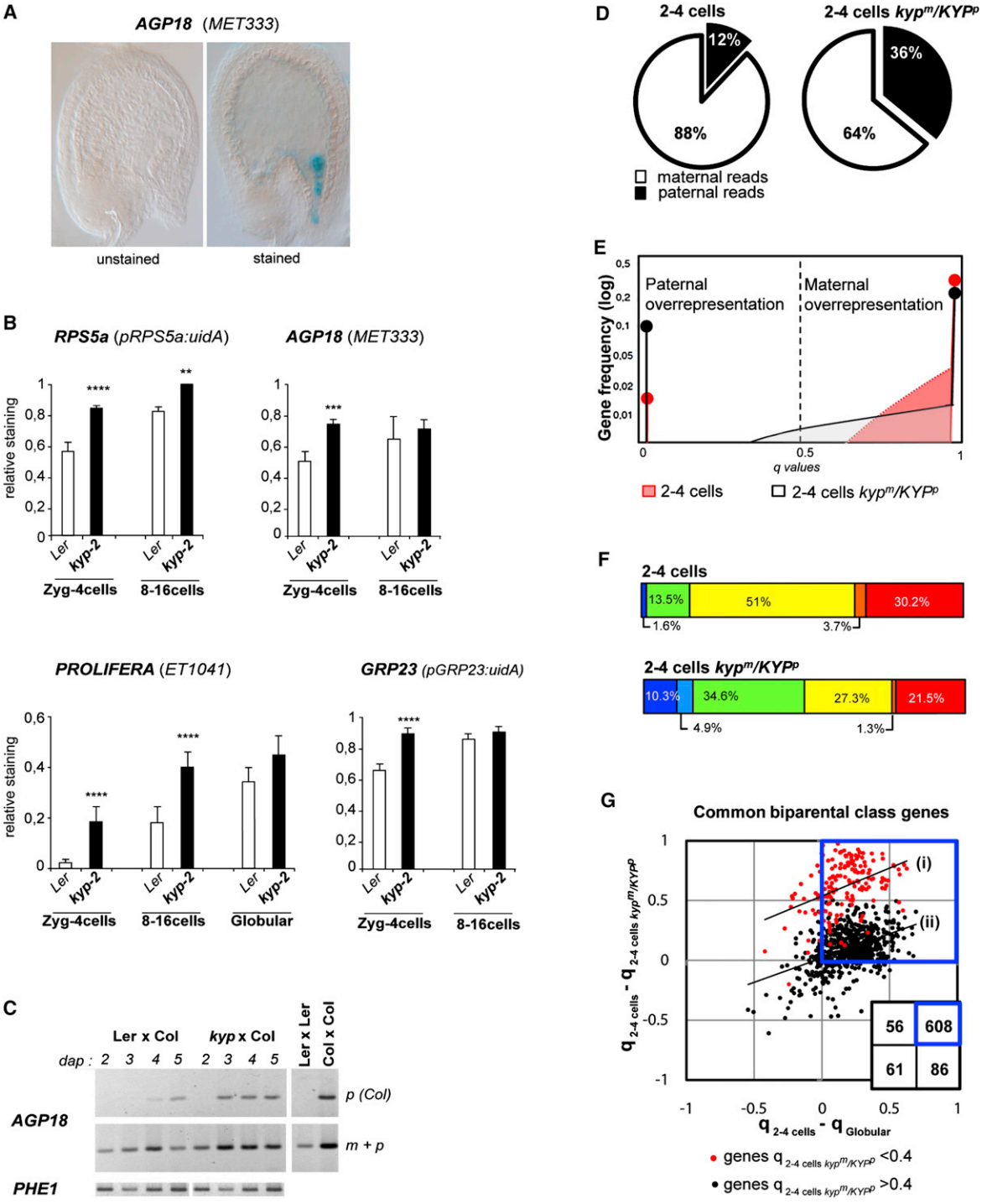


Figure 4. Maternal *KYP* Activity Controls Parental Contributions to the Early Embryo

(A) Histochemical staining for GUS activity from a paternally inherited embryo marker (*MET333*), showing the typical stained and unstained seeds as scored in the graphs in (B).

(B) Expression of paternal markers, in wild-type L*er* or *kyp-2* maternal backgrounds scored as described in Figure 3. See also Figure S3A, Table S3, and Table S4.

(C) Allele-specific RT-PCR of endogenous gene *AGP18* paternal transcripts in siliques harvested 1–5 days after pollination (dap) inheriting a maternal *kyp* mutation, as compared to the wild-type. Selective amplification of the paternal allele (p, top), amplification of maternal and paternal alleles (m + p, middle), control amplification of paternally expressed *PHE1* mRNA (bottom).

(D–G) Allele-specific transcriptome profiling in isolated 2–4 cells embryos inheriting a maternal *kyp-2* mutation compared to 2–4 cell wild-type embryos. Data and legends are as in Figure 1.

silencing pathway (Brodersen and Voinnet, 2006), did not alter the activation kinetics of the paternal reporters (Figure 5A). Importantly, paternal inheritance of mutant *KYP*, *CMT3*, or *NRPD1b* components showed no effect on paternal marker expression (Figure S4D), confirming a specific maternal role for the RdDM pathway in paternal marker regulation.

Several lines of evidence indicate that derepression of the paternal reporters was not linked to their transgenic nature. First, we confirmed precocious detection of paternal transcripts for the endogenous *AGP18* locus in embryos inheriting a maternal *kyp* mutation using allele-specific RT-PCR (Figure 4C). Second, several transgenic reporters under the control of transposon enhancers active in pollen (Slotkin et al., 2009) remained paternally undetectable in embryos inheriting a maternal *cmt3* mutation (Figure S4E), whereas ectopic maternal activation of the same reporters was reported in *cmt3* embryo sacs (Pillot et al., 2010a). Consistently, the profiling confirmed that the *kyp* mutation did not massively derepress transposons and repeats in the embryo (Table S1). Together with the genome-wide analysis of over 3000 endogenous loci in kyp^m/KYP^p embryos, these results indicate that the maternally inherited components of the RdDM pathway are involved in controlling the genome-wide transcriptional dynamics of paternally inherited alleles.

Consistent with the proposed global role for the RdDM pathway in transcriptional control in early embryos, we observed that *nrpd1a1b* mutant zygotes showed an abnormally high level of active PolII in their nucleus (Figure 5B). This was associated with abnormal deposition of the repressive H3K9me2 marks (Figure S5). The epigenetic and transcriptional states of zygotic nuclei in this RdDM mutant is thus in stark contrast to wild-type zygotes, which have a relatively quiescent transcriptional state (Pillot et al., 2010b).

Surprisingly, despite their effect on gene expression, RdDM mutants have not been reported to cause embryo lethality. However, this does not exclude subtle defects and, indeed, transient patterning defects were observed in embryos lacking maternal CMT3 activity (Pillot et al., 2010b). Similarly, early *kyp* embryos showed abnormal division planes in the embryo and suspensor cells, suggesting a role in early embryonic patterning (Figure S6).

Ovules Are Enriched in 24 nt siRNAs Targeting Gene-Coding Sequences

Our results suggest a novel role for the RdDM pathway in the regulation of genic regions, as this pathway had previously been associated mostly with transposon and repeat silencing. To verify the presence of maternal small RNAs targeting protein-coding regions, we profiled a library of small RNAs generated from manually dissected mature ovules before fertilization. We reasoned that if this were a maintenance mechanism, the siRNAs had to be produced before fertilization. Our analysis revealed a large fraction of siRNAs targeting genic regions (comprising protein-coding sequences [CDS] and 500 bp of putative 5′ regulatory regions of genes) in ovules as compared to whole inflorescence (Lu et al., 2005) (Figures 6A and 6B). This increase was not due to 21 nucleotide (nt) sRNAs (represented mainly by DCL1-dependent miRNAs and siRNAs) but was associated with the 24 nt fraction, whose biogenesis is dependent on PolIV, RDR2, and DCL3 (Brodersen and Voinnet, 2006). Overrepresentation of 24 nt siRNAs derived from CDS in ovules was correlated with the transcriptional control of individual loci mediated by maternal KYP activity in the embryo: genes showing a transition from maternal expression in the wild-type to biallelic or paternal expression in kyp^m/KYP^p embryos (*kyp*-responsive genes; Figure 6C) showed significantly more matching siRNAs than genes that remained maternally expressed in kyp^m/KYP^p embryos (*kyp*-unresponsive; Figure 6C). This finding is consistent with a role of the 24 nt siRNAs in regulating paternally inherited alleles.

Maternal CAF-1 and Histone H3 Turnover Regulate Transcriptional Activation of Paternal Alleles

Mutant analyses showed that lack of maternal RdDM components induced the precocious transcriptional activation of paternal alleles. It is unknown whether paternal alleles are in a state permissive to transcription or whether additional epigenetic reprogramming events are necessary for their activation. In the course of our genetic screen for mutants affecting paternal reporter expression, we identified MULTIPLE SUPPRESSOR OF IRA1 (MSI1) (Hennig et al., 2003) to be necessary for their expression. In embryos inheriting a maternal *msi1* mutation, the activation of paternal reporter genes was markedly delayed (Figures 7A and 7B and Figure S7A) compared to wild-type embryos. Reduction of paternal transcript levels in *msi1* mutants was confirmed by RT-PCR for the endogenous *GRP23* locus (Figure 7C).

MSI1 participates in several protein complexes including the CAF1 complex, a component of chromatin organization ensuring mitotic stability (Ono et al., 2006). CAF1 is formed by FASCIATA1 (FAS1), FAS2, and MSI1, and functions as an H3/H4-specific chaperone facilitating nucleosome assembly during replication (Hennig et al., 2005; Kaya et al., 2001). Consistently, maternal loss of another subunit of the CAF1 complex, FAS2, showed a similar effect as *msi1* (Figure 7A and Figure S7B). MSI1 is also a subunit of the MEA-FIE *Polycomb* group (PcG) complex active in seeds (Köhler et al., 2003), but a mutation affecting

(D) Distribution of maternal versus paternal reads as in Figure 1A.
(E) q distribution as in Figure 1B. A total of 3125 genes were identified by informative reads in mutant embryos.
(F) Composite diagram representation of the gene distribution in (E) as in Figure 1C.
(G) Scatter plot distribution of 811 biparental-class genes commonly detected in wild-type 2–4 cell embryos, globular embryos, and 2–4 cell kyp^m/KYP^p embryos. The difference in maternal contribution between stages ($q_{globular} - q_{2–4\ cell}$) or genotype ($q_{2–4\ cell\ WT} - q_{2–4\ cell\ kyp/KYP}$) is plotted on the x and y axis, respectively. q values were calculated for each transcript as follows: maternal reads/maternal + paternal reads. Differences >0 mean a higher paternal contribution compared to wild-type 2-4 cell embryos. Blue frame: genes with a correlated increased paternal contribution in both globular and kyp^m/KYP^p 2–4 cell embryos. Two groups of genes are delineated according to their relative maternal contribution (q) in kyp^m/KYP^p embryos as indicated (red, blue). Linear regressions: (i), $y = 0.47x + 0.54$; $R^2 = 0.17$; (158 genes); and (ii), $y = 0.41x + 0.02$; $R^2 = 0.20$ (653 genes).
Related to Figure S1, Figures S3B–S3F, Table S1, and Table S2.

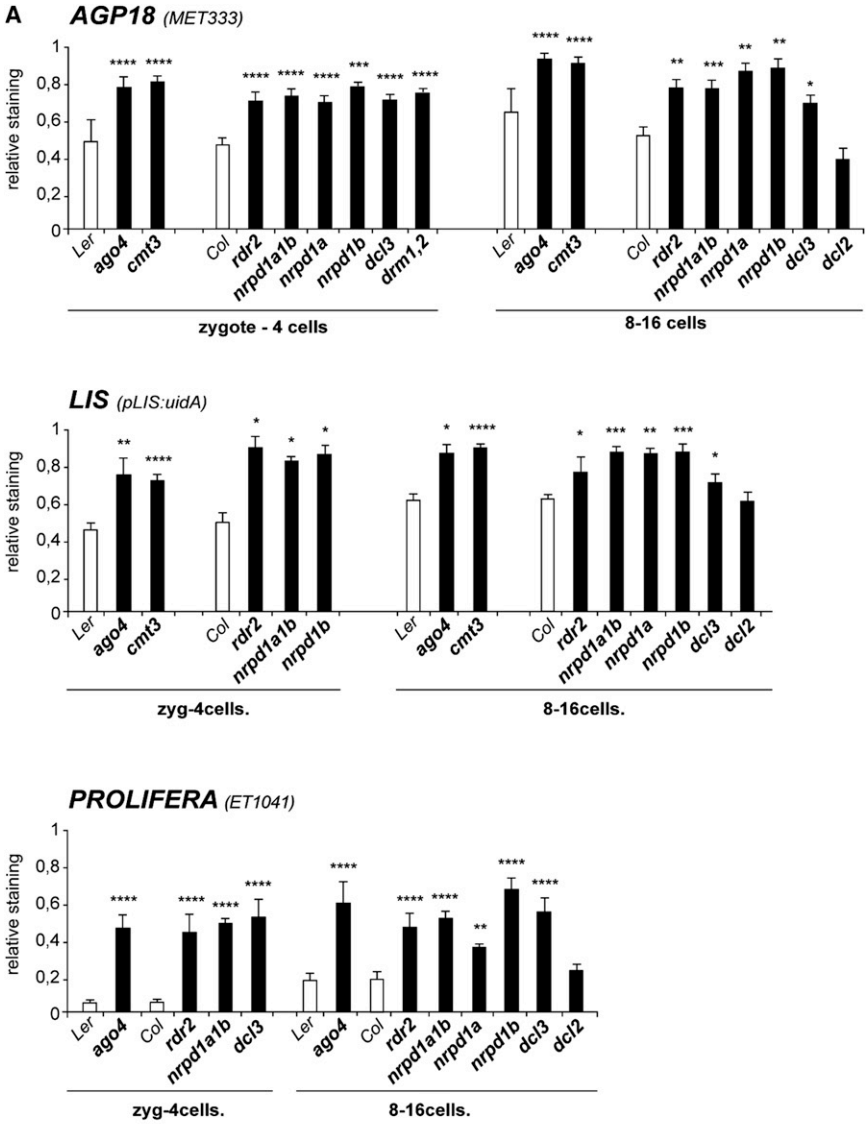


Figure 5. A Functional RdDM Pathway Is Required for Maternal Control of Paternal Embryo Markers and Maintenance of Transcriptional Quiescence in the Zygote

(A) Expression of paternal markers in embryos, scored as in Figure 3 and Figure 4, in maternal mutant backgrounds lacking activity of RdDM components as indicated. See also Figure S4, Figure S6, Table S3, Table S4, and Table S5.

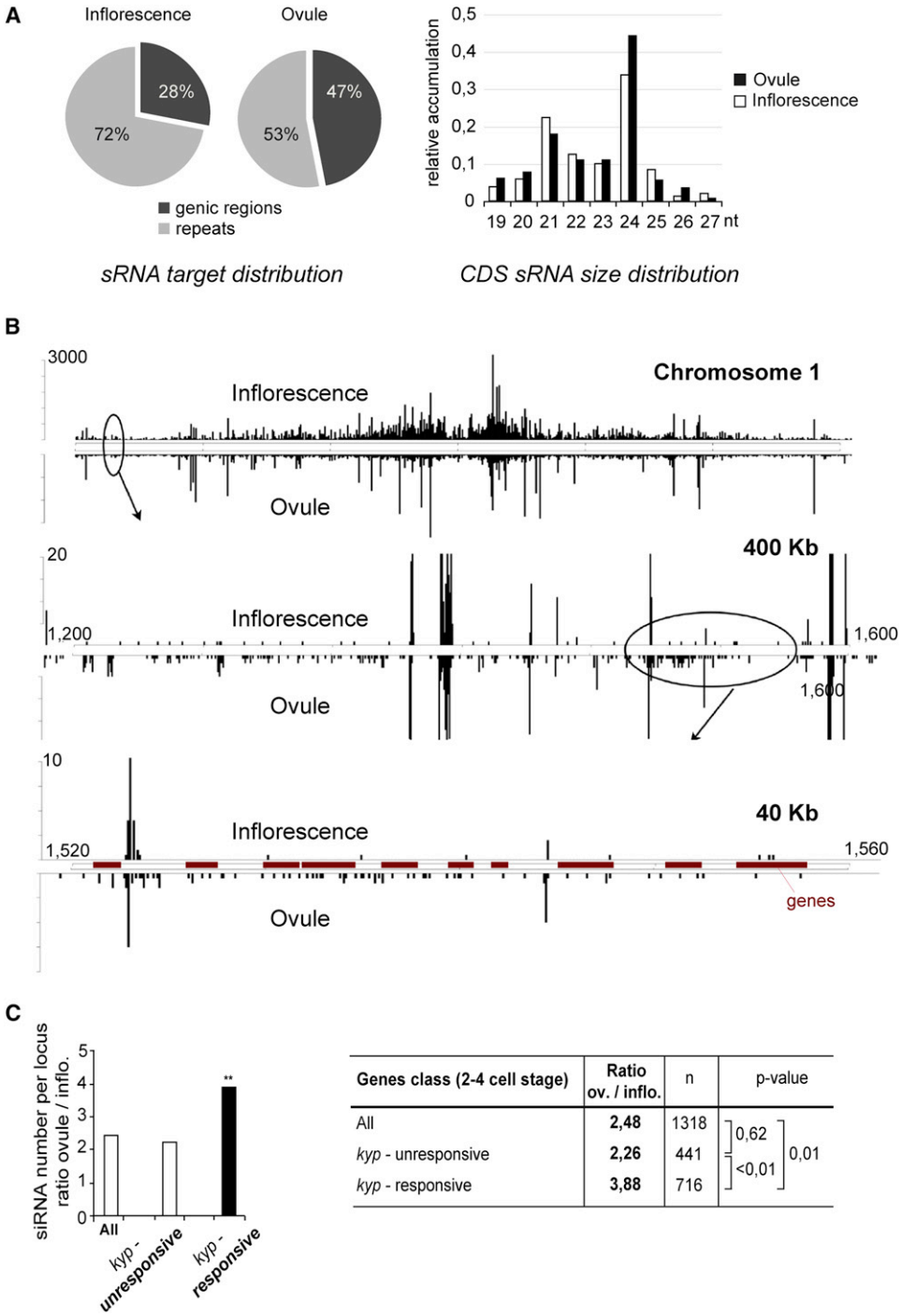
(B) Immunolocalization of the active form of RNA polymerase II (H5) in PolIV/PolV mutant zygotes (*nrpd1a1b*) compared to the wild-type (WT). DNA was counterstained with DAPI. See also Figure S5.

the MEDEA (MEA) subunit had no effect on paternal expression at the globular stage (Figure S7C). These results strongly suggest that the CAF1 complex is maternally required to activate transcription of the paternal genome, likely via histone turnover. CAF1 may regulate the incorporation of specific histone variants controlling transcriptional activity in plants, as it is the case in animals. For example, in the animal germ line H3.3 variants are incorporated at actively transcribed loci (Ooi et al., 2006).

A similar role may be proposed in plants, because we observed that maternal mutations affecting the H3.3 variants HTR4 and HTR5 (Okada et al., 2005) significantly delayed the activation of paternal markers in the embryo (Figure S7D). These results indicate a role for CAF1 and H3/H3.3 turnover in the transcriptional activation of paternal—and possibly maternal—alleles after fertilization.

DISCUSSION

We provided a genome-wide view of the parental contributions to early embryogenesis in *Arabidopsis*. At early stages, the maternal transcriptome clearly predominates: although 68% of the genes are biparentally expressed, their maternal transcripts are overrepresented. Further, >30% of the genes show an exclusively maternal contribution. During the transition from the 2–4 cell to the globular stage the paternal contribution increases. We showed that these parental contributions are maternally controlled by two antagonistic regulatory pathways regulating the onset of paternal and, at least partially, maternal zygotic transcription. Maternal dominance at early stages results from downregulation of paternal alleles at loci throughout the genome via the chromatin siRNA pathway, linked to RNA-directed DNA and histone methylation. In addition, transcriptional activation of paternal alleles involves histone exchange, possibly via the replacement of H3.3 variants, for which rapid turnover is observed after fertilization (Ingouff et al., 2010). Release of silencing might also involve passive DNA-demethylation during mitoses or the activity of DNA- or histone-demethylases. Although additional investigations are required to refine the mechanistic role of these events in the control of the zygotic genome, our results suggest that flowering plants evolved



Genes class (2-4 cell stage)	Ratio ov. / inflo.	n	p-value
All	**2,48**	1318	0,62
kyp - unresponsive	**2,26**	441	<0,01
kyp - responsive	**3,88**	716	0,01

Figure 6. Ovules Are Enriched in Small RNAs Targeting Genic Regions

(A) Deep sequencing of small RNA (sRNAs) from mature ovules reveals increased targeting to genic regions (comprising protein coding sequences (CDS) and putative 5′-regulatory regions 500 bp upstream of ATG) as compared to sRNAs from inflorescence (pie charts). Size distribution of CDS-targeted sRNAs showed an increase in the 24 nt fraction in ovules as compared to inflorescence sRNA libraries (histogram).

(B) Comparative mapping of sRNA distribution in inflorescence (upper line) and ovules (lower line) libraries, exemplified by chromosome 1 (top), with a 400 kb zoom (middle), and 40 kb zoom (bottom). Boxes represent protein coding units (genes). CDS-specific 24 nt sRNA were distributed between + and − strands in the proportion of 63(+):37(−) in the ovule library.

(C) Average number of 24 nt siRNA per locus for maternal- and biparental-class genes showing no significant changes (*kyp*-unresponsive genes) or reduced q values (increased paternal contribution) in kyp^{m}/KYP^{p} embryos compared to wild-type (*kyp*-responsive genes). t test p values (table and graph, *p < 0.01) refers to differences in siRNA mean number between groups.

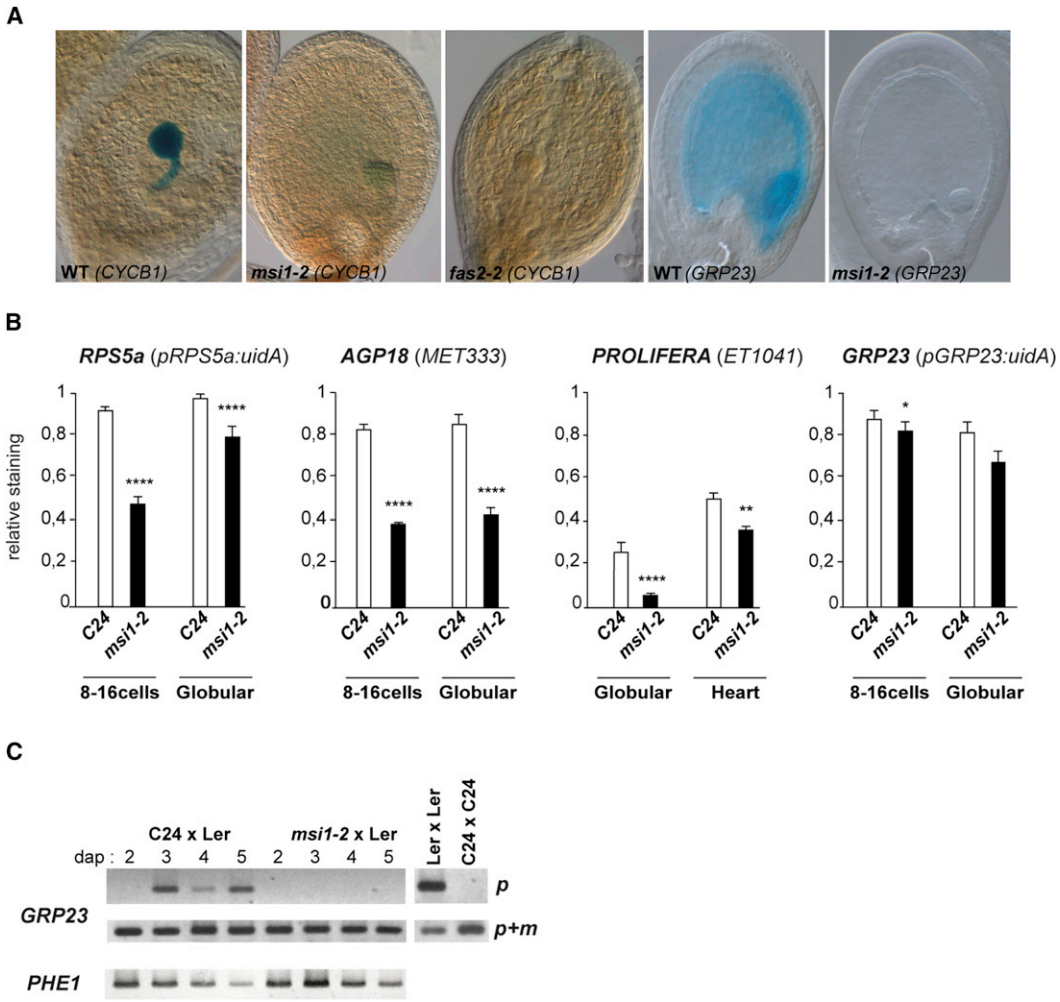


Figure 7. The Maternal CAF1 Nucleosome Assembly Complex Controls Paternal Gene Activation

(A) Paternal marker expression in *msi1* or *fas2* maternally mutant embryos compared to wild-type.

(B) Paternal marker expression scored as in Figure 3, Figure 4, and Figure 5, in *msi1* or *fas2* maternal background. See also Figure S7, Table S3, Table S4, and Table S5.

(C) Allele-specific RT-PCR shows reduced endogenous *GRP23* paternal transcript levels in siliques collected 2–5 dap inheriting a maternal *msi1-2* mutation, as compared to the wild-type. Selective amplification of the paternal allele (p), or of both maternal and paternal alleles (m + p), control amplification of paternally expressed *PHE1*.

strategies to regulate early embryogenesis that are similar to those described in animals (Baroux et al., 2008; Tadros and Lipshitz, 2009).

Maternal Effect and Zygotic Functions during Early Embryogenesis

The relative contribution and dynamic changes of parental transcripts we observed in early *Arabidopsis* embryos may result from a combination of de novo transcription postfertilization and transcripts carried over from the egg or sperm, although their respective abundance is unknown. 1331 and 621 genes with respective maternal and paternal contributions (of 3399 and 1482 represented on the ATH1 microarray) were consistently detected as present in egg- or sperm-specific microarray experiments (Borges et al., 2008; Wuest et al., 2010). Although these overlaps between pre- and postfertilization transcriptomes indicate a carryover from egg and sperm cells, potentially influencing early embryo development (Bayer et al., 2009; Pillot et al., 2010b), the expression status of these genes in the early embryo remains unknown. For many loci, stored and de novo expressed transcripts likely coexist during early embryogenesis (Tadros and Lipshitz, 2009), as it was shown by chromosomal deletions in *Drosophila* (De Renzis et al., 2007). Whenever exactly these genes may be expressed, the vast majority of transcripts are derived from the maternal genome, providing extensive maternal control over early development, explaining the existence of numerous maternal effect genes (reviewed in Baroux et al., 2008). However, it is equally clear that there is a paternal contribution to early embryogenesis from many loci, sometimes exclusively, consistent with paternal effect and zygotic genes acting early after fertilization (Bayer

et al., 2009, Meyer and Scholten, 2007; Ronceret et al., 2005, 2008; Scholten et al., 2002; Tzafrir et al., 2004; Weijers et al., 2001).

Epigenetic Pathways Controlling Parental Contributions Are Distinct from Those Regulating Genomic Imprinting

In plants and mammals, certain genes are regulated by genomic imprinting and are expressed monoallelically only from one parental allele (reviewed in Raissig et al., 2011). One could imagine that for maternally expressed, imprinted loci, the RdDM pathway repressing paternal alleles stays in place throughout development, thus leading to monoallelic maternal expression. However, the maternal regulatory pathways we uncovered complement—and act beyond—the regulation of genes by genomic imprinting: neither *KYP* nor *CMT3* regulates the imprinted *FIS2* locus (Jullien et al., 2006b). Conversely, we determined that *dcl3*, a mutation affecting the RdDM pathway, does not alter paternal silencing of the imprinted *MEA* gene, nor do mutations affecting the CAF1 subunit FAS2 (Table S5). Instead, the MEA-FIE PcG complex maintains silencing of the paternal *MEA* allele via H3K27 methylation (Baroux et al., 2006; Jullien et al., 2006a). This complex is not required for global repression of the paternal genome because a mutation affecting the MSI1 PcG subunit did not show precocious activation of paternal markers, but instead had a delaying effect. Thus, the maternal mechanisms controlling the timing of paternal genome activation described here are distinct from those regulating genomic imprinting. Furthermore, although parental contributions were described in the late endosperm (Hsieh et al., 2011), their regulation at early stages still needs to be elucidated at a genome-wide level. Possibly, different transcriptional controls are in place because the endosperm is actively transcribed soon after fertilization (Aw et al., 2010; Pillot et al., 2010b), owing to its differential targeting by a global DNA demethylase pathway that may counteracts the RdDM pathway (Gehring et al., 2009; Hsieh et al., 2009).

The RdDM Pathway Plays a Role in Early Seed Development

We have shown that maternal mutations affecting the PolIV subunit NRPD1a de-repressed the activity of paternal markers and modified the transcriptional status of the zygotic genome. Thus, we propose that PolIV-dependent 24 nt maternal siRNAs epigenetically control the transcriptional status of paternal, and possibly also maternal, loci throughout the genome. Whether an epigenetic dimorphism establishes differential susceptibility of the parental alleles to siRNA-based regulation is a challenging question that remains to be addressed. Maternal siRNAs were recently detected at later stages of seed development in the endosperm (Mosher et al., 2009), and thus our findings extend their role to early stages of embryogenesis, particularly in regulating protein-coding sequences. The endosperm and its progenitor, the central cell, or alternatively maternal sporophytic tissue (Olmedo-Monfil et al., 2010) have recently been proposed as a potential source of mobile maternal siRNAs driving silencing in the egg cell and the embryo (reviewed, e.g., in Bourc'his and Voinnet, 2010; Feng et al., 2010), an attractive hypothesis awaiting confirmation.

Perturbation of the maternal RdDM pathway leads to transient patterning defects as reported for embryos lacking maternal CMT3 (Pillot et al., 2010b) or, as described here, KYP activity. Thus, the pathway seems to fine-tune the expression of early patterning genes. Alternatively, phenotypic aberrations might be revealed in embryos inheriting a divergent paternal genome distinct from the maternal background, which provides the epigenetic control on zygotic genome expression. Further studies are awaited to determine if RdDM mutants might represent sensitized backgrounds to hybridization and how these maternal pathways affect out-breeding species. Maternal siRNAs, particularly those targeting transposable elements, have been proposed to act in heterosis and inter-specific hybridization (Chen, 2010; Martienssen, 2010). Twenty-four nucleotide siRNAs targeting coding regions are also downregulated in hybrid offspring (Groszmann et al., 2011). Our data suggest that siRNA-based mechanisms also target protein-coding sequences of the early embryonic genome, to control chromatin-based parental interactions during the epigenetic reprogramming that occurs after fertilization.

EXPERIMENTAL PROCEDURES

Plant Material

Arabidopsis thaliana accessions Columbia-0 (Col), Landsberg *erecta* (L*er*), C24, WS or Nossen (No) were used as wild-type controls according to the mutant's background. The marker lines and mutants are listed in Table S3 and genotyping assays are described in Extended Experimental Procedures.

Embryonic cDNA Libraries, Sequencing, and Allele-Specific Transcriptome Analysis

The full method is described in Extended Experimental Procedures. In brief, embryos were released from the seeds by gentle pressure and isolated under an inverted microscope using a microcapillary. Total RNA was extracted using the PicoPure RNA Isolation Kit (Arcturus) and 300–700 pg was amplified in a linear fashion using the WT-Ovation Pico RNA Amplification System (NuGEN Technologies). After second strand cDNA synthesis and library preparation, 50 bases sequence reads were generated by SOLiD v3 (Applied Biosystems) and aligned to the *Arabidopsis* Col genome (TAIR8.0). Unique reads mapping full length in the exome were sorted for the presence of L*er* annotated polymorphisms (Borevitz et al., 2007).

The analysis of parental contributions was performed using a likelihood-based model fitting best the observed distribution of maternal and paternal reads per transcripts. A detailed explanation is provided in Extended Experimental Procedures.

Marker Gene Analysis

The activity of paternal markers following crosses to wild-type or mutant females was assayed by histochemical staining of the *uidA* reporter gene product (the GUS enzyme) as described in Extended Experimental Procedures. The number of GUS-positive progeny was scored for each developmental stage and differences were assessed using two-tailed Fisher's exact test. For allele-specific RT-PCR, LNA-modified primers targeting a SNP from one parental transcript were used. Details on the reactions and primer sequences are provided in Extended Experimental Procedures.

Profiling of Small RNAs in Ovules

In brief, total RNA was extracted from dissected mature ovules (Peiffer et al., 2008) and a small RNA library was prepared and sequenced using Illumina Genome Analyzer. After filtering, reads were mapped against *Arabidopsis* Col reference sequence (TAIR 8) and compared to inflorescence small RNA reads (Lu et al., 2005), analyzed using the same procedure. Mapping, occurrence information, normalization, and graphical displays were computed using R. Genic regions and repeat target analysis was done using TAIR 8 genome release and ftp://ftpmips.helmholtz-muenchen.de/plants/cress/.

Whole-Mount Immunolocalization

Immunodetection was performed essentially as described (Pillot et al., 2010b), using antibodies from Abcam: anti-H3K9me2 (#ab1220) and anti-phosphoS2 RNA PolII [H5] (#ab24758) specific to the active form of PolII (Palancade and Bensaude, 2003). Images were captured on a laser scanning confocal microscope (Leica SP2) and maximum-intensity projections of selected optical sections generated. Details are provided in Extended Experimental Procedures.

ACCESSION NUMBERS

Data from embryo SOLiD profiling are accessible under GEO accession number GSE24198. Data from ovule small RNA profiling are accessible under GEO accession number GSE28627. Sequence data from this article can be found in the *Arabidopsis* Genome Initiative or GenBank/EMBL databases under the following accession numbers: At5g13960 (*KRYPTONITE*), At5g14620 (*DRM2*), At5g15380 (*DRM1*), At1g69770 (*CMT3*), At5g49160 (*MET1*), At5g66750 (*DDM1*), At4g11130 (*RDR2*), At3g43920 (*DCL3*), At3g03300 (*DCL2*), At2g27040 (*AGO4*), At1g63020 (*NRPD1A/POLIVA*), At2g40030 (*NRPD1b/NRPE1/POLV*), At5g58230 (*MSI1*), At5g64630 (*FAS1*), At5g64630 (*FAS2*), At1g02580 (*MEDEA*), At4g40030 (*HTR4*), At4g40040 (*HTR5*), At4g37450 (*AGP18*), At4g02060 (*PROLIFERA*), At4g37490 (*CYCB1;1*), At3g11940 (*RPS5a*), At2g241500 (*LACHESIS*), At1g10270 (*GRP23*).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, five tables, and one movie and can be found with this article online at doi:10.1016/j.cell.2011.04.014.

ACKNOWLEDGMENTS

We thank the *Arabidopsis* Stock Centers (NASC/ABRC), J. Carrington, R. Gross-Hardt, R. Martienssen, F. Meins, T. Lagrange, J. Paszowski, and E. Richards for seeds; C. Michaud, N. Duran-Figueroa, and V. Olmedo-Monfil for technical support; T. Nawy for advice on embryo microdissection; S. Schreiber, L. Bossen, and M. Schilhabel for sequencing support at Kiel; and M. Schmid for help in GEO data submission. We are grateful to M. Arteaga-Vazquez, O. Hamant, and T. Lagrange for helpful comments on earlier versions of the manuscript and to S. Kessler and S. Gillmor for critical reading. C.B., M.R., S.G., and A.S. were supported by the University of Zürich and grants of Swiss National Science Foundation and the European Union (FP5 ApoTool Project) to U.G. D.A., O.L., and D.G. were supported by the Institut de Recherche pour le Développement and the Agence Nationale de la Recherche, and T.L. by CNRS and a starting grant of the European Research Council. P.R. is supported by the Clusters of Excellence: "The Future Ocean" and "Inflammation at Interfaces." J.-P.V.-C. is supported by CONACyT and is an International Scholar of the Howard Hughes Medical Institute.

Received: January 21, 2011
Revised: March 28, 2011
Accepted: April 15, 2011
Published: May 26, 2011

REFERENCES

Andéol, Y. (1994). Early transcription in different animal species: implication for transition from maternal to zygotic control in development. Roux. Arch. Dev. Biol. *204*, 3–10.

Aw, S.J., Hamamura, Y., Chen, Z., Schnittger, A., and Berger, F. (2010). Sperm entry is sufficient to trigger division of the central cell but the paternal genome is required for endosperm development in Arabidopsis. Development *137*, 2683–2690.

Baroux, C., Autran, D., Gillmor, C.S., Grimanelli, D., and Grossniklaus, U. (2008). The maternal to zygotic transition in animals and plants. Cold Spring Harb. Symp. Quant. Biol. *73*, 89–100.

Baroux, C., Blanvillain, R., and Gallois, P. (2001). Paternally inherited transgenes are down-regulated but retain low activity during early embryogenesis in Arabidopsis. FEBS Lett. *509*, 11–16.

Baroux, C., Gagliardini, V., Page, D.R., and Grossniklaus, U. (2006). Dynamic regulatory interactions of Polycomb group genes: MEDEA autoregulation is required for imprinted gene expression in Arabidopsis. Genes Dev. *20*, 1081–1086.

Bayer, M., Nawy, T., Giglione, C., Galli, M., Meinnel, T., and Lukowitz, W. (2009). Paternal control of embryonic patterning in Arabidopsis thaliana. Science *323*, 1485–1488.

Borevitz, J.O., Hazen, S.P., Michael, T.P., Morris, G.P., Baxter, I.R., Hu, T.T., Chen, H., Werner, J.D., Nordborg, M., Salt, D.E., et al. (2007). Genome-wide patterns of single-feature polymorphisms in Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA *104*, 12057–12062.

Borges, F., Gomes, G., Gardner, R., Moreno, N., McCormick, S., Feijo, J.A., and Becker, J.D. (2008). Comparative transcriptomics of Arabidopsis sperm cells. Plant Physiol. *148*, 1168–1181.

Bourc'his, D., and Voinnet, O. (2010). A small-RNA perspective on gametogenesis, fertilization, and early zygotic development. Science *330*, 617–622.

Brodersen, P., and Voinnet, O. (2006). The diversity of RNA silencing pathways in plants. Trends Genet. *22*, 268–280.

Bushell, C., Spielman, M., and Scott, R.J. (2003). The basis of natural and artificial postzygotic hybridization barriers in Arabidopsis species. Plant Cell *15*, 1430–1442.

Chaudhury, A.M., and Berger, F. (2001). Maternal control of seed development. Semin. Cell Dev. Biol. *12*, 381–386.

Chen, Z.J. (2010). Molecular mechanisms of polyploidy and hybrid vigor. Trends Plant Sci. *15*, 57–71.

De Renzis, S., Elemento, O., Tavazoie, S., and Wieschaus, E.F. (2007). Unmasking activation of the zygotic genome using chromosomal deletions in the Drosophila embryo. PLoS Biol. *5*, e117.

Feng, S., Jacobsen, S.E., and Reik, W. (2010). Epigenetic reprogramming in plant and animal development. Science *330*, 622–627.

Gehring, M., Bubb, K.L., and Henikoff, S. (2009). Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science *324*, 1447–1451.

Grimanelli, D., Perotti, E., Ramirez, J., and Leblanc, O. (2005). Timing of the maternal-to-zygotic transition during early seed development in maize. Plant Cell *17*, 1061–1072.

Groszmann, M., Greaves, I.K., Albertyn, Z.I., Scofield, G.N., Peacock, W.J., and Dennis, E.S. (2011). Changes in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to hybrid vigor. Proc. Natl. Acad. Sci. USA *108*, 2617–2622.

Hennig, L., Bouveret, R., and Gruissem, W. (2005). MSI1-like proteins: an escort service for chromatin assembly and remodeling complexes. Trends Cell Biol. *15*, 295–302.

Hennig, L., Taranto, P., Walser, M., Schonrock, N., and Gruissem, W. (2003). Arabidopsis MSI1 is required for epigenetic maintenance of reproductive development. Development *130*, 2555–2565.

Hsieh, T.F., Ibarra, C.A., Silva, P., Zemach, A., Eshed-Williams, L., Fischer, R.L., and Zilberman, D. (2009). Genome-wide demethylation of Arabidopsis endosperm. Science *324*, 1451–1454.

Hsieh, T.F., Shin, J., Uzawa, R., Silva, P., Cohen, S., Bauer, M.J., Hashimoto, M., Kirkbride, R.C., Harada, J.J., Zilberman, D., et al. (2011). Regulation of imprinted gene expression in Arabidopsis endosperm. Proc. Natl. Acad. Sci. USA *108*, 1755–1762.

Ingouff, M., Rademacher, S., Holec, S., Soljic, L., Xin, N., Readshaw, A., Foo, S.H., Lahouze, B., Sprunck, S., and Berger, F. (2010). Zygotic resetting of the HISTONE 3 variant repertoire participates in epigenetic reprogramming in *Arabidopsis*. Curr. Biol. *20*, 2137–2143.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature *416*, 556–560.

Jahnke, S., Sarholz, B., Thiemann, A., Kuhr, V., Gutierrez-Marcos, J.F., Geiger, H.H., Piepho, H.P., and Scholten, S. (2010). Heterosis in early seed development: a comparative study of F1 embryo and endosperm tissues 6 days after fertilization. Theor. Appl. Genet. *120*, 389–400.

Jullien, P.E., Katz, A., Oliva, M., Ohad, N., and Berger, F. (2006a). Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in *Arabidopsis*. Curr. Biol. *16*, 486–492.

Jullien, P.E., Kinoshita, T., Ohad, N., and Berger, F. (2006b). Maintenance of DNA methylation during the Arabidopsis life cycle is essential for parental imprinting. Plant Cell *18*, 1360–1372.

Kaya, H., Shibahara, K.I., Taoka, K.I., Iwabuchi, M., Stillman, B., and Araki, T. (2001). FASCIATA genes for Chromatin Assembly Factor-1 in Arabidopsis maintain the cellular organization of apical meristems. Cell *104*, 131–142.

Köhler, C., Hennig, L., Bouveret, R., Gheyselinck, J., Grossniklaus, U., and Gruissem, W. (2003). Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. EMBO J. *22*, 4804–4814.

Lu, C., Tej, S.S., Luo, S., Haudenschild, C.D., Meyers, B.C., and Green, P.J. (2005). Elucidation of the small RNA component of the transcriptome. Science *309*, 1567–1569.

Martienssen, R.A. (2010). Heterochromatin, small RNA and post-fertilization dysgenesis in allopolyploid and interploid hybrids of Arabidopsis. New Phytol. *186*, 46–53.

Meyer, S., and Scholten, S. (2007). Equivalent parental contribution to early plant zygotic development. Curr. Biol. *17*, 1686–1691.

Mosher, R.A., Melnyk, C.W., Kelly, K.A., Dunn, R.M., Studholme, D.J., and Baulcombe, D.C. (2009). Uniparental expression of PolIV-dependent siRNAs in developing endosperm of Arabidopsis. Nature *460*, 283–286.

Ngo, Q.A., Moore, J.M., Baskar, R., Grossniklaus, U., and Sundaresan, V. (2007). Arabidopsis GLAUCE promotes fertilization-independent endosperm development and expression of paternally inherited alleles. Development *134*, 4107–4117.

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.

Olmedo-Monfil, V., Duran-Figueroa, N., Arteaga-Vazquez, M., Demesa-Arevalo, E., Autran, D., Grimanelli, D., Slotkin, R.K., Martienssen, R.A., and Vielle-Calzada, J.P. (2010). Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature *464*, 628–632.

Ono, T., Kaya, H., Takeda, S., Abe, M., Ogawa, Y., Kato, M., Kakutani, T., Mittelsten Scheid, O., Araki, T., and Shibahara, K. (2006). Chromatin Assembly Factor 1 ensures the stable maintenance of silent chromatin states in Arabidopsis. Genes Cells *11*, 153–162.

Ooi, S.L., Priess, J.R., and Henikoff, S. (2006). Histone H3.3 variant dynamics in the germline of Caenorhabditis elegans. PLoS Genet. *2*, e97.

Pahlavani, M.H., and Abolhasani, K. (2006). Xenia effect on seed and embryo size in cotton (Gossypium hirsutum L.). J. Appl. Genet. *47*, 331–335.

Palancade, B., and Bensaude, O. (2003). Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation. Eur. J. Biochem. *270*, 3859–3870.

Peiffer, J.A., Kaushik, S., Sakai, H., Arteaga-Vazquez, M., Sanchez-Leon, N., Ghazal, H., Vielle-Calzada, J.P., and Meyers, B.C. (2008). A spatial dissection of the Arabidopsis floral transcriptome by MPSS. BMC Plant Biol. *8*, 43.

Pillot, M., Autran, D., Leblanc, O., and Grimanelli, D. (2010a). A role for CHROMOMETHYLASE3 in mediating transposon and euchromatin silencing during egg cell reprogramming in Arabidopsis. Plant Signal. Behav. *5*, 1167–1170.

Pillot, M., Baroux, C., Vazquez, M.A., Autran, D., Leblanc, O., Vielle-Calzada, J.P., Grossniklaus, U., and Grimanelli, D. (2010b). Embryo and endosperm inherit distinct chromatin and transcriptional states from the female gametes in Arabidopsis. Plant Cell *22*, 307–320.

Raissig, M.T., Baroux, C., and Grossniklaus, U. (2011). Regulation and flexibility of genomic imprinting during seed development. Plant Cell *23*, 16–26.

Ronceret, A., Guilleminot, J., Lincker, F., Gadea-Vacas, J., Delorme, V., Bechtold, N., Pelletier, G., Delseny, M., Chaboute, M.E., and Devic, M. (2005). Genetic analysis of two Arabidopsis DNA polymerase epsilon subunits during early embryogenesis. Plant J. *44*, 223–236.

Ronceret, A., Gadea-Vacas, J., Guilleminot, J., Lincker, F., Delorme, V., Lahmy, S., Pelletier, G., Chaboute, M.E., and Devic, M. (2008). The first zygotic division in Arabidopsis requires de novo transcription of thymidylate kinase. Plant J. *53*, 776–789.

Scholten, S., Lörz, H., and Kranz, E. (2002). Paternal mRNA and protein synthesis coincides with male chromatin decondensation in maize zygotes. Plant J. *32*, 221–231.

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzic, M., Becker, J.D., Feijo, J.A., and Martienssen, R.A. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell *136*, 461–472.

Tadros, W., and Lipshitz, H.D. (2009). The maternal-to-zygotic transition: a play in two acts. Development *136*, 3033–3042.

Tzafrir, I., Pena-Muralla, R., Dickerman, A., Berg, M., Rogers, R., Hutchens, S., Sweeney, T.C., McElver, J., Aux, G., Patton, D., et al. (2004). Identification of genes required for embryo development in Arabidopsis. Plant Physiol. *135*, 1206–1220.

Vielle-Calzada, J.P., Baskar, R., and Grossniklaus, U. (2000). Delayed activation of the paternal genome during seed development. Nature *404*, 91–94.

Weijers, D., Geldner, N., Offringa, R., and Jürgens, G. (2001). Seed development: early paternal gene activity in Arabidopsis. Nature *414*, 709–710.

Wuest, S.E., Vijverberg, K., Schmidt, A., Weiss, M., Gheyselinck, J., Lohr, M., Wellmer, F., Rahnenfuhrer, J., von Mering, C., and Grossniklaus, U. (2010). Arabidopsis female gametophyte gene expression map reveals similarities between plant and animal gametes. Curr. Biol. *20*, 506–512.

EXTENDED EXPERIMENTAL PROCEDURES

Plant Material

Arabidopsis thaliana accessions Columbia-0 (Col), Landsberg *erecta* (L*er*), C24, WS or Nossen (No) were used as wild-type controls depending on the mutant investigated. The reporter lines used are listed in Table S3 with the corresponding reference. The *pAtRP-S5A:uidA* reporter was reconstructed in a different vector backbone (pCAMBIA1391Z) using the same promoter as originally described (Weijers et al., 2001) and transformed into the L*er* accession. The mutants used are listed and referenced in Table S3. All mutants were homozygous for a recessive null mutation except for *ddm1-2/DDM1*, *met1-3/MET1*, *msi1-2/MSI1* and *mea-2/MEA*. For the latter, heterozygous mutants were genotyped in a segregating population. For *ddm1-2* a cleaved amplified polymorphic sequence (CAPS) marker (gift from E. Richards) was used. A 100 bp amplicon produced with the primers 5′-gttggacagtgtggta aattccgct-3′ and 5′-gagctacgagccatgggtttgtgaaacgta-3′ (Tm 56°C, 40 cycles) was ethanol precipitated and digested with RsaI (NEB, Ipswich CA, USA) for 3 hr at 37°C. *msi1-2/MSI1*, *mea-2/MEA*, *met1-3/MET1* plants were genotyped as described (Baroux et al., 2006; Hennig et al., 2003; Saze et al., 2003). For the *htr4-1* mutant (line N582765, T-DNA insertion in the second exon of *At4g40030*) (Okada et al., 2005), abolished transcription was verified by RT-PCR using the primers 5′-tggctcgtaccaag-caaaccgctcg-3′ and 5′-acggactagcctctgaaatggcagtt-3′ (Tm 62°C, 35 cycles), targeting exon2 and exon 3, respectively. Equal loading was controled by amplifying *ACTIN11* mRNA as described (Baroux et al., 2006).

Preparation of Embryonic cDNA Libraries and Sequencing

For embryo isolation 3–5 siliques resulting from crosses between L*er* wild-type or *kyp-2/kyp-2* (L*er*) mutant mothers and Col wild-type fathers were harvested 2.5 days (2–4 cell embryos) and 4 days (globular) after pollination. Seeds were dissected and immersed in 20 μl isolation buffer (first-strand cDNA synthesis buffer [Invitrogen], 1.6 U/μl RNase Out [Invitrogen], 1 mM DTT), in a round-bottom 2 ml Eppendorf tube. Seeds were gently crushed with a plastic pestle to release the embryos. 400 μl isolation buffer was added to the extract and 5 x 50 μl droplets were placed on 6-well printed slides previously coated with 1% BSA. 50 μl fresh isolation buffer was placed on the remaining well for washing the isolated embryos (see below).

One slide was placed on an inverted microscope (Nikon TMS) and the droplets were screened at magnification 10x. Embryos were isolated with a siliconized, manually drawn, and freshly BSA-coated glass capillary fixed to a micromanipulator and linked to a 200 μl pipette with a rubber tube. The calibration wheel of the pipette was used to create a slight vacuum in the capillary to collect the embryos with as little solution as possible. The embryos were released in a clean drop of buffer and collected again (in ca. 2–3 μl) before release in 100 μl RNA extraction buffer. For profiling of wild-type transcriptomes (L*er* x Col) 28 and 4 embryos at the 2–4 cell stage and globular stage, respectively, were isolated. 25 mutant embryos from *kyp-2* x Col crosses were isolated. Plants were grown, crossed and harvested at the same time. RNA extraction was performed using the PicoPure RNA Isolation Kit (Arcturus) according to the manufacturer's instructions. Quality of the total embryonic RNA was assessed using Agilent RNA 6000 Pico Kit on the Agilent 2100 BioAnalyzer (Agilent Technologies, Germany) and we estimated a yield of 700 pg to 4 ng total RNA in each sample. 300–700 pg of total RNA was amplified using the WT-Ovation Pico RNA Amplification System (NuGEN Technologies, USA). The amplification technology is inspired from Philipps and Eberwine (1996) and performs a linear isothermal amplification of mRNA species. Unlike PCR-based exponential amplification, this linear amplification approach is carried out by replication of only the original transcripts, not replication of copies. Amplification of our samples produced 6–10 μg single-stranded cDNA. To create the second strand, poly(A) tails were added to 1 μg of the amplified cDNA library (10 pMol DNA ends) using 20U of Terminal Transferase (New England Biolabs, USA) and 0.2 mM dATPs in the provided buffer. The second strand was amplified during 1 PCR cycle on a thermal cycler (30″ at 95°C, 2 min at 50°C, 20 min at 72°C) using the Ex Taq Polymerase (TaKaRa, Japan) and oligo(dT)$_{12\text{-}18}$ Primer (Invitrogen).

Typically 200–500 ng of cDNA was used for SOLiD system's express fragment library preparation. Using the Covaris S2 system (Covaris, Inc.), cDNA (0.5–3 kb) was sheared into 80–130 bp short fragments according to the protocol. The ends of the target DNA were repaired and subsequently ligated to SOLiD P1 and P2 adaptors. After ligation the library was enriched by PCR (7 cycles) and a size selecting gel was run to remove any short fragments. The resulting ligated population was the SOLiD Fragment Library ready for emulsion PCR. Emulsion PCR reactions were performed according to the manufacturer's recommendation (Applied Biosystems, USA) by mixing 170 pg libraries with 0.8 billion 1 mm-diameter beads with P1 primers (ABI) covalently attached to their surfaces. Sequence reads of 50 bases length were generated by SOLiD v3 (Applied Biosystems, USA).

Allele-Specific Transcriptome Analysis

50 base reads generated by SOLiD v3 were aligned to the TAIR8.0 version of the *Arabidopsis* Col genome using the SOLiD System Analysis Pipeline Tool (Corona Lite 4.0r2.0, Applied Biosystems, USA), allowing up to 4 color-space mismatches with the additional rule ''count valid adjacent errors as single errors.'' For transcriptome profiles, reads were excluded if they mapped to more than one genomic position, mapped at splice junctions or were partially overlapping. Consequently, full-length reads uniquely mapping inside a transcript were taken.

Arabidopsis L*er* SNPs (ftp://ftp.Arabidopsis.org/Polymorphisms/Ecker_ler.homozygous_snp.txt) (Borevitz et al., 2007) were used to identify reads matching L*er* sequences. We only used SNPs that were biallelic and have exactly one defined allele for Col and one

defined allele for L*er*. Under such constraints, we removed 24 SNPs from the published list for our calling procedure. To assign the reads to a L*er* or Col allele, we used the SNP calling pipeline (consensus caller) of Corona Lite 4.0r2.0. This procedure delivers the number of reads matching the reference sequence and the alternative sequence when at least 3 alternative alleles could be found. For each base change (alternative sequence) at a given position, probability and confidence scores are calculated and are used by the calling algorithm to categorize heterozygous and homozygous SNPs. In addition, the alternative sequence must be met by at least 3 reads with independent starting points. Consequently, in our case SNPs sequenced from one parent only (for instance L*er* if Col is used as the reference) would not be detected. To correct for this, we ran the SNP calling procedure twice, using a reciprocal set-up where the reference sequence was either L*er* or Col. To this aim, we generated two references with the new SNP list: one with the Col allele at all 304,978 genomic positions and one with the L*er* allele. Next, we checked for inconsistencies and only SNP calls showing the exact predicted L*er* or Col base were taken, SNP calls showing a different base were excluded (32 SNPs). To calculate the coverage at each SNP position, the results from the reciprocal SNP calling were merged, choosing the read count with highest number of unique start points then highest coverage. To calculate the allele-specific coverage (Col, L*er*), the coverage per SNPs per transcript was added following a procedure to eliminate redundant read counts, when those were covering 2 (or more) SNP. The procedure consisted in (1) sorting the SNPs by highest number of unique start points first, highest number of coverage second, (2) SNPs were interrogated, starting with the highest covered SNP, for their position and only SNPs more than 49 bp apart from a previously chosen SNP are kept, (3) we summarized Col- and L*er*-specific coverage and start points of all SNPs kept. When SNPs matched an annotation with two or more entries (transcript version), the entry with the highest coverage was used for calculating the transcript level.

Analysis of the Parental Distribution of Embryonically Expressed Genes Covered by SOLiD Reads

We assume that all gene expression patterns are either biparental, uniparental maternal, or uniparental paternal. A gene with both maternal and paternal transcripts is necessarily biparentally expressed. However, when transcripts from a single parent are detected but only a few reads have been sequenced, it is not immediately clear whether the gene is uni- or biparentally expressed, and the probability of missing one parental contribution in the sample needs to be considered. The relative probability of being uni- versus biparentally expressed was obtained using the following analysis. We note q the proportion of maternal transcripts: q = 0 for only paternal expression, q = 1 for only maternal expression and 0 < q < 1 for biparental expression. We build a model to adjust the distribution of q. We note θ_{pat}, θ_{mat}, θ_{bip} the proportion of genes that are paternally, maternally, or biparentally expressed, respectively. We assume that q values for genes that are biparentally expressed is Beta distributed (with parameters θ_a and θ_b). We note m_i and p_i the observed number of paternal and maternal transcripts for gene i. We note **m** and **p** the vector of all m_i and p_i and θ the vector of parameters to be estimated. The likelihood of the data can then be written:

$$L(\mathbf{m},\mathbf{p}\mid\boldsymbol{\theta})=\prod_i\begin{cases}\theta_{pat}+\theta_{bip}\int_0^1\beta(\theta_a,\theta_b;x)\mathrm{B}(p_i,x;0)dx & \text{if } m_i=0\\ \theta_{bip}\int_0^1\beta(\theta_a,\theta_b;x)\mathrm{B}(p_i+m_i,x;m_i)dx & \text{if } m_ip_i>0\\ \theta_{mat}+\theta_{bip}\int_0^1\beta(\theta_a,\theta_b;x)\mathrm{B}(m_i,x;m_i)dx & \text{if } p_i=0\end{cases}\quad(1)$$

where $\beta(\theta_a, \theta_b ; x)$ denotes the probability to draw x in a Beta distribution with parameters θ_a and θ_b and where B(n, x; k) denotes the probability to draw k success among n trials with a probability of success x (i.e., in a binomial distribution with parameters n and x). For consistency, we assume that the mode of q distribution for biparentally expressed genes is not 0 or 1, which entails that $\theta_a > 1$ and $\theta_b > 1$. If no maternal transcripts are present, the gene may be paternally (with probability θ_{pat}) or biparentally expressed (with probability θ_{bip}). In the latter case, the probability to not sample any maternal transcript is evaluated by the corresponding binomial distribution integrated over the distribution of q within the biparentally expressed genes (the Beta distribution). This leads to the first line of Equation 1. Similar reasoning yields the two other lines in the equation. The relative probability of being uni- versus biparentally expressed can then be calculated. A gene with no maternal transcript is exclusively paternally expressed with probability

$$P_i=\frac{1}{1+\int_0^1\beta(\hat{\theta}_a,\hat{\theta}_b;x)\mathrm{B}(p_i,x;0)dx}\quad(2)$$

and biparentally expressed with probability $B_i = 1 - P_i$. Similarly, a gene with no paternal transcript is exclusively maternally expressed with probability

$$M_i=\frac{1}{1+\int_0^1\beta(\hat{\theta}_a,\hat{\theta}_b;x)\mathrm{B}(m_i,x;m_i)dx}\quad(3)$$

and biparentally expressed with probability $B_i = 1 - M_i$. In Equations 2 and 3, the hat denotes the maximum likelihood estimates of the parameters that have been obtained by maximizing $L(\mathbf{m},\mathbf{p}\mid\boldsymbol{\theta})$. We performed this analysis independently for the three samples

analyzed (2–4 cell wild-type embryos, globular embryos, and 2–4 cell *kyp/KYP* embryos). P_i, M_i, and B_i values for all genes are provided in Table S2.

To obtain a global view of the variations of the parental expression between the samples analyzed, we computed transition matrices. More precisely, we computed the overall proportion of genes being maternally, biparentally or paternally expressed in sample 1 becoming maternally, biparentally or paternally expressed in sample 2. For instance the proportion of maternally expressed genes in 2–4 cell wild-type embryos (WT24) becoming biparentally expressed in Globular embryos (WT Glob) was computed as:

$$\sum_{i=1}^{\#\ genes} M_{i(\mathrm{WT24})} B_{i(\mathrm{WTGlob})}, \tag{4}$$

where M_i and B_i are given by Equations 2 and 3. Thus, these transition matrices account for the uncertainty of assigning each gene to maternal, biparental, or paternal categories. Two matrices were computed, for (2–4 cell wild-type -> globular) and (2–4 cell wild-type -> 2–4 cell *kyp/KYP*) transitions.

Histochemical Detection of the uidA Reporter Gene Product (GUS Staining)

Developing siliques were cut longitudinally and fixed in ice-cold 90% acetone for < 1 hr at −20°C. After washing three times with 100 mM phosphate buffer (100 mM Na2HPO4, 100 mM NaH_2PO_4), the tissue was immersed in staining solution (0.1% Triton X-100, 10 mM EDTA, 0.5 mM Ferrocyanide, 0.5 mM Ferricyanide and 4 mM 5 -bromo-4-chloro-3-indolyl-beta-d-glucuronic acid cyclohexyl-ammonium salt (X-gluc, Biosynth AG, Staad, CH) in 100 mM phosphate buffer) and vacuum-infiltrated for 5 min. The staining reaction was carried out for 2 days at 37°C except for ET1041 (4 days) and *pGRP23:uidA* (2 hr). The staining solution was removed and the samples mounted in clearing solution (40 g of chloralhydrate (Sigma, Steinheim, DE) dissolved in 5 ml Glycerol, 1 ml Lactic acid, and 10 ml water).

Quantification of Paternal Reporter Gene Expression following Crosses

Wild-type and mutant plants were pollinated with the marker lines listed in Table S3, 2 days after emasculation. In the case of *msi1-2/MSI1* and C24, however, pollination was carried out 1 day after emasculation to limit the formation of autonomous seed development (Hennig et al., 2003). Developing siliques were harvested at different days after pollination and stained as described above. The seeds showing GUS staining were scored under a Leica HC microscope (Leica Microsystems, Wetzler GMBH, DE) or Zeiss Axio Imager microscope (Carl Zeiss MicroImaging GmbH, DE). Imaging was done with a CCD camera (Magnafire - Optronics, Goletta, USA) and images edited using Graphic Converter (lemkeSOFT). The seeds were scored according to four classes corresponding to the following developmental stages: zygote to 4 cell stage (embryo proper), octant to 16 cell stage, globular stage, and heart stage. In the case of *msi1-2/MSI1* and C24, the earliest class (zygote to 4 cell stage) was not considered because of potential bias due to parthenogenetic embryo development in the *msi1* mutant (Guitton and Berger, 2005). Average and standard error of the relative proportion of GUS staining seeds in the different developmental classes were calculated from independent biological replicates (Table S4). For each replicate, wild-type and mutant samples were processed in parallel (cross, GUS staining, scoring) to ensure comparable results. Differences in the number of GUS-positive seeds, either between wild-type and mutant samples or between consecutive developmental stages, were tested for statistical significance using the two-tail Fisher's exact test (http://www.langsrud.com/fisher.htm). The details of the scoring results are given in Table S4.

Allele-Specific RT-PCR

Total RNA was isolated (Trizol reagent, Invitrogen) from siliques collected at 2, 3, 4, and 5 days after pollination and treated with DNase I (Invitrogen). cDNA was produced using SuperSript III (Invitrogen) and polyT primers following the manufacturer's instructions. *AGP18* (*At4g37450*) CDS shows a single-nucleotide polymorphism (SNP) between the L*er* and Col ecotypes. An LNA-modified primer was designed (AGP18-Col-R: 5′-gcagttggagttttcgccggagc+c-3′) ("+" before the base indicates the LNA modified base) that, together with a nonspecific primer (AGP18-F5′-ggccaatctcctatctcttctccga-3′) specifically amplifies the Col allele. In a cross between a wild-type L*er* or *kyp-2* female and a Col pollen donor, this primer pair specifically detects the paternal Col allele following 40 cycles with 62°C annealing temperature. As an internal control, both parental alleles were amplified using AGP18-F and AGP18-R (5′-gcagttggacttttgccggagct-3′). Similarly, a SNP in *GRP23* (*At1g10270*) allowed distinguishing the C24 maternal and L*er* paternal alleles in crosses between C24 wild-type or *msi1-2* females and L*er* pollen donors. GRP23-Ler-R (5′-cggtggctgttgccctgccgt+c-3′) and GRP23-F (5′- gcaggtcaaacagcaggaggag-3′) specifically amplified the paternal L*er* allele using 39 cycles with 63°C annealing temperature. Amplification of both parental alleles was obtained using GRP23-F and GRP23-R (5′-cggtggctgttgccctgccgtt-3′) primers. Control PCR reactions were carried out on samples without reverse transcriptase to confirm the absence of genomic DNA. RT-PCR analysis of the *PHE1* gene was done using primers described previously (Köhler et al., 2005) and 36 cycles of amplification with 62°C annealing temperature.

Profiling of Ovule Small RNAs

Mature ovules were collected using a custom micro-pump (Peiffer et al., 2008), avoiding contamination by placenta and silique valve tissues. Total RNA was extracted using the Trizol reagent according to the manufacturer's instructions. A small RNA library was prepared and sequenced using current Illumina Genome Analyzer protocols (http://www.illumina.com). Approx. 10 millions reads were obtained, which were filtered for adaptator removal, size (reads between 19 and 27 nt were conserved), mismatches, and redondancy using custom PERL scripts and the MEGA-BLAST algorythm (http://blast.ncbi.nlm.nih.gov/). The library finally contained 1,193,800 unique reads, which were mapped against *Arabidopsis* Col reference sequence (TAIR 8) using the BOWTIE software (http://bowtie-bio.sourceforge.net/index.shtml) (Langmead et al., 2009). The same prodecure was used to analyze a 995,650 read inflorescence small RNA library, sequenced by B. Meyers' group (library code: FLR, http://mpss.udel.edu/at/tiny_library.php?lib=1) (Lu et al., 2005). Mapping and occurrence informations, after normalization of both libraries to 1 million, were compiled and graphical displays were produced using R. For relative CDS and repeat targeting analysis, two different reference databases (indexes), CDS and Repeats, were built based on TAIR 8 information (ftp://ftp.Arabidopsis.org/home/tair/Genes/TAIR8_genome_release/) and repeat annotation available (ftp://ftpmips.helmholtz-muenchen.de/plants/cress/), and compared after mapping with BOWTIE.

Whole-Mount Immunolocalization

All antibodies were obtained from Abcam (Cambridge, UK). The H5 (#ab24758) raised against the active form of PolII specifically targets the CTD of the main sub-unit of PolII when it is phosphorylated on serine 2, which occurs during transcript elongation. As shown previously (Palancade and Bensaude, 2003), the CTD phosphorylation pattern is modified as PolII engages in transcript elongation and the H5 antibody detects Pol II molecules during active transcription. Chromatin analysis was performed using H3K9me2 (#ab1220) or H3K9me3 (#ab71999) antibodies. Young siliques were fixed for 3 to 4 hr in 4% paraformaldehyde in 1xPBS with 2% Triton, washed twice in 1xPBS. Young seeds were dissected and embedded in acrylamide on slides as described (Bass et al., 1997). Samples were digested with 1% driselase, 0,5% cellulase, 1% pectolyase (all from Sigma) in 1XPBS with 1%BSA for 25 min to 1 hr at 37°C, subsequently rinsed 3 times in 1xPBS, and permeabilized for 1 to 2 hr in 1xPBS with 2% Triton. The primary antibodies were applied at the following dilutions: 1:400 for H3K9me2 and H3K9me3; 1:200 for H5, overnight at 4°C. The slides were washed day-long in 1xPBS, 0,2% Triton, and coated with secondary antibody (Alexa Fluor 488 conjugate, Molecular Probes) used at a 1:400 dilution. After washing in 1xPBS; 0,2% Triton for a minimum of 6 hr, the slides were incubated with DAPI (1ug/ml in 1xPBS) for 1h, washed for 1h in 1xPBS, and mounted in PROLONG medium (Molecular Probes). Images were captured on a laser scanning confocal microscope (Leica SP2) equipped for DAPI (405nm) and FITC (488 nm) excitation and either 40X or 63X objectives. Maximum-intensity projections of selected optical sections were generated, and edited using Graphic Converter (lemkeSOFT).

SUPPLEMENTAL REFERENCES

Acosta-Garcia, G., and Vielle-Calzada, J.P. (2004). A classical arabinogalactan protein is essential for the initiation of female gametogenesis in *Arabidopsis.* Plant Cell *16*, 2614–2628.

Bass, H.W., Marshall, W.F., Sedat, J.W., Agard, D.A., and Cande, W.Z. (1997). Telomeres cluster de novo before the initiation of synapsis: a three-dimensional spatial analysis of telomere positions before and during meiotic prophase. J. Cell Biol. *137*, 5–18.

Blevins, T., Rajeswaran, R., Shivaprasad, P.V., Beknazariants, D., Si-Ammour, A., Park, H.S., Vazquez, F., Robertson, D., Meins, F., Jr., Hohn, T., et al. (2006). Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. Nucleic Acids Res. *34*, 6233–6246.

Cao, X., and Jacobsen, S.E. (2002). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proc. Natl. Acad. Sci. USA *99* (*Suppl 4*), 16491–16498.

Chan, S.W., Henderson, I.R., Zhang, X., Shah, G., Chien, J.S., and Jacobsen, S.E. (2006). RNAi, DRD1, and histone methylation actively target developmentally important non-CG DNA methylation in *Arabidopsis.* PLoS Genet. *2*, e83.

Colon-Carmona, A., You, R., Haimovitch-Gal, T., and Doerner, P. (1999). Technical advance: spatio-temporal analysis of mitotic activity with a labile cyclin-GUS fusion protein. Plant J. *20*, 503–508.

Ding, Y.H., Liu, N.Y., Tang, Z.S., Liu, J., and Yang, W.C. (2006). *Arabidopsis* GLUTAMINE-RICH PROTEIN23 is essential for early embryogenesis and encodes a novel nuclear PPR motif protein that interacts with RNA polymerase II subunit III. Plant Cell *18*, 815–830.

Phillips, J., and Eberwine, J.H. (1996). Antisense RNA Amplification: A linear amplification method for analyzing the mRNA population from single living cells. Methods *10*, 283–288.

Ferreira, P.C., Hemerly, A.S., Engler, J.D., van Montagu, M., Engler, G., and Inze, D. (1994). Developmental expression of the *Arabidopsis* cyclin gene cyc1At. Plant Cell *6*, 1763–1774.

Gross-Hardt, R., Kägi, C., Baumann, N., Moore, J.M., Baskar, R., Gagliano, W.B., Jurgens, G., and Grossniklaus, U. (2007). LACHESIS restricts gametic cell fate in the female gametophyte of *Arabidopsis.* PLoS Biol. *5*, e47.

Grossniklaus, U., Vielle-Calzada, J.P., Hoeppner, M.A., and Gagliano, W.B. (1998). Maternal control of embryogenesis by MEDEA, a Polycomb group gene in *Arabidopsis.* Science *280*, 446–450.

Guitton, A.E., and Berger, F. (2005). Loss of function of MULTICOPY SUPPRESSOR OF IRA1 produces nonviable parthenogenetic embryos in *Arabidopsis.* Curr. Biol. *15*, 750–754.

Jackson, J.P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., Singh, P.B., Cheng, X., Schubert, I., Jenuwein, T., and Jacobsen, S.E. (2004). Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in *Arabidopsis thaliana.* Chromosoma *112*, 308–315.

Jeddeloh, J.A., Stokes, T.L., and Richards, E.J. (1999). Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nat. Genet. *22*, 94–97.

Köhler, C., Page, D.R., Gagliardini, V., and Grossniklaus, U. (2005). The *Arabidopsis thaliana* MEDEA Polycomb group protein controls expression of PHERES1 by parental imprinting. Nat. Genet. *37*, 28–30.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lindroth, A.M., Shultis, D., Jasencakova, Z., Fuchs, J., Johnson, L., Schubert, D., Patnaik, D., Pradhan, S., Goodrich, J., Schubert, I., et al. (2004). Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. EMBO J. *23*, 4286–4296.

Pontier, D., Yahubyan, G., Vega, D., Bulski, A., Saez-Vasquez, J., Hakimi, M.A., Lerbs-Mache, S., Colot, V., and Lagrange, T. (2005). Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis.* Genes Dev. *19*, 2030–2040.

Saze, H., Mittelsten Scheid, O., and Paszkowski, J. (2003). Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. Nat. Genet. *34*, 65–69.

Sung, S., He, Y., Eshoo, T.W., Tamada, Y., Johnson, L., Nakahigashi, K., Goto, K., Jacobsen, S.E., and Amasino, R.M. (2006). Epigenetic maintenance of the vernalized state in *Arabidopsis thaliana* requires LIKE HETEROCHROMATIN PROTEIN1. Nat. Genet. *38*, 706–710.

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G., and Colot, V. (2007). *Arabidopsis* TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. PLoS Genet. *3*, e86.

Weijers, D., Franke-van Dijk, M., Vencken, R.J., Quint, A., Hooykaas, P., and Offringa, R. (2001). An *Arabidopsis* Minute-like phenotype caused by a semi-dominant mutation in a RIBOSOMAL PROTEIN S5 gene. Development *128*, 4289–4299.

Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., and Carrington, J.C. (2004). Genetic and functional diversification of small RNA pathways in plants. PLoS Biol. *2*, E104.

Zilberman, D., Cao, X., and Jacobsen, S.E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science *299*, 716–719.

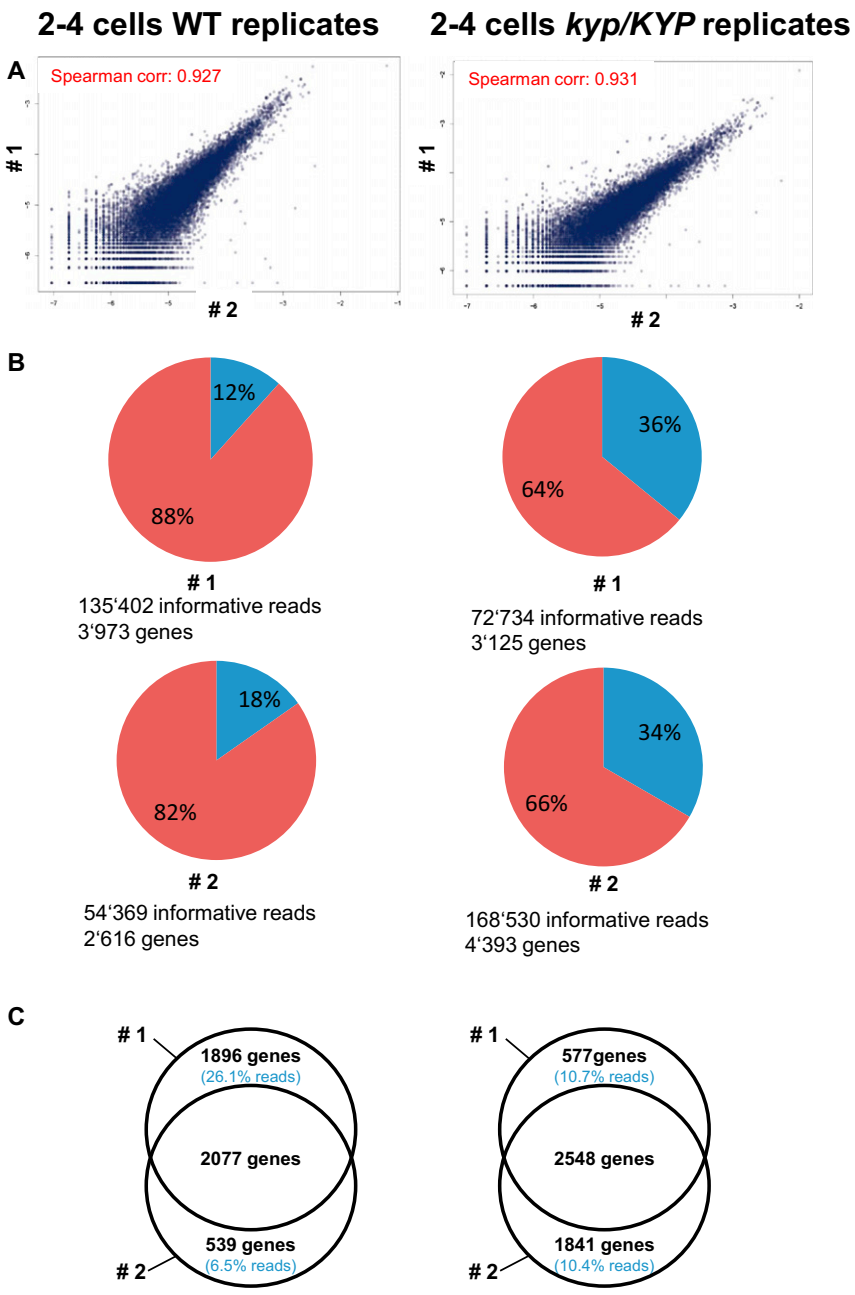


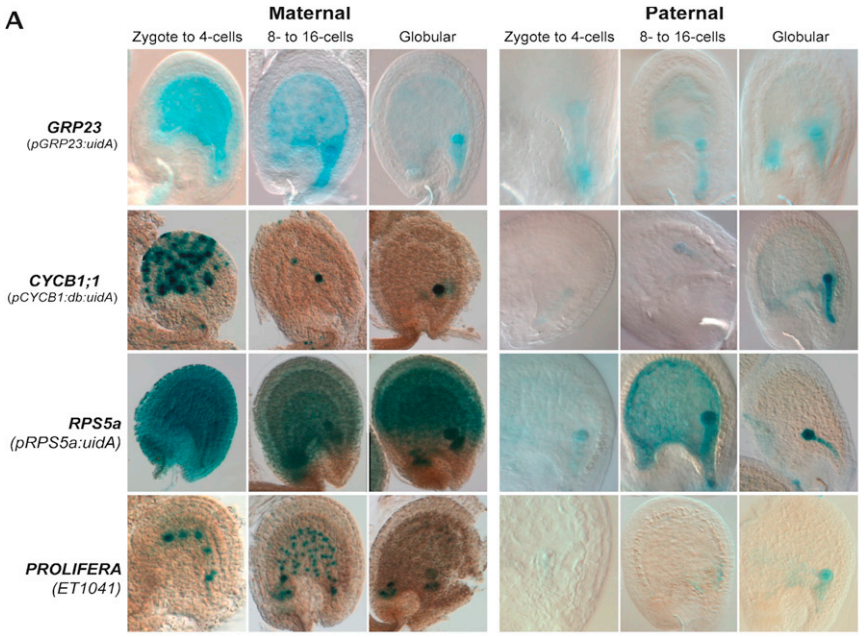
Figure S1. Replicate Transcriptome Profiling from Wild-Type and Mutant Embryos at the 2–4 Cell Stage, Related to Figure 1 and Figure 4

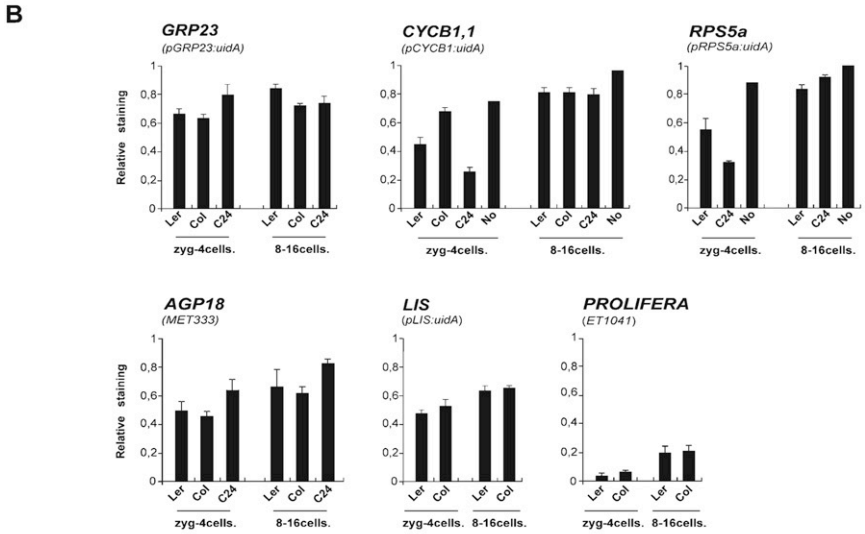
Allele-specific profiling of wild-type and *kyp/KYP* mutant embryos at the 2-4 cell stage was performed in two independent replicates (independent embryo collection, RNA extraction, amplification and sequencing).

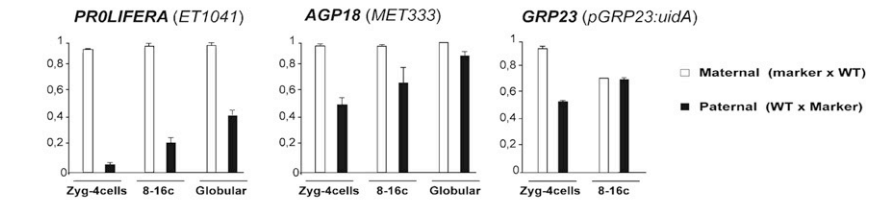
(A) The replicate transcriptomes show overall good correlation. The transcript levels describing the embryonic transcriptomes of replicates #1 and #2 were plotted in log2 scale and the Spearman correlation is indicated.

(B) The parental read distribution is similar in each replicate. The distribution of paternal (blue) and maternal (red) reads is presented as in Figure 1. the number of informative reads and genes identified is indicated. Replicates #1 are those reproduced in Figure 1. Because of better coverage in the wild-type, the allele-specific analysis was carried out in depth for replicates #1.

(C) The genes identified by informative reads overlap well between replicates. Common genes are covered by the majority of reads (89.6%–93.5% of replicates #2). The Venn diagrams show the overlap between the informative transcriptomes (i.e., identified by informative reads covering SNP regions) of each replicates. Note that the lack of overlap does not necessarily indicate lack of expression but coverage below our detection threshold on the selected SNPs (i.e., genes appearing specific to one replicate can show read coverage in regions outside the SNPs). The genes common between replicates are covered by a majority of the reads: the 2077 common genes in WT replicates are covered by 93.5% reads of replicate #2 ; the 2548 common genes in *kyp/KYP* replicates are covered by 89.6% reads of replicate #2.







	Zygote - 4 cells			8 - 16 cells			Globular		
Cross	Relative staining	SE (rep)	n	Relative staining	SE (rep)	n	Relative staining	SE (rep)	n
ET1041 x Ler	**0,93**	0,00 (3)	289	**0,95**	0,02 (2)	53	**0,96**	0,02 (2)	35
Ler x ET1041	**0,03**	0,01 (6)	207	**0,19**	0,03 (7)	129	**0,40**	0,03 (7)	305
MET333 x Ler	**0,97**	0,01 (3)	64	**0,96**	0,01 (4)	56	**1,00**	0,00 (4)	32
Ler x MET333	**0,50**	0,06 (5)	214	**0,66**	0,13 (4)	128	**0,88**	0,03 (4)	110
pGRP23:uidA x Col	**0,96**	0,01 (2)	243	**0,67**	na (1)	27			
Col x pGRP23:uidA	**0,51**	0,00 (9)	293	**0,66**	0,00 (9)	368			

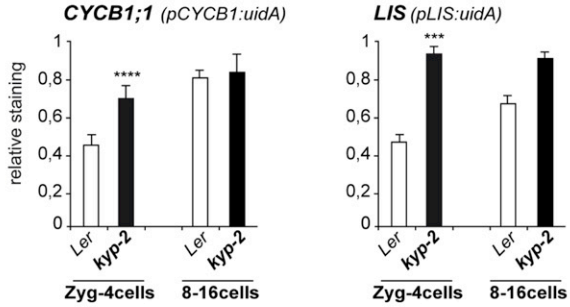
Figure S2. Maternal and Paternal Expression of Embryo Markers in Wild-Type, Related to Figure 3

(A) Differential expression of the indicated reporter transgenes (Table S3), when transmitted maternally (right panel) or paternally (left panel) was detected using histochemical detection of the *uidA* gene product (GUS) at different stages.

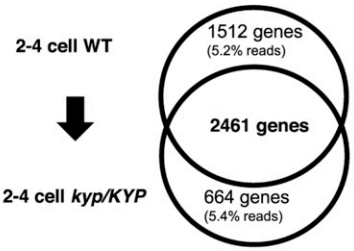
(B) Quantification of paternal expression in *Arabidopsis* accessions. The proportion of seeds showing GUS staining (relative staining) was compared in the progeny of crosses between wild-type plants from the indicated accessions: Landsberg *erecta* (L*er*), Columbia (Col), C24, Nossen (No), and the indicated marker lines. Error bars represent standard error (SE). Detailed analysis is provided in Table S4.

(C) Quantification of maternal and paternal expression. The proportion of seeds showing GUS staining (relative staining) was compared in the progeny of reciprocal crosses between wild-type and the indicated marker lines. SE, standard error. (rep), replicates. n, total number of seeds scored. See also Table S4.

A Paternal marker expression in wild-type and kyp mutant mother



B Transcriptome comparisons



C Class transitions among common genes

Number of genes

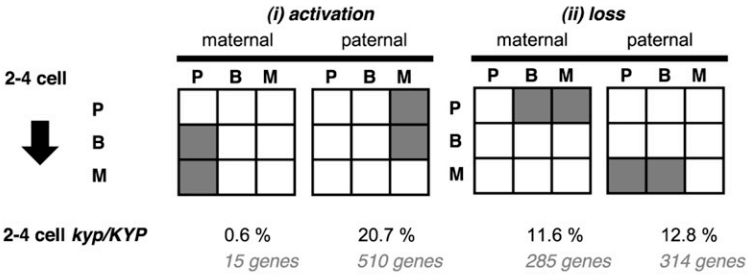
2-4 cell kyp/KYP \ 2-4 cell WT	P	B	M
P	5	174	111
B	8	1116	399
M	7	307	333

% of genes (heatmap)

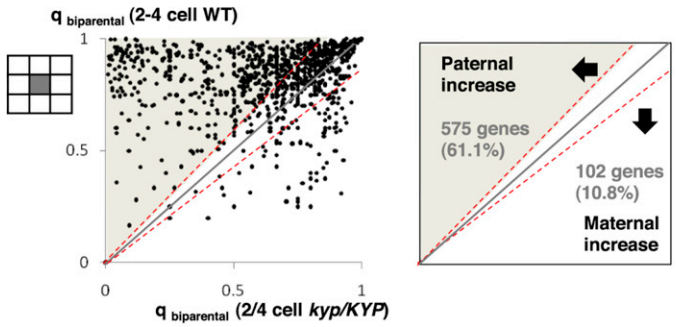
	P	B	M
P	0.2	7.1	4.5
B	0.3	45.4	16.2
M	0.3	12.5	13.6

P Paternal Class (q=0)
M Maternal Class (q=1)
B Biparental Class (0<q<1)

D Activation *vs* loss of one parental contribution



E Relative increase of paternal or maternal contribution among biparental-class genes (B > B transition)



F Summary: paternal alleles are preferentially upregulated in *kyp* embryos

	De novo detected	Relative increase (>10%)	Total
Paternal	510	575	1085 / 2461
Maternal	15	102	117 / 2461

Figure S3. Derepression of Paternal Markers and Changes of Parental Contributions Induced by a Maternal *kyp* Mutation, Related to Figure 4

(A) Embryos showing GUS staining were scored in the progeny of crosses between wild-type (L*er*) or *kyp-2* mutant females and the indicated paternal markers, at the developmental stages indicated above the graphs. Two-tailed Fisher's exact tests were carried out to assess the differences between wild-type and mutant samples. The levels of significance are indicated (*: $p < 0,05$; **: $p < 0,01$; ***: $p < 0,001$; ****: $p < 0,0001$). Error bars represent SE. Detailed analysis is provided in Table S4.

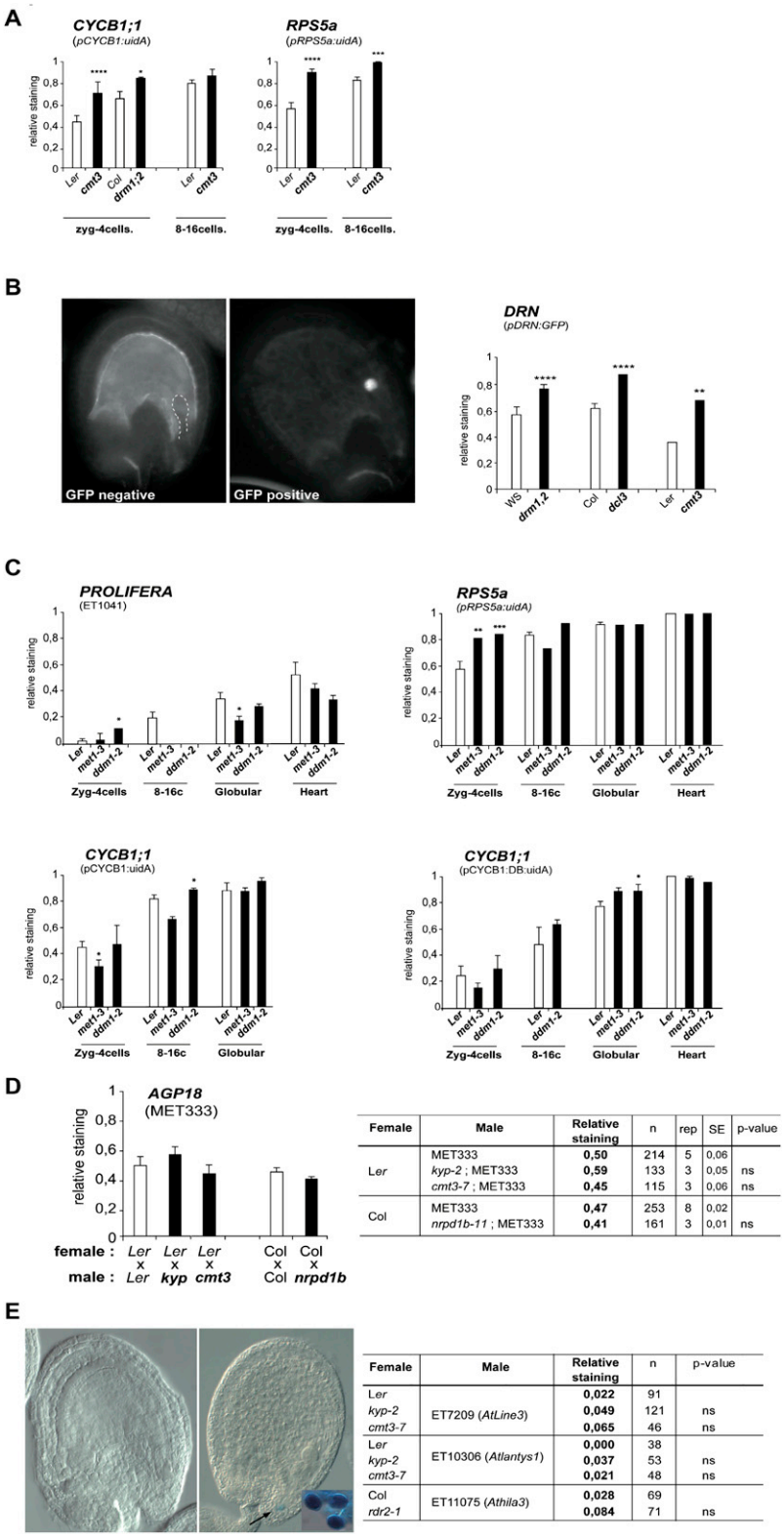
(B) Venn diagrams show the number of genes shared between the SNP-tagged transcriptomes (see Experimental Procedures) of 2-4 cell WT and 2-4 cell *kyp/KYP* embryos. Genes specifically detected in one sample only are covered by 5% reads, as indicated. Genes commonly detected in both samples, by contrast, are covered by the majority of reads (95%).

(C) The transition tables describe the changes for the common genes in the parental-class distributions (P, B, M, see legend in gray box) that were induced by a maternal *kyp* mutation (2-4 cell WT > 2-4 cell *kyp/KYP* transition). The transition matrices were calculated using parental-class probabilities (Pi, Bi, Mi) as described in Experimental Procedures.

(D) Several subsets of class transitions clearly illustrate cases of (i) activation/de novo expression and (ii) decreased expression /loss of one parental allele in our transcriptomes. Note that loss may correspond to a decrease in SNP coverage (see Experimental Procedures) falling below our detection threshold, rather than the absolute loss of transcript. The % indicated is the sum of the % genes in (C) falling in the gray transitions.

(E) A large fraction of common genes (45.4%) remain in the biparental class in both genotypes. The proportion of maternal transcripts was calculated for the 940 common genes effectively sequenced on both alleles. The calculated q values were plotted as indicated (left panel). The scheme (right panel) show the relative changes in parental contributions toward either higher paternal or higher maternal representation. Only genes with changes deviating from 10% (above/below dashed lines) were counted. 263 genes showed no or less than 10% change.

(F) The table summarizes the number of genes affected on either the maternal or the paternal allele in embryos inheriting a maternal *kyp* mutation. Shown are the genes with a novel contribution of one parental allele (de novo) in mutant embryos as reported in (D) and biparental-class genes showing a relative increase of one parental contribution as shown in E. Note that for the latter, only genes are counted that show a change >10% compared to in wild-type embryos (genes above/below the red lines).



Female	Male	Relative staining	n	rep	SE	p-value
Ler	MET333	**0,50**	214	5	0,06	
	kyp-2 ; MET333	**0,59**	133	3	0,05	ns
	cmt3-7 ; MET333	**0,45**	115	3	0,06	ns
Col	MET333	**0,47**	253	8	0,02	
	nrpd1b-11 ; MET333	**0,41**	161	3	0,01	ns

Female	Male	Relative staining	n	p-value
Ler	ET7209 (*AtLine3*)	**0,022**	91	
kyp-2		**0,049**	121	ns
cmt3-7		**0,065**	46	ns
Ler	ET10306 (*Atlantys1*)	**0,000**	38	
kyp-2		**0,037**	53	ns
cmt3-7		**0,021**	48	ns
Col	ET11075 (*Athila3*)	**0,028**	69	
rdr2-1		**0,084**	71	ns

Figure S4. Maternal *CMT3* and *DRM2*, but not *MET1* and *DDM1*, Control Early Expression of Paternally Transmitted Markers, Related to Figure 5

Additional control experiments showing maternal but not paternal effect of RdDM mutations and absence of effects on transposon enhancer trap markers.

(A) Embryos showing GUS staining were scored in the progeny of crosses between wild-type (Col or L*er*) or mutant females and the indicated paternal markers, at the developmental stages indicated above the graphs. Two-tailed Fisher's exact tests were carried out to assess the differences between wild-type and mutant samples. The levels of significance are indicated (*: $p < 0,05$; **: $p < 0,01$; ***: $p < 0,001$; ****: $p < 0,0001$). Error bars represent SE. Detailed analysis is provided in Table S4.

(B) The *drm1-1,drm2-1* double mutant line contains a *GUS* transgene (see allele reference Table S3), a GFP embryo marker was therefore used to test the role of these genes on paternal gene activation. Paternal expression of *pDRN:GFP* observed 1.5 days after pollination is shown, longer exposure was used to visualize the young embryo devoided of GFP signal (outlined). The proportion of GFP positive embryos was scored in the progeny of the indicated crosses, 1.5 days after pollination. Stage distributions were similar among mutant and control populations, ranging from the zygote to octant stages. *dcl3-1*, Col, *cmt3-7* and L*er* females were also tested in parallel as controls. Two-tailed Fisher's exact tests were carried out to assess the differences between wild-type and mutant samples. The levels of significance are indicated (**: $p < 0,01$; ***: $p < 0,001$;****: $p < 0,0001$). Error bars represent SE. Detailed analysis is provided in Table S4. Marker and mutant lines are described in Table S3.

(C) Embryos showing positive GUS staining were scored in the progeny of crosses between wild-type (L*er*), *met1-3*/*MET1* or *ddm1-2*/*DDM1* females and the indicated paternal markers, at different developmental stages as noted below graphs. Two-tailed Fisher's exact tests were carried out to assess the differences between wild-type and mutant samples. The levels of significance are indicated (*: $p < 0,05$; **: $p < 0,01$). Error bars represent SE. Detailed analysis is provided in Table S4.

(D) Absence of paternal effects of *kyp* and *nrpd1b* mutations on paternally transmitted markers. The proportion of seeds showing GUS staining (relative staining) was compared in the progeny of crosses between wild-type females and male with or without the mutations, together with the paternal markers lines, as indicated. n, total number of seeds scored. Two-tailed Fisher's exact tests showed no significant differences (ns). Error bars represent SE.

(E) Paternally transmitted transposon enhancer trap lines are not reactivated in seeds derived from *kyp*, *cmt3* or *rdr2* mothers. Paternally transmitted enhancer trap line *ET7209* monitoring expression of *AtLine3* family transposon (Slotkin et al., 2009), as for *ET10306* (*Atlantys1*) and *ET11075* (*Athila3*), did not show expression in the early seed in most cases, as described (Slotkin et al., 2009) (left picture). Faint staining at the suspensor base was detected occasionally (right picture, arrow), and scored as GUS positive. Strong pollen staining was detected in all lines (inset in left picture). The table displays the proportion of early seeds (zygote to globular stages) showing GUS staining in crosses between wild-type or mutant females and three transposon GUS marker lines, as indicated. Staining in the embryo proper was never observed, and the percentage of seed with signal at suspensor base (see right picture) remained below 10% for all mutants tested. n, total number of seeds scored. Two-tailed Fisher's exact tests showed no significant differences (ns). Error bars represent SE.

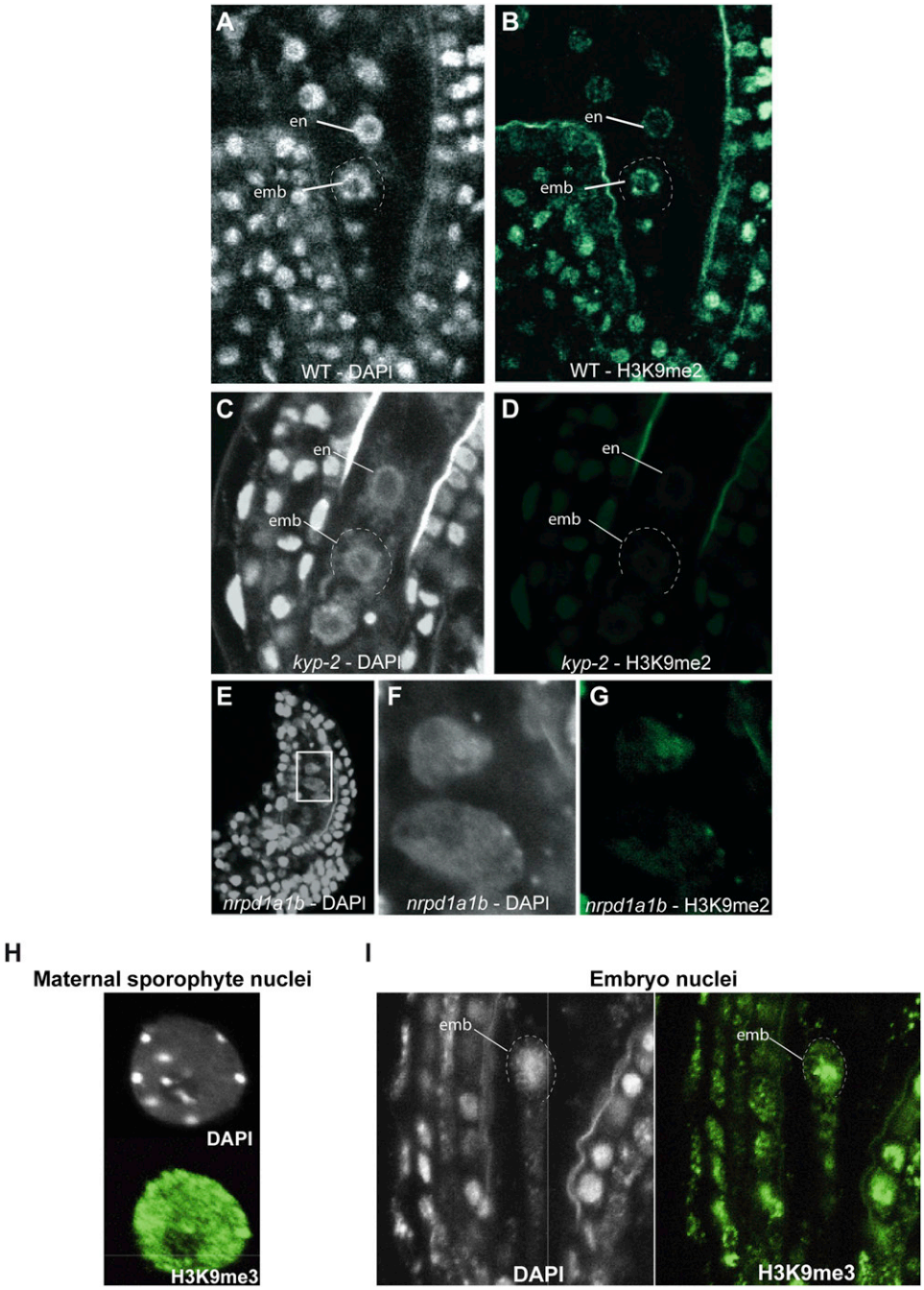
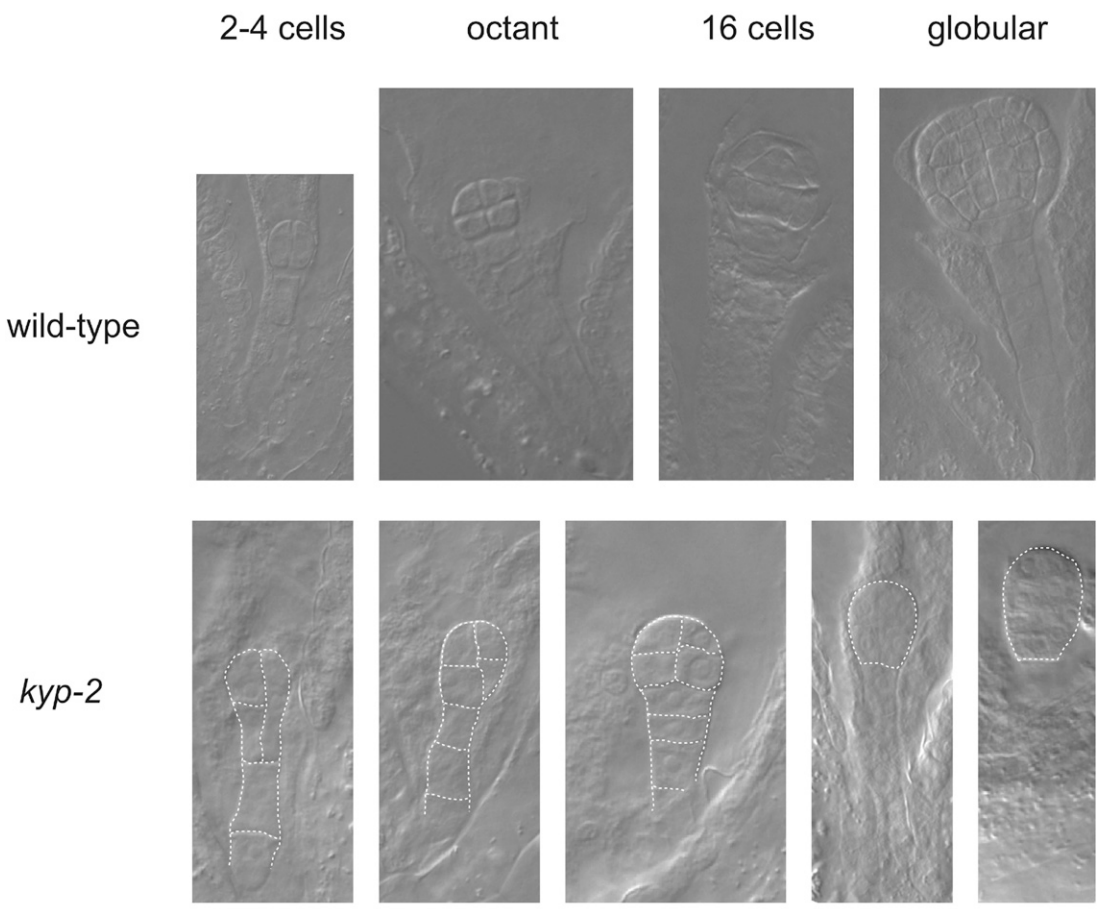


Figure S5. Control Immunodetection of H3K9me2 Showing Absence of Signals in *kyp* Embryos and Altered Distribution of *nrpd1* Mutant Embryos, by Contrast to H3K9me3 which Remain Unaffected in Both Mutants, Related to Figure 5

(A–G) Immunolocalization of histone H3 di-methylation on lysine 9 (H3K9me2) shows a strong signal in heterochromatic chromocenters in wild-type (A and B), while almost no signal in *kyp-2*, as described previously (Jackson et al., 2004) (C and D), and an altered, dispersed signal in *nrpd1a-1/nrpd1b-11* (*nrpd1a1b*) double mutant nuclei (E–G).

(H and I) Immunolocalization of histone H3 tri-methylation on lysine 9 (H3K9me3) showed that, by contrast to H3K9me2, H3K9me3 is present in both somatic (tegument, H) and embryonic chromatin (I). Specifically, it is homogenously distributed in euchromatin as described previously (Turck et al., 2007; Sung et al., 2006) in both wild-type, *kyp-2* and *nrpd1a-1/nrpd1b-11* backgrounds. en, endosperm. emb, embryo. The DNA was counterstained with DAPI.



	abnormal embryos	n	rep.
L*er*	2,6 % (+/- 0,2)	421	3
kyp-2	11,8 % (+/- 3,5)	260	3

Figure S6. Transient Patterning Defects in *kyp* Mutant Embryos, Related to Figure 4

Wild-type (L*er*) and *kyp-2* mutant early embryos were cleared using Herr's solution. Representative examples of the abnormal phenotypes observed in *kyp-2* are shown. Wild-type and abnormal phenotypical classes were scored in independent plants (rep.).

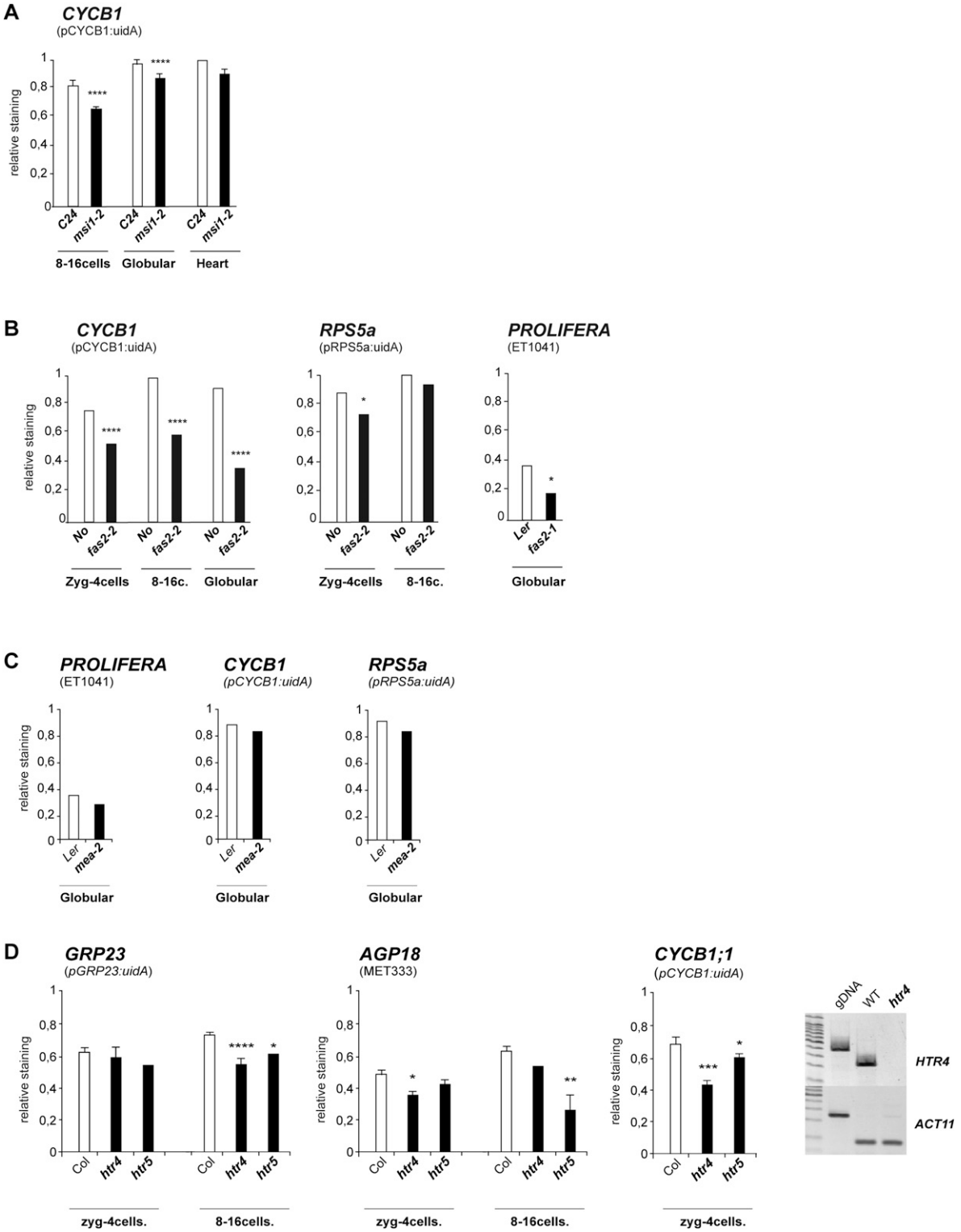


Figure S7. Maternal *MSI1* and *FAS2,* but not *MEA*, Control the Activation of Paternally Transmitted Markers and Require Maternal H3.3 Variants, Related to Figure 7

(A) Embryos showing GUS staining were scored in the progeny of crosses between wild-type (*Ler*), *msi1-2* females and the indicated paternal markers, at different developmental stages. Two-tailed Fisher's exact tests were carried out to assess the differences between the wild-type and mutant samples. The levels of significance are indicated (*: $p < 0,05$; **: $p < 0,01$; ***: $p < 0,001$; ****: $p < 0,0001$). Error bars represent SE. Detailed analysis is provided in Table S4.

(B) Embryos showing GUS staining were scored in the progeny of crosses between wild-type (No) or *fas2* females and the indicated paternal markers, at different developmental stages. Quantifications displayed as above. Detailed analysis is provided in Table S4.
(C) *MEA* does not delay expression of paternally transmitted markers at the globular stage. Embryos showing GUS staining were scored in the progeny of crosses between wild-type (L*er*) or *mea-2*/ *MEA* females and the indicated paternal markers, at the globular stage. Two-tailed Fisher's exact tests showed no significant differences. Detailed analysis is provided in Table S4.
(D) Maternal histone H3.3 variants control the activation of paternal markers. The graphs show embryos scored for GUS staining following crosses between wild-type (Col) or mutant females and the indicated paternal markers. Two-tailed Fisher's exact tests were carried out to assess the differences between wild-type and mutant samples as in A. Error bars represent SE. *htr4* and *htr5* are loss of function mutants in the *HTR4* and *HTR5* genes, which are close homologs encoding a H3.3 variant and likely show functional redundancy (Okada et al., 2005, Table S3). Absence of *HTR4* (*At4g40030*) mRNA in *htr4-1* insertion allele was confirmed by RT-PCR (see gel picture) performed on young siliques from wild-type (WT) and *htr4-1* (*htr4*) plants using intron-spanning primers. *ACTIN11* (*ACT11*) was amplified as a loading control. Amplification on genomic DNA (gDNA) is shown in the gel picture to the left.