Genome-Wide Transcript Profiling of Endosperm without Paternal Contribution Identifies Parent-of-Origin-Dependent Regulation of AGAMOUS-LIKE36

Reza Shirzadi1, Ellen D. Andersen1, Katrine N. Bjerkan1, Barbara M. Gloeckle2, Maren Heese2, Alexander Ungru3, Per Winge4, Csaba Koncz5,6, Reidunn B. Aalen1, Arp Schnittger2,3, Paul E. Grini1*

1 Department of Molecular Biosciences (IMBV), University of Oslo, Oslo, Norway, 2 Department of Molecular Mechanisms of Phenotypic Plasticity, Institut de Biologie Moléculaire des Plantes (IBMP), UPR2357 du CNRS, Université de Strasbourg, Strasbourg, France, 3 Ungruppe am Max-Planck-Institute for Plant Breeding Research, Department of Botany III, University of Cologne, Cologne, Germany, 4 Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway, 5 Max-Planck-Institute for Plant Breeding Research, Cologne, Germany, 6 Institute of Plant Biology, Biological Research Center of Hungarian Academy of Sciences, Szeged, Hungary

Abstract

Seed development in angiosperms is dependent on the interplay among different transcriptional programs operating in the embryo, the endosperm, and the maternally-derived seed coat. In angiosperms, the embryo and the endosperm are products of double fertilization during which the two pollen sperm cells fuse with the egg cell and the central cell of the female gametophyte. In Arabidopsis, analyses of mutants in the cell-cycle regulator CYCLIN DEPENDENT KINASE A;1 (CDKA;1) have revealed the importance of a paternal genome for the effective development of the endosperm and ultimately the seed. Here we have exploited cdka;1 fertilization as a novel tool for the identification of seed regulators and factors involved in parent-of-origin-specific regulation during seed development. We have generated genome-wide transcription profiles of cdka;1 fertilized seeds and identified approximately 600 genes that are downregulated in the absence of a paternal genome. Among those, AGAMOUS-LIKE (AGL) genes encoding Type-I MADS-box transcription factors were significantly overrepresented. Here, AGL36 was chosen for an in-depth study and shown to be imprinted. We demonstrate that AGL36 parent-of-origin-dependent expression is controlled by the activity of METHYLTRANSFERASE1 (MET1) maintenance DNA methyltransferase and DEMETER (DME) DNA glycosylase. Interestingly, our data also show that the active maternal allele of AGL36 is regulated throughout endosperm development by components of the FIS Polycomb Repressive Complex 2 (PRC2), revealing a new type of dual epigenetic regulation in seeds.

Citation: Shirzadi R, Andersen ED, Bjerkan KN, Gloeckle BM, Heese M, et al. (2011) Genome-Wide Transcript Profiling of Endosperm without Paternal Contribution Identifies Parent-of-Origin–Dependent Regulation of AGAMOUS-LIKE36. PLoS Genet 7(2): e1001303. doi:10.1371/journal.pgen.1001303

Editor: Mathilde Grelon, Institut Jean-Pierre Bourgin, INRA de Versailles, France

Received June 10, 2010; Accepted January 11, 2011; Published February 17, 2011

Copyright: © 2011 Shirzadi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Services provided by the Norwegian Arabidopsis Research Centre, which is a part of the Research Council of Norway’s National Program for Research in Functional Genomics (www.fuge.no). RS, EDA, KNB, and PEG were supported by grants from the Norwegian Research Council (www.rcn.no) and ERA-Net Plant Genomics (www.erapg.org) (166057/V40, 183190/S10, 182903/S10). AU is a fellow of the Konrad-Adenauer-Stiftung. BMG is supported by a PhD fellowship of the Studienstiftung des deutschen Volkes. This work was supported by an EMBO Young Investigator grant to AS, an ATIP grant from the Centre National de la Recherche Scientifique to AS, an ERC starting grant from the European Union to AS, a Norwegian Research Council FUGE Young Investigator Starting Grant to PEG, and an ERA-Net Plant Genomics Grant to AS and PEG. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: p.e.grini@imbv.uio.no

Introduction

Seed development is a tightly regulated process that is controlled, both before and after fertilization and requires tight coordination of parental gene expression [1]. A paradigm for the importance of balanced parental contribution is the observation that certain genes in the developing offspring of flowering plants are exclusively or preferentially expressed from only one of the two parental genomes, a phenomenon called genomic imprinting that has also been observed in mammals [2,3]. The relevance of parent-of-origin effects was first found in interploidy crosses [4]. Typically, an increase in the paternal genome results in larger seeds, while the opposite is observed if the maternal gene dosage is higher than normal [5]. This is in agreement with the parental conflict theory, which implies that fathers direct maximal amount of maternal resources to their own offspring and thereby promote growth. Mothers on the other hand would seek to distribute the resources equally among all their offspring, and balance their resource between themselves and their offspring. Thus, maternal factors are thought to dampen growth [6].

In mammals, imprinted genes are often involved in growth control [7–10]. In Arabidopsis, the endosperm is the major tissue regulating the flow of nutrients to the embryo, and is therefore a likely site for parent-of-origin dependent gene expression.

Imprinting results from differences in epigenetic marks, involving DNA methylation and post-translational modifications of histones on the parental alleles [11,12]. Trimethylation of lysine 27 on histone H3 (H3K27me3) leading to repression of gene expression, has been found to be a particularly important imprinting mechanism in plants. In Arabidopsis seeds, H3K27me3 mark is set by the FIS Polycomb Repressive Complex 2 (PRC2), which consists of at least four components; the histone methyl-
Author Summary

Seeds of flowering plants consist of three different organisms that develop in parallel. In contrast to animals, a double fertilization event takes place in plants, producing two fertilization products, the embryo and the endosperm. Imprinting, the parent-of-origin-specific expression of genes, typically takes place in the mammalian placenta and in the plant endosperm. A prevailing hypothesis predicts that a parental tug-of-war on the allocation of available resources to the developing progeny has led to the evolution of imprinting systems where genes expressed from the mother dampen growth whereas genes expressed from the father are growth enhancers. The number of imprinted genes identified in plants is low compared to mammals, and this precludes the elucidation of the epigenetic mechanisms responsible for this specialized expression system. Here, we have used genome-wide transcript profiling of endosperm without paternal contribution to identify seed regulators and, among these, imprinted genes. We identified a cluster of downregulated MADS-box transcription factors, including AGL36, that was subsequently shown to be imprinted by an epigenetic mechanism involving the DNA methylase MET1 and the glycosylase DME. In addition, the expression of the active AGL36 allele was dampened by the FIS Polycomb Repressive Complex, identifying a novel mode of regulation of imprinted genes.

transferase MEDEA (MEA), FERTILIZATION INDEPENDENT SEED 2 (FIS2), FERTILIZATION INDEPENDENT ENDOSPERM (FIE), and MULTICOPY SUPPRESSOR OF IRA 1 (MS1). The corresponding genes were identified in screens for autonomous endosperm development, indicating that the FIS complex acts as a repressor of endosperm development prior to fertilization [13–17].

An equally important regulatory mechanism in imprinting is DNA methylation resulting from the activity of several different DNA methyltransferase enzymes, where each has specificity for cytosine (C) in certain sequence contexts. So far, imprinting has been shown to be under the influence of MET1, the major Arabidopsis maintenance DNA methyltransferase involved in CG-methylation [11,18–20]. DNA demethylation can be achieved either by a passive process i.e. the repression of MET1 expression [21,22], or by an active mechanism involving DNA glycosylase enzymes such as DME [23]. Several lines of evidence show that DME, which is expressed in the central cell of the female gametophyte, is necessary for maternal-specific gene expression in the endosperm [11,18,19,24].

So far, only about a dozen genes in Arabidopsis have been identified to have parental-specific gene expression, and they illustrate different modes of imprinting [3]. MEA, ARABIDOPSIS FORMIN HOMOLOGUE 5 (AFH5) and PHERES 1 (PHR1) are imprinted by the action of FIS PRC2, where only the latter is paternally expressed [13,25–31]. FIS2, FLOWERING WAGENINGEN (FWA) and MATERNALLY EXPRESSED PAB C-TERM (MPC) are all maternally expressed and regulated by the dual action of MET1 and DME [11,19,24,32–34]. Recently, five novel imprinted genes, HOMEODOMAIN GLABROUS 3 (HDG3), HDG8, HDG9, At5g62110 and ATM1B52R2 were identified by differential DNA methylation in embryo and endosperm [35].

In comparison to Arabidopsis, more than 100 genes have been shown to have a uniparental or preferential parental expression pattern in mammals [36–39]. This suggests that additional genes in Arabidopsis are imprinted. Furthermore, the low number of known imprinted genes in plants precludes the identification of general principles in this kind of gene expression control and thus, the identification of further imprinted genes is pivotal. Moreover, the targets of imprinted genes, as well as genomic pathways and regulatory modules influenced by imprinted genes are largely unknown.

Here, we have designed a microarray strategy for the identification of seed regulators by exploiting the cdka;1 mutation. Using this approach, we have identified a cluster of previously uncharacterized AGAMS-LIKE (AGL) Type-I MADS-box transcription factors that are downregulated in endosperm with no paternal contribution. Here, we report that AGL36 is imprinted by the dual action of MET1 and DME. In addition, AGL36 is regulated throughout endosperm development in its maternal expression cycle by the Polycomb FIS-complex, thereby identifying a novel mode of regulation for imprinted genes.

Results
cdka;1 is a tool to identify key seed regulators

Here we have used cdka;1 as a tool to identify factors sensitive to the vital parental gene balance in the endosperm. In heterozygous cdka;1 mutants, the second pollen mitosis is either missing or is severely delayed. However, mutant pollen can successfully fertilize the egg cell while leaving the central cell unfertilized [40,41]. A detailed analysis by Aw and colleagues has revealed that a second sperm cell is delivered to the central cell, but that karyogamy does not take place [42]. Although not properly fertilized, the majority of the central cells in cdka;1 fertilized ovules (70–90%) are triggered to initiate endosperm proliferation [40,42,43]. Thus, fertilization by cdka;1 sperm cells creates a unique situation where endosperm initially develops without any paternal contribution (in the following also referred to as cdka;1P). The endosperm, however remains under-developed, and ultimately the seed aborts, further demonstrating the importance of the paternal contribution to the endosperm for proper seed development. Since activation of maternal alleles by loss of maternal FIS PRC2 could rescue seed lethality [43], we hypothesized that the disturbance of parental gene balance in the endosperm is the main cause leading to developmental arrest of cdka;1P at 3–4 days after pollination (DAP).

To identify factors and mechanisms sensitive to such an imbalance in gene dosage in the endosperm and with that likely key regulators of seed development, we performed microarray transcript profiling of cdka;1 fertilized seeds at 3 DAP (Figure S1A). Due to the heterozygous nature of the cdka;1 mutant line used, a transcript that is absent in cdka;1P seeds will lead to a reduction of maximal 50% in the genome profiling experiment. For example, genes that are only expressed from the paternal genome would show such reduced expression levels (Figure S1B). Likewise, maternally expressed genes that require activation by a paternally expressed gene(s) would be downregulated (Figure S1C), whereas genes that are acted upon by paternally expressed repressors were expected to be upregulated in the microarray screen (Figure S1D).

When we compared the transcriptional profiles of Ler x cdka;1 versus Ler x Col seeds 3 DAP, we detected 17223 nuclear genes that were expressed in all biological replicates of both mutant (cdka;1 set) and wild-type (WT set) seed profiles. Our result is in good agreement with a set of genes identified by Goldberg & Harada laboratories (GH) in globular stage seeds of Arabidopsis Ws-0 plants as 68% of our genes were also identified by GH, and our gene set included >90% of the GH globular seed gene set (Figure 1A; http://seedgenenetwork.net, [44]).

To further validate the quality of our dataset, we examined the expression pattern of genes known to be preferentially expressed
Figure 1. Analysis of cdka;1 microarray profiles. (A) Venn diagram representing overlap of genes expressed in globular stage seeds of Arabidopsis Ws-0 plants (red) and genes expressed in 3 DAP seeds from Ler plants pollinated with Col cdka;1 pollen (green). As high as 67.8% of the cdka;1 set and 90.7% of the GH set genes were found in the overlap. Gene numbers refer to the reference set of genes (see material and methods). GH: Goldberg & Harada laboratories (http://www.seedgenenetwork.net). (B) Boxplot showing the reduced relative expression of known maternally imprinted (blue background) and paternally imprinted (pink background) genes in the Ler x cdka;1 versus Ler x Col seeds. Calculations are based on values taken from three independent biological replicas. (C) GO functional classification of microarray expression data. Deregulated genes identified in the microarray experiment were functionally classified regarding their molecular function using the GO Slim classification system (http://www.arabidopsis.org/tools/bulk/go/). The total number of unique GO-term:locus assignments for each group is indicated (n). The functional classifications of all genes present on the microarray (On microarray) and all genes having a present call (All expressed) have been included for comparison. The cut-off for deregulation is ≤ 0.8 for the downregulated group, and ≥ 1.5 and ≥ 1.2 for the upregulated groups. doi:10.1371/journal.pgen.1001303.g001
from the paternal allele. To date, only three genes have been identified that show a predominant paternal expression pattern; PHE1, HDG3 and AGL62110, where all three genes were found to be downregulated in our arrays (Figure S1E), supporting our working hypothesis that paternally expressed genes can be detected amongst downregulated genes. In addition, out of seven imprinted maternally expressed genes present in our microarray sets, four were also detected as downregulated (Figure S1E). This could reflect required activation by maternal factors (Figure S1C), or be a result of more complex deregulation in response to change in gene dosage. To exclude array artifacts we tested all down-regulated genes by means of real-time PCR and could confirm their deregulation (Figure 1B).

Due to the background noise in the microarray experiment, modest but reproducible downregulation of arithmetic ratios (ar) ranging from 0.5 to 1.0 will produce False Discovery Rates (FDR, see materials and methods) with insignificant q values. Since the absence of paternally expressed genes was the simplest hypothesis to account for downregulation, we defined a functional limit for screening purposes that allowed us to detect two out of three known paternally expressed genes in the array. Both PHE1 and HDG3 are detected at q values of 0.35 and a downregulation cutoff of 0.8 (ar). Consequently these values were chosen and used to filter the microarray data.

Using these criteria, a set of 602 genes was extracted (q≤0.35 and ar ≤0.8), subsequently called Down 0.8. For upregulation, we worked with two gene sets. For the first set, Up 1.2, we used parameters equivalent to the downregulated set (q≤0.35 and ar ≥1.2), which resulted in a set of 1030 genes. For the second set, Up 1.5, resulting in 323 genes, we chose ar ≥1.5, a threshold for deregulation commonly used in genome-wide expression studies (Table S3).

To test whether the deregulated genes could preferentially be attributed to a certain seed structure, we compared our data to gene sets expressed in different seed regions and compartments of globular stage seeds using data generated by Goldberg & Harada (GH) laboratories available at http://seedgenenetwerk.net [44]. The overlap between the upregulated gene sets and the GH embryo, seed coat and endosperm was significantly lower than expected for independent sets of genes, indicating that among the upregulated genes we preferentially find those that are below the detection limit of the GH analyses. However looking at the downregulated genes, the picture was different. While we found slightly less overlap than expected by chance for the GH embryo set, the overlap was clearly larger than expected by chance for the GH embryo, slightly less overlap than expected by chance for the GH embryo set. While we found slightly less overlap than expected by chance for the GH embryo set, the overlap was clearly larger than expected by chance for the GH embryo set.

The overlap between the upregulated gene sets and the GH embryo, seed coat and endosperm was significantly lower than expected for independent sets of genes, indicating that among the upregulated genes we preferentially find those that are below the detection limit of the GH analyses. However looking at the downregulated genes, the picture was different. While we found slightly less overlap than expected by chance for the GH embryo set, the overlap was clearly larger than expected by chance for the GH embryo set.

Analysis of AGL36 imprinting by MET1

AGL36 Imprinting by MET1

AGL36 is only expressed from its maternal allele

Real-time PCR measurement of AGL36 relative expression level three days after pollination (3 DAP) in Ler ovules fertilized with either Col or cdka;1 pollen confirmed that AGL36 expression was reduced in cdka;1 fertilized seeds, (27% when normalized towards ACT11, and 36% when normalized towards GAPA) compared to wild-type seeds (Figure 2A).

To determine whether AGL36 has parental-specific expression, we took advantage of an AGL36 Single Nucleotide Polymorphism (SNP) existing between the Col and Ler ecotypes. This SNP allows the PCR product of Col cDNA to be digested by AluNI, leaving the Ler cDNA PCR product intact (Figure 2B). We performed reciprocal crosses between Col and Ler ecotypes, and analyzed the digested RT-PCR fragments on an Agilent Bioanalyzer Lab-on-a-Chip, allowing accurate measurement of fragment sizes and their concentrations. When Col paternal is crossed with Ler paternal, we only detected the Col bands (165 bp+234 bp) after AluNI digestion, indicating only maternal expression (Figure 2C). Similarly, in the reciprocal cross when Ler maternal is fertilized with Col paternal pollen, the cDNA PCR digest resulted only in an undigested band (399 bp) originating from Ler, indicative of maternal expression (Figure 2C). This testified that AGL36 was only expressed from the maternal genome after fertilization and thus identified as a novel imprinted gene.

AGL36 is imprinted throughout early seed development

AGL36 expression level in wild-type seeds (Ler x Col) at different stages of seed development was monitored over a period of 12 days after pollination. Initially, a low expression level was detected (1 DAP), followed by a rapid increase and subsequent peak in AGL36 expression at 4 DAP, when the embryo is at the late globular stage of development, before declining (Figure 3A). At the embryo heart stage, corresponding to 6 DAP, AGL36 expression had decreased to similar levels as 1 DAP. To address whether AGL36 imprinting is maintained throughout its expression cycle, we performed a SNP analysis of the RT-PCR product obtained from Ler x Col crosses harvested during 1 to 12 DAP (Figure 3B). We found that AGL36 expression is originating from the maternal genome (Ler) throughout the experiment. By plotting the molarities of the maternal band obtained by Agilent Bioanalyzer, an expression profile closely identical to the pattern obtained in the real-time PCR analysis was found (Figure 3C).

To rule out that the observed maternal expression is due to expression of AGL36 in the ovule integument, which is a maternal tissue, we generated a reporter construct consisting of 1752 bp of the AGL36 promoter region fused to a GUS reporter (pAGL36::GUS) (Figure 4A). Single-copy lines carrying this construct were
used in reciprocal crosses with wild-type Ler and Col plants to examine GUS expression at 3 and 6 DAP. When inherited maternally, pAGL36::GUS expression in the seed was indeed found to be restricted only to the fertilization product (Figure 4B, Figure S7D). In the reciprocal cross, when pAGL36::GUS was inherited from the paternal genome, no GUS expression was detected, (Figure 4C, Figure S7E). Consistent with the SNP analysis, this demonstrated that AGL36 was imprinted and only maternally active throughout its expression cycle. Furthermore, the 1.7 Kb promoter fragment used in this analysis appears to be sufficient to confer parent-of-origin specific expression of the reporter.

**AGL36 is not required for seed survival**

To further investigate the biological function of AGL36, we screened the Koncz T-DNA collection for insertions [48]. We identified a mutant line, agl36-1, harboring a single T-DNA insert in the 3’UTR of AGL36, denoted as agl36-1. To determine the impact of this mutation on seed survival, we harvested seeds from agl36-1 and wild-type plants and compared their germination rates. The results showed no significant difference in seed germination between agl36-1 and wild-type, indicating that AGL36 is not required for seed survival.

**Figure 2. AGL36 is only expressed from the maternal genome.**

(A) Real-time PCR analysis showing AGL36 expression in Ler x cdka1 relative to Ler x Col seeds 3 DAP. Gray bars represent Ler x Col expression levels, black bars represent the Ler x cdka1 expression levels. Left section: AGL36 normalized to ACT11 levels. Right section: AGL36 normalized to GAPA levels. Average values from three independent biological replicates are shown. Error bars indicate standard deviation (STDEV). (B) Schematic overview of AGL36 SNP analysis. The presence of a SNP between Col and Ler ecotypes (C-T conversion respectively) allows the amplified AGL36 cDNA PCR product from the Col ecotype to be digested with AlwNI restriction enzyme, while the Ler ecotype remains undigested. (C) AGL36 is maternally expressed. Seeds obtained from Col x Ler and Ler x Col crosses were harvested at 3 DAP followed by AGL36 RT-PCR, AlwNI digestion and subsequent Bioanalyzer analysis. Genomic Col and Ler were included as controls (Left section, first two lanes). Digestion products of two independent biological replicates of maternal Col x Ler pollen crosses produced only Col bands, indicating maternal expression (Left section). The intensities of the bands are represented as concentrations (nmol/L), and create a basis for comparison. 100 ng DNA was used as template for each PCR reaction.

doi:10.1371/journal.pgen.1001303.g002

**Figure 3. AGL36 is imprinted throughout its expression cycle.**

(A) AGL36 expression profile. Calculations were done using 3 DAP values as reference point, giving other obtained expression values relative to the 3 DAP expression level. Samples were taken at 1, 2, 3, 4, 6, 9, and 12 DAP. The graph represents the average relative expression values obtained from two independent biological parallels where the RNA from each biological sample gave rise to two independent cDNA syntheses (technical replica). The indicated STDEV is derived from the two independent biological parallels. The AGL36 transcript levels were normalized to ACT11 levels. (B) RT-PCR digest of the SNP containing region analyzed by the Bioanalyzer show that AGL36 imprinting is maintained throughout seed development. Samples were taken at time-points as indicated for each lane. A representative light micrograph of each DAP stage is shown. Only maternal (Ler) AGL36 expression was found when present. Genomic Ler and Col DNA were included as controls. 100 ng DNA/cDNA was used as template for each PCR reaction. The intensities of the bands are represented as concentrations (nmol/L). Note, weak paternal bands obtained at 2 DAP were below the detection limit for measurement on our instrument (0.1 ng/µL 0.4 nmol/L). Intensities below the detection point of the instrument are indicated as b.d. The displayed SNP picture is representing one of four independent runs (2BR and 2TR). (C) Visual representation of the obtained intensities of maternal bands in B) represented as concentrations (nmol/L).

doi:10.1371/journal.pgen.1001303.g003

**AGL36 Imprinting by MET1**

used in reciprocal crosses with wild-type Ler and Col plants to examine GUS expression at 3 and 6 DAP. When inherited maternally, pAGL36::GUS expression in the seed was indeed found to be restricted only to the fertilization product (Figure 4B, Figure S7D). In the reciprocal cross, when pAGL36::GUS was inherited from the paternal genome, no GUS expression was detected, (Figure 4C, Figure S7E). Consistent with the SNP analysis, this demonstrated that AGL36 was imprinted and only maternally active throughout its expression cycle. Furthermore, the 1.7 Kb promoter fragment used in this analysis appears to be sufficient to confer parent-of-origin specific expression of the reporter.

**AGL36 is not required for seed survival**

To further investigate the biological function of AGL36, we screened the Koncz T-DNA collection for insertions [48]. We identified a mutant line, agl36-1, harboring a single T-DNA insert in the 3’UTR of AGL36, denoted as agl36-1. To determine the impact of this mutation on seed survival, we harvested seeds from agl36-1 and wild-type plants and compared their germination rates. The results showed no significant difference in seed germination between agl36-1 and wild-type, indicating that AGL36 is not required for seed survival.
insertion 16 bp upstream of the AGL36 ATG start codon (Figure S3A). The agl36-1 line showed Mendelian segregation of the T-DNA insertion, as 75% of the plants were resistant to Hygromycin (N = 1025, χ² = 0.83, Table S1).

To test the transmission through the male and female gametes directly, reciprocal crosses of both hemizygous and homozygous agl36-1 mutant plants with wild-type plants were performed (Table S1). In a reciprocal cross, a hemizygous mutant will segregate 50% of the T-DNA resistance marker if the disrupted gene is not vital for gametophyte transmission or function. Thus, gametophyte transmission of the T-DNA resistance marker if the disrupted gene is not vital for gametophyte transmission or function. Thus, gametophyte transmission would result in 50% resistance and 50% sensitivity in the progeny. If the gene is required for gametophyte transmission, the progeny would consist of 100% resistance or 0% resistance. Therefore, the transmission of the T-DNA resistance marker can be used to determine whether a gene is required for gametophyte transmission. In a reciprocal cross with agl36-1, 75% of the plants were resistant to Hygromycin (N = 1025, χ² = 0.83, Table S1).

The position of the T-DNA insertion in agl36-1 predicts AGL36 expression failure, and indeed real-time PCR analyses of 3 DAP seeds of homozygous agl36-1−/− plants compared to Col wild-type indicate a 1000-fold AGL36 downregulation in the mutant seeds (Figure S5B). In line with an imprinted and maternal-only expression of AGL36, close to 50% reduction of the transcript level was observed in 3 DAP hemizygous agl36-1−/+ seeds (Figure S5B). We thereby concluded that agl36-1 represents a loss-of-function allele of AGL36.

Although depletion of AGL36 did not interfere with the fitness of the mutant allele in our experimental system, we have shown that AGL36 is specifically expressed from the maternal allele in the fertilization product, in a time frame between 2 and 6 DAP. To investigate whether this was reflected morphologically or developmentally in the developing seed, we compared embryo and endosperm development in wild-type and homozygous agl36-1−/− seeds within the AGL36 expression time frame.

After fertilization of the egg and the central cell, the endosperm in Arabidopsis undergoes three syncytial rounds of nuclear divisions before the first asymmetric division of the zygote that creates the apical embryo proper and the basal suspensor that connects the embryo proper and the maternal tissue (Figure S5C). At the 2 DAP stage, no obvious difference could be observed between wild-type and agl36-1−/− seeds, both typically harboring a 1–2 cell embryo proper and a 16–32 nucleated endosperm (Figure S5C, left section). The embryo continues to divide through radial, longitudinal and transverse divisions to produce the so-called globular stage at 4 DAP (Figure S5C, middle section). The endosperm also undergoes 3–4 syncytial nuclear divisions and remains uncellularized as cell proliferation at the upper half of the embryo forms the cotyledon primordia at the so-called heart stage at 6 DAP (Figure S5C, right section). Although the main AGL36 expression peak occurs during this time frame, no obvious deviation between wild-type and agl36-1−/− could be observed at these stages. Similarly, using an endosperm specific pFIS2::GUS reporter [33], a wild-type endosperm division pattern was observed in agl36-1−/− seeds (Figure S5D).

**MET1 is required for AGL36 imprinting**

The majority of imprinted, maternally expressed genes identified in Arabidopsis so far have been shown to be paternally silenced by mechanisms involving symmetric CG methylation, maintained by MET1 [11, 18, 19]. Although not directly linked to imprinting, methylation can also be directed by CHROMOMETHYLASE 3 (CMT3) that has specificity for CNG, and members of the DOMAINS REARRANGED METHYLTRANSFERASE (DRM) family; DRM1 and DRM2, that are mainly responsible for asymmetric CHH methylation [50]. In order to address the involvement of DNA methylation in the regulation of maternal AGL36 expression, we performed SNP analyses of 3 DAP ovules from reciprocal crosses with mutants that have been shown to be involved in DNA methylation. In the SNP RT-PCR analysis of mutant pollen crossed to wild-type, paternal AGL36 expression is expected if the tested mutants are involved in AGL36 imprinting.

CMT3 DNA methylation has been reported to be guided to specific sites by KRYPTONITE (KYP) H3K9 methylation [51]. When mutant cmt3-7 and kyp-2 pollen were crossed to Col wild-type plants, no difference in AGL36 expression was observed (Figure 5A). In the reciprocal cross with cmt3-7 also no difference could be detected compared to wild-type expression (Figure S6).

DRM1 and DRM2 are mainly responsible for asymmetric DNA CHH methylation [50] and rely on small interfering RNAs, processed by ARGONAUTE4 (AGO4), for target template guidance [52]. In our assays, fertilization by pollen lacking DRM1;DRM2 and pollen lacking AGO4 had no effect on the
AGL36 expression pattern (Figure 5A). Likewise, AGL36 expression in the reciprocal cross was identical to wild-type (Figure S6). DECREASE IN DNA METHYLATION (DDM1) is involved in maintenance of DNA methylation [53]. In our SNP RT-PCR analyses where mutant ddm1-2 pollen was used to fertilize wild-type ovules, no paternal AGL36 expression could be detected in these crosses. (B) Upper panel: Hemizygous met1-4/+ in the Col ecotype reciprocally crossed with Ler wild-type plants show that the paternal allele of AGL36 is expressed when met1-4 is crossed as male. No paternal bands are observed when met1-4 is used as the maternal cross partner. Using pollen from first generation met1-4/+ homozygous plants in a Ler x met1-4/+ cross gives rise to prominent bands of the digested paternal Col expression product. In the reciprocal met1-4/+ x Ler cross no paternal expression can be detected. Lower panel: FWA control using the same tissue as above showing imprinted expression of FWA in dependence of MET1. Paternal FWA expression was observed when plants hemizygous and homozygous for met1-4 were crossed as male to Ler wild-type. doi:10.1371/journal.pgen.1001303.g005

AGL36 expression in the reciprocal cross was identified as wild-type (Figure S6).

DECREASE IN DNA METHYLATION (DDM1) is involved in maintenance of DNA methylation [53]. In our SNP RT-PCR analyses where mutant ddm1-2 pollen was used to fertilize wild-type ovules, paternal AGL36 expression was not activated (Figure 5A). In summary, CMT3, KYP, DRM1;DRM2, AGO4 and DDM1 appear not to be involved in the establishment nor maintenance of AGL36 imprinting (Figure 5A, Figure S6).

However, paternal AGL36 expression was detected when plants hemizygous for the met1-4 mutation were used as pollen donor in crosses with wild-type Ler (Figure 5B). In the reciprocal cross, using met1+/− as the maternal partner, no AGL36 expression from the paternal genome could be observed (Figure 5B). Furthermore, we performed crosses using pollen from homozygous met1-4/+ parents. When first generation homozygous met1 plants were used as pollen donor on wild-type plants, prominent AGL36 expression from the paternal Col genome could be observed (Figure 5B). This strongly suggests that the repression of the paternal copy of AGL36 is lifted due to the met1-4 mutation, and that MET1 is required for maintaining paternal inactivation of AGL36. In the reciprocal crosses, only expression from the maternal genome could be detected, both in the heterozygous and the homozygous met1-4 situation, further substantiating the requirement of MET1 in the male germline in order to maintain AGL36 imprinting (Figure 5B). Paternal AGL36 expression levels using homozygous met1-4 as the maternal cross partner appeared to be equal to maternal levels in the reciprocal crosses (Figure 5B). This opens for the interpretation that DNA methylation is not required for the regulation of maternal AGL36 expression.

Silencing of vegetative AGL36 expression involves MET1

In public expression databases, AGL36 is reported to be expressed in the seed and more precisely in the endosperm [54] (Figure S4). In order to monitor AGL36 expression in vegetative tissues and its dependence on DNA methylation, we performed a real-time PCR experiment on vegetative tissues from reciprocal Ler x Col crosses and homozygous met1-4 tissues. In biological replicates of progenies from both reciprocal crosses, weak AGL36 expression ranging from 1–6% of the seed expression level could
be detected in seedlings, leaves and flowers (Figure 6A). This showed that AGL36 was expressed throughout the plant life cycle, although at very low levels. In the same experiment, we monitored expression in met1-4 tissues. AGL36 expression levels were 50–90-fold higher in met1-4 leaves compared to seed expression levels (Figure 6A). In a direct comparison, expression levels were elevated 2000-fold in homozygous met1-4 leaves compared to wild-type Col x Ler flowers (Figure 6B). In flowers, the upregulation was more than 20-fold in met1-4 compared to wild-type Col x Ler flowers (Figure 6C). In conclusion, these data showed that silencing of AGL36 in vegetative tissues involves MET1, suggesting that the absence of maintenance DNA methylation elevates vegetative AGL36 expression beyond the maternal expression levels found in seeds.

Figure 6. AGL36 expression is controlled by MET1 and DME. (A) Real-time PCR of AGL36 expression levels in seeds, seedlings, leaves and flowers in F1 progenies of Ler x Col (gray bars), and Col x Ler (black bars) plants. AGL36 expression level in leaves and flowers of met1-4 selfed plants in the Col ecotype are shown in the two rightmost bars. All expression levels shown are relative to the AGL36 expression levels in 3 DAP seeds. (B–C) Expression levels of AGL36 in leaves (B) and flowers (C) of met1-4 plants (Col) relative to expression level in F1 progenies of Col x Ler plants. (A–C) represent the average relative expression values obtained from two independent biological parallels (BR) where each gave rise to four independent cDNA syntheses (TR). STDEV is derived from the two BRs. ACT11 is the reference gene used. (D) Parental dependence of AGL36 expression in seeds and flowers. SNP analyses of 3 DAP seeds (left) and F1 hybrid flowers (right) obtained from Col x Ler and Ler x Col crosses. The amplified SNP containing region of AGL36 was AlwN1I digested and analyzed in a Bioanalyzer. Seeds express AGL36 only from the maternal genome (left). Flowers express AGL36 biparentally (Note both Ler and Col bands (right)). The electropherograms represent one BR. Peaks are representing the bands shown in the graph. Asterisk; digested Col product, Arrowhead; undigested Ler product. (E) Real-time PCR expression levels of FWA, FIS2 and AGL36 in dme-6 x Col vs. wild-type seeds 3 and 6 DAP. Graphs represent the average relative expression from four independent BRs. Values for FIS2 are calculated based on 3 BRs as the value for the fourth BR was clearly out of range. Samples used in the first BR gave rise to two TRs. STDEV is derived from the independent BRs. ACT11 is the reference gene used.

doi:10.1371/journal.pgen.1001303.g006
AGL36 is biparentally expressed in vegetative tissues

In order to investigate the parental expression pattern of AGL36 in vegetative tissues, we performed SNP analyses of flowers from F1 hybrids of Ler and Col reciprocal crosses. In both reciprocal crosses, AGL36 appeared to be expressed equally from the parental Ler and Col genomes, indicating biparental expression in flowers (Figure 6D). This indicates that parental-specific expression, i.e. imprinting of AGL36, as expected, only takes place in the seed and that a low basal biparental expression is present throughout the plant life cycle. Interestingly, bi-allelic expression in flowers suggests that further silencing of AGL36 takes place in the male germline before uniparental expression in the seed (Figure 6D).

AGL36 is controlled by DEMETER

According to our data, the action of MET1 suppresses AGL36 expression throughout the vegetative phase and this suppression is maintained in the fertilization product through the male germline. AGL36 imprinting thus requires specific activation of the maternal allele. DNA demethylation by DME has previously been shown to mediate maternal-specific gene expression in the endosperm [11,18,19,24], and we therefore investigated AGL36 expression in dme-6 mutant plants. Since dme cannot be maintained in a homozygous state, we harvested siliques of dme-6/+ heterozygous plants pollinated with Col pollen at 3 and 6 DAP. We monitored the relative expression by means of real-time PCR using FWA and FIS2 as controls. At 3 DAP, both controls were downregulated by 69±6.09% and 53±8.90% respectively (Figure 6E), in line with a lack of functional DME in 50% of the seeds in heterozygous dme-6/+ plants. AGL36 was downregulated in a similar manner as FIS2 (41±25.0%), suggesting that DME is indeed involved in early activation of the maternal AGL36 allele.

Expression of maternal AGL36 is regulated by the PRC2 FIS-complex

We also tested the expression of FWA and FIS2 in 6 DAP samples and found that their downregulation were sustained as predicted (Figure 6E). However, to our surprise AGL36 expression in dme-6/+ seeds was elevated more than 50-fold (Figure 6E). This result was unexpected, and implicated a more intricate regulation of AGL36.

DME is required for the activation of MEA, the core histone H3K27 methyltransferase (HMTase) of the PRC2 FIS-complex [46,53,56]. To determine whether PRC2 FIS is involved in the regulation of AGL36, we analyzed the relative expression of AGL36 over time (1 to 12 DAP) in mea mutant seeds compared to wild-type (Figure 7A). While AGL36 expression in wild-type seeds was at its maximum at 4 DAP, we observed that AGL36 expression in mea seeds surpassed the maximum levels of wild-type at 4 DAP, and reached its highest levels at around 6 DAP. At this point, the AGL36 relative expression in mea mutant seeds was approximately 40-fold higher than wild-type expression at the same stage, and 7-fold higher than the maximum AGL36 level found in wild-type seeds at 4 DAP (Figure 7A). Our data thus indicate that the FIS-complex is indeed a repressor of AGL36 expression, and could also explain the elevated AGL36 expression level in 3 DAP dme-6/+ seeds (Figure 6E). In line with these findings, we found highly elevated AGL36 relative expression levels in mutant seeds from three different mutant alleles of mea (Figure 7C). Similar results were also obtained with mutants of other components of the FIS PRC2 complex (FIS2, FIE, and MS1, data not shown).

To investigate whether FIS activity was exerted on the maternal and/or paternal allele of AGL36, we performed SNP analyses on the RT-PCR product of AGL36 obtained from mea mutant plants (in Ler background) pollinated with Col wild-type pollen. We found that AGL36 is expressed only from its maternal allele in the mea background throughout the duration of our experiment (Figure 7B). In comparison to the expression pattern in wild-type (Figure 3B), strong ectopic maternal expression was also observed at 9 and 12 DAP stages. No paternal expression could be observed in these stages. By plotting the molarities of the maternal band detected by the Agilent Bioanalyzer, an expression profile for the maternal allele could be generated (Figure 7B, lower panel). This demonstrated that in the absence of MEA, AGL36 expression continues to increase after 4 DAP, and although the intensity decreases from 6 DAP, high level of AGL36 is maintained at 12 DAP. Hence, the FIS-complex represses the maternal allele of AGL36 after the 3 DAP stage.

To further substantiate that maternal AGL36 expression is regulated by the maternal action of MEA, we crossed mea mutant plants with pollen expressing the pAGL36::GUS reporter line. Here, no obvious activation of the paternal transgene could be observed at 3 DAP (Figure S7A). Surprisingly, at 6 DAP, corresponding to embryo heart stage, weak expression of the paternal copy in the embryo could be found (Figure S7A). In addition, we performed reciprocal crosses with the pAGL36::GUS reporter line in mutant mea background. When the transgene was contributed from the female side in mea background, a GUS signal was found in 3 DAP stages that increased drastically up to 6 DAP (Figure S7B). In the reciprocal cross however, no expression could be observed (Figure S7C).

The E(z) class of H3K27 histone methyltransferases (HMTases) in Arabidopsis consists of MEA, SWINGER (SWN) and CURLY LEAF (CLF) that participate in different PRC2 complexes. To test whether AGL36 repression is a specific function of FIS-MEA PRC2, we analyzed AGL36 expression in homozygous swn-4 and clf-2 seeds. For mutants of both HMTases values similar to the wild-type situation were found, and in conclusion AGL36 appear to be specifically regulated by FIS-MEA PRC2 (Figure 7C).

In summary, maternal AGL36 expression appears to be repressed specifically by the maternal action of FIS PRC2.

PRC2 acts on a subset of MET1/DME–regulated genes

For all genes known to be imprinted by PRC2, the FIS-complex is involved in the repression of the silenced allele [25-27,30,56]. Our data suggest that silencing of the paternal AGL36 allele requires ME1 whereas the maternal allele is activated by DME. Modulation of female AGL36 expression by PRC2 thus represents a novel mechanism in this type of gene expression system, and adds an additional level of parent-of-origin specific gene expression to the scheme. In order to investigate if this regulation applies to other genes imprinted by the dual action of MET1/ DME [11,18,19], we analyzed the relative expression levels of FWA, FIS2, AGL36 and MPC in a mea mutant. At 3 DAP expression levels were unchanged or slightly downregulated (0.40-0.99) for all genes tested (Figure 7D). However, while the expression of FWA and FIS2 remained stable at 6 DAP, AGL36 and MPC levels were elevated up to 80-fold (Figure 7D). Thus, genes imprinted by means of MET1/DME can be divided in two classes based on their dependence of FIS PRC2 for additional regulation of the expressed allele. Whereas one class appears not to be regulated by FIS PRC2, the other class depends on the action of the FIS-complex for developmental regulation of its expression.

Discussion

We have performed genome-wide microarray transcript profiling of seeds with only maternal endosperm as a screening
method to identify novel regulators of seed development. Previous experiments have shown that a paternal genomic contribution is essential in wild-type Arabidopsis plants for successful seed development. Thus, our working hypothesis was that in the absence of the paternal genome in the endosperm, key regulators of seed development are not present or not effectively transcribed. Using selection criteria that allowed for the identification of known paternally expressed genes, we extracted a set of downregulated genes that significantly overlapped with a set of endosperm expressed genes identified by Goldberg & Harada laboratories. The GO-Slim term Transcription factor activity was overrepresented in both down- and up-regulated gene sets, and a closer analysis revealed a striking overrepresentation of the Type-I MADS-box class among the downregulated transcription factors. With the selection criteria used, each detected gene could be a false positive at a probability of 0.35 at the highest, and thus a thorough examination of candidate genes, as performed in this report for AGL36, will be required.

MADS-box transcription factors play important roles in developmental control and signal transduction pathways in most if not all eukaryotes [57]. They are divided into two groups: the very well studied Type-II group (46 genes) including the MIKC class with important regulators such as AGAMOUS, and the Type-I group (61 genes), on which there is very limited information related to function [54,58,59]. Emerging data suggest that Type-I MADS-box genes differ from Type-II genes by being involved in female gametophyte and seed development [46,60–62]. In addition they were found to be only weakly expressed, and most members of this group contain no introns [63].

A comprehensive interaction study with members of the Arabidopsis MADS-box protein family by de Folter and colleagues indicated a complex network of interactions between these proteins (Figure S4). It revealed for instance that PHE1 interacts with AGL62, which in turn interacts with both AGL36 and AGL80. AGL62 itself is regulated by the FIS-complex, and functions as a suppressor of endosperm cellularization [46,59]. PHE1 and AGL36 on the other hand both interact with AGL28. In addition, mutant analysis has shown that AGL80 function is
required for the expression of DME in the central cell, and is therefore an upstream regulator of FIS PRC2 [60]. Moreover, AGL61 is required for central cell development, and there is evidence that a heterodimerization between AGL61 and AGL90 is necessary for AGL61 translocation to the nucleus [59,62]. PHE1 expression is upregulated in A. thaliana (At) x A. arenosa (Aa) incompatible hybrids due to loss of maternal PHE1 silencing, and introduction of phe1 could improve seed viability in semi-compatible 4xAx 2xAa crosses [64]. In A. thaliana, expression of a PHE1 antisense construct (MEAPromoter::PHE1) could partially restore the seed abortion phenotype in mea mutants [29]. Peculiarly, PHE1 loss-of-function has no phenotypic effect in A. thaliana [56]. However, given the high sequence similarity within the MY class of Type-I MADS-box factors, it is possible that MEAPromoter::PHE1 silenced not only PHE1 but also many other MY class genes. Taken together, it seems likely that additional Type-I MADS-box factors are upregulated in mea mutants and a collective downregulation by antisense PHE1 would thus restore some of the defects in mea.

In the cluster of Type-I AGL proteins identified in our screen we also found a large overlap with genes recently shown to be upregulated in incompatibility balanced At x Aa crosses compared to semi-compatible At x Aa maternal excess crosses (AGL35, AGL36, PHE1, PHE2, AGL62, AGL90) [65]. In accordance, mutations of both AGL62 and AGL90 partially restore seed lethality in incompatibility balanced At x Aa crosses, accompanied with selective transmission of the mutant alleles. This array of genes was also found to be upregulated in a PRC2 fsi2 mutant [65]. In addition, AGL36, AGL62, AGL90 and PHE1 were commonly upregulated in transcriptional profiles of At paternal excess crosses using tetraploid or unreduced jason (jas) pollen [66].

Together with these recent findings, the network of interactions with AGL62 (AGL36, PHE1, PHE2, AGL90) and PHE1 (AGL40, AGL62) and interactors of these proteins (AGL40, AGL45 and AGL90) strongly suggest that the here identified cluster of Type-I AGL proteins plays key roles in parent-of-origin dependent regulation of seed development. An in-depth study of different members of this group will therefore be of great value in understanding this process, and aid the identification of novel imprinted genes.

AGL36 imprinting requires MET1

Here, we report that AGL36 is a novel imprinted gene that is only expressed from its maternal allele in the endosperm. Silencing of the paternal allele requires the action of MET1, as paternal expression is restored in met1 mutants.

In public high-density DNA methylation maps prepared from wild-type seedlings (http://signal.salk.edu), both the AGL36 transcribed region and the 5’ and 3’ regulatory regions are decorated by CG methylation. In line with this, AGL36 was expressed at very low levels in vegetative tissues. Transcript levels however, were highly elevated in the absence of MET1, in accordance with the virtual absence of CG methylation in met1 mutants (http://signal.salk.edu) [67].

AGL36 is expressed from both parental alleles at low levels in vegetative tissues, which show that AGL36 imprinting occurs specifically in the endosperm. Other imprinted genes in Arabidopsis have been shown to have biallelic expression in the embryo and other vegetative tissues [11,34,68]. However, for most imprinted genes this issue is not clarified [3]. Since paternal AGL36 expression is absent in the seed, it suggests that further silencing of AGL36 takes place by entry into the male germline. Moreover, silencing in the female germline must be lifted to allow AGL36 expression in the seed. Alternatively, maintenance methylation and further silencing do not take place on the AGL36 gene in the female gametophyte. The majority of previously described imprinted genes are regulated by a dual switch of methylation and demethylation involving MET1 and DME [11,18-20,35]. Here we have shown that AGL36 expression is reduced in a dme mutant, indicating that DME has an activating function towards AGL36. In accordance with this, mutants of CMT3, KYP, AG04, DDML and DRML1/2 had no effect on paternal AGL36 expression suggesting that maintenance and repression by MET1 and activation by DME is sufficient for AGL36 imprinting.

In our SNP analyses, a weak paternal signal was observed only at the 2DAP stage. This was interpreted as an artifact since the signal was absent both before and after this stage. If this is a real paternal signal, it suggests an alternative hypothesis where silencing is achieved in the endosperm post fertilization. Further analyses are however required to support this.

In two recent studies, the genome-wide methylation profile of the seed was dissected by comparing cytosine methylation in wild-type embryos to wild-type and dme endosperm. This showed that endosperm development, and hence the activity of endosperm-specific genes, is marked by an extensive demethylation of the maternal genome, especially at specific transposon sequences [35,69]. According to the Zilberman Lab Genome Browser (http://zlab.pmb.berkeley.edu/browser/), such demethylation indeed takes place in the 5’ and 3’ regulatory regions of AGL36. Methylation patterns are regained in the dme mutant, supporting our data that AGL36 is maternally activated through the action of DME.

In an elegant approach by Gehring and colleagues, novel imprinted genes have recently been identified by the prediction of Differentially Methylated Regions (DMRs) between embryo and endosperm. In support of our findings, significant DMRs were also mapped to 5’ and 3’ regulatory regions of AGL36 [35].

Imprinting could be demonstrated in transgenic µAGL36:GUS seeds, thus indicating that the 1752 bp promoter fragment used is sufficient for parent-of-origin specific expression. The genomic environment of imprinted genes is highly correlated with transposable elements (TE), and imprinting has been postulated to be an evolutionary byproduct of silencing of invading transposons [23,69,70]. For instance, methylation of a SINE-related tandem repeat structure in the 5’-region correlates with FWA expression [32,71], and DMRs in MEA, PHE1, HDG2 and HDG9 map to TE [35]. In line with this, a variety of remnants of TE reside in both the 5’and 3’ regulatory regions of AGL36 (Figure 4A). The 1752 bp µAGL36:GUS promoter fragment harbors remnants of Helitrons and parts of an Arnoldy TE. An 800 bp DMR maps immediately (78 bp) upstream of the AGL36 transcriptional start site overlapping the Helitron TE[s] [35], Mary Gehring, personal communication). Clearly, the 1752 bp 5’ region is sufficient for basal AGL36 imprinting, and similar to the abovementioned examples, AGL36 DMRs map to TE. Further investigations will be needed to elucidate the role and the mechanisms of additional 5’and 3’DMRs as well as the involvement of small RNAs in AGL36 imprinting [72].

The PRC2 FIS-complex regulates maternal expression of AGL36

Distinct from the expression pattern of AGL36 that subsides at the time of cellularization in wild-type seeds, AGL36 maternal expression in mea mutant seeds was highly elevated and sustained throughout seed development. Recently, Walia et al. also reported AGL36 upregulation obtained in five days old seeds from selfed fsi2+/- plants [65]. Our results show that FIS-complex mediated repression acts exclusively on the expression of the maternal allele
of AGL36. The paternal allele was efficiently silenced throughout endosperm development.

Surprisingly, weak paternal pAGL36::GUS expression could be observed in 6 DAP early heart stage embryos when the mother was homozygous for mea. MEA1 has been shown to have biallelic expression in the embryo [28], and thus the observed paternal expression in hemizygous zygotes is not caused by the lack of functional MEA. This could hint to dosage-dependent regulation of paternal AGL36 expression by MEA, directly or indirectly, but in lack of further experiments this remains speculation.

Different PRC2 complexes can regulate common genes [30]. However, in mutants of CLF and SWN, the paralogues of MEA, no significant effect on AGL36 expression levels was found, indicating that AGL36 regulation is specific to PRC2FIS. H3K27 trimethylation mediates PRC2s repressive function, and in a whole-genome assay for H3K27 methylation more than 4400 target genes were detected [73] (Daniel Bouyer, personal communication). AGL36 was however not part of this set of genes. Since this material was obtained from seedlings and may not reflect the situation in the seed, it is not known whether AGL36 is a direct target of H3K27 trimethylation.

AGL36 identifies a dual regulation mechanism by DME and the FIS PRC2-complex

Repression of the maternal AGL36 allele identifies a novel means of dual epigenetic regulation of imprinted genes. In this scenario, the expressed maternal AGL36 allele is antagonistically activated by DME and repressed by PRC2FIS. To our knowledge, this is the first report of an imprinted gene where the expressed allele is concurrently regulated by a repressive epigenetic mark.

We asked whether this type of regulation was specific for AGL36 by investigating the fis mutant for expression of three other imprinted genes that are activated by DME. We found that these genes fall into two distinct groups; FWA and FIS2 which were largely unaffected by the lack of FIS, and MPC along with AGL36 which showed strong upregulation. This suggests that additional PRC2 regulation of DME-activated alleles defines a common mechanism that applies to a subset of imprinted genes.

In Arabidopsis, three imprinted genes, MEA, PHE1 and Adf15 are known to have their silenced allele repressed by PRC2FIS, and two of these genes, MEA and PHE1 are additionally regulated by DNA methylation [55]. In these cases however, the repressed allele is silenced by PRC2 whereas the active allele is regulated by DNA methylation [74]. Here, we show that AGL36 defines a novel type of regulation where the same allele is activated by DME and repressed by PRC2FIS in a sequential fashion. This suggests that maternal AGL36 expression after DME activation needs to be dampened and developmentally regulated by PRC2FIS, in accordance with the strong AGL36 expression observed in hypomethylated met1-1 plants. Interestingly, DME is required to activate both PRC2MEA and AGL36, and is thus a key player in developmental tuning of parent-of-origin specific AGL36 expression.

The role of AGL36 in seed development

AGL36 was identified in our transcript profiling as a downregulated gene when the paternal contribution to the endosperm was absent. A simple hypothesis to account for this regulation would be that AGL36 is under the control of one or more paternally expressed factor(s) that activate the maternal allele of AGL36. The identity of such factor(s) remain unknown, and was not approached in this work, but a simple prediction from this hypothesis is that AGL36 would be upregulated in paternal excess interlopidy crosses. In a recent report, AGL36 is indeed upregulated in such crosses, as well as in crosses with unreduced diploid jas pollen [66]. Such parental cross-talk is however likely to involve complex genetic and epigenetic regulatory mechanisms, and the mechanism that cause the observed transcriptional response of AGL36 and other previously described imprinted genes in cdka;1P seeds remains to be clarified.

In our study, we have shown that AGL36 is only maternally expressed. Our current model suggests that the paternal allele is silenced by the action of MET1 and the maternal allele activated by DME (Figure 8). In addition, we have also shown that PRC2FIS regulates the expression of the maternal AGL36 allele at the transition between proliferation and cellularization (Figure 8).

Although AGL36 is identified as a novel target of the imprinting machinery in Arabidopsis, we have limited knowledge about its function during plant and seed development. Since expression of AGL36 and its interacting partners coincide with the transition of endosperm from proliferation to differentiation, we speculate that it plays an important role in this process. This is in agreement with recent findings [65], showing that suppression of an AGL cluster including AGL36 is critical for successful transition of endosperm from syncytial to cellularized stage.

In this work we have identified a novel imprinted gene that is controlled by a novel type of dual epigenetic regulation in the seed. This underscores the importance of further investigations to identify imprinted genes in order to unravel the complex network of epigenetic regulation of parent-of-origin effects in seed development.

Materials and Methods

Plant strains and growth conditions

All plant lines used in these experiments were obtained from the Nottingham Arabidopsis Stock Centre (NASC) unless otherwise stated. The mutant lines cdka;1-1 (SALK_106809; [40,41]), ddm1-2 (a kind gift from E. Richards; [53]), dme-6 (GK-252E03-014577; Figure S9), mea-8 (SAIL_55_C04; [75]), mea-9 (SAIL_724_E07; Figure S9), met1-4 (SAIL_809_E03; [76] and swn-1 (SALK_109121; [77]) were in the Col accession. The mutant lines agld-1 (N6364; [78]), df-2 (N290; [79]), cmt3-7 (N6365; [80], fmi1 (a kind gift from A. Chaudhury; [14]) and bsp-2 (N6367; [51]) were in the Ler accession. The drm1 drm2 (N6366; [81]) line was in the ws-2 accession. Mutants used in this study were genotyped using gene-specific and T-DNA specific primers as described in Table S2. The ddm1-2 mutant line was genotyped by an allele-specific PCR test using dCAPS primers DDM1f and ddm1-2Rsa, as described by [68], allowing digestion of the PCR fragment of the ddm1-2 allele with Rsal restriction endonuclease, generating a ~130 bp band.

We obtained the agld-1 allele from the Koncz collection [48]. Allele-specific PCR, using the primers HOOK1 (left border T-DNA primer) and AGL36 AS2 KONCZ (genomic AGL36 primer), was carried out to verify the T-DNA insertion, followed by sequencing analysis of the PCR product using the HOOK1 primer. The left border of the insertion was verified to be 16 bp upstream of the ATG start codon of AGL36. In addition, there is an 11 bp long DNA filler located between the genomic sequence and the T-DNA sequence.

Arabidopsis seeds were surface-sterilized using EtOH, bleach and Tween20 prior to plating out on MS-2 plates [82] supplemented with 2% Sucrose, containing the correct selection when necessary. Seeds on the MS-2 plates were stratified at 4°C O.N. before they were incubated for 14 days at 18°C to germinate. The seedlings were then put on soil and grown in long day conditions (16 hr light) at 18°C.

Seed isolation, RNA extraction, and cDNA synthesis

To increase tissue specificity, siliques were cut open and seeds were isolated directly in tubes containing pre-chilled ceramic
beads (Roche MagNA Lyser Green Beads). Isolated tissues were stored at −80°C. Homogenization was performed by the addition of Lysis buffer containing β-ME (Sigma Spectrum Plant Total RNA Kit) directly to the samples, followed by 3 × 15 second intervals of homogenization using a MagNA Lyser Instrument (Roche). To prevent RNA degradation, samples were chilled on ice two minutes between each homogenization interval. After the last homogenization step, the samples were centrifuged at 4°C for one minute prior to the transfer of the lysate to a new 1.5 ml tube. RNA extraction was performed according to the Sigma Plant Total RNA Kit protocol, except that all centrifugation steps were done at 4°C and not at room temperature as indicated in the protocol. RNA was eluted in 50 μl volume. cDNA was synthesized by first preparing the RNA for real-time PCR by treatment with DNase I (Sigma) followed by Reverse Transcription using Oligo(dT) and SuperScript III Reverse Transcriptase (Invitrogen) according to the manufacturer’s protocols. The synthesized cDNA was purified utilizing QIAquick PCR Purification kit (Qiagen) and eluted in 30 μl volume prior to measurement of cDNA concentration using a NanoDrop 1000 Spectrophotometer.

**Microarray analysis**

Plants were grown and seeds isolated as described above. Total RNA was isolated as described above. For microarray analysis, three biological replicates were generated, each consisting of approximately 35 hand-pollinated siliques from ten different plants. The microarray experiment was conducted by the NARC Microarray Service in Trondheim. Microarray slides were printed by the Norwegian Microarray Consortium (Trondheim, Norway). A custom made *Arabidopsis* chip with 32567 unique 70-mer oligo probes was used in the experiments. Total RNA (15 mg) and Super-Script III reverse transcriptase (Invitrogen) were used in a reverse transcription reaction. A 3DNA Array 350 kit with Cy3- and Cy5-labelled dendrimers (Genisphere Inc.) was used for labeling. Hybridizations were performed in a Slide Booster Hybridization Station (Advalytix), and the slides were washed according to the manufacturers’ descriptions (Genisphere and Advalytix). The slides were scanned at 10 mm resolution on a G2505B Agilent DNA microarray scanner (Agilent Technologies). The resulting image files were processed using GenePix 5.1 software (Axon Instruments). Spots identified as not found or manually flagged out as bad were filtered out. Spots with more than 50% saturated pixels were also excluded. The data sets were log-transformed and normalized using the print-tip Loess approach [83]. Within-array replicated measurements for the same gene were merged by taking the average between the replicates. The data were then scaled so that all array data sets had the same median absolute deviation. The differentially expressed genes were identified using the Limma software package [84]. The resulting set of p-values were used to compute the q-values as described [85].

The microarray data generated in this publication have been deposited in NCBI’s Gene Expression Omnibus and are accessible through GEO Series accession number GSE24809 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE24809).

**Bioinformatics analyses**

We defined the following sub-sets for our microarray data (see Table S3): All expressed = all genes having a present call (17223

---

**Figure 8. A model for the imprinted expression of AGL36.** Paternal silencing of AGL36 is maintained by the action of MET1 in the male germline prior to fertilization (1). The maternal copy of AGL36 is repressed after fertilization, either directly or indirectly, by the maternally expressed FIS-complex (2) (only MEA is indicated in the figure). DME is expressed only in the central cell and activates the maternal allele of MEA and AGL36 (3 & 4). The autoregulatory repression of maternal MEA upon paternal MEA is indicated (5). Solid lines indicate direct regulation. Dashed lines indicate possible indirect regulation. doi:10.1371/journal.pgen.1001303.g008
genes); Down 0.8 = in Ler x cdka;1 downregulated genes with $q \leq 0.35$ and arithmetic ratio $|ar| \leq 0.0$ (602 genes); Up 1.5 = in Ler x cdka;1 upregulated genes with $q \leq 0.35$ and $ar \leq 1.5$ (323 genes); Up 1.2 = in Ler x cdka;1 upregulated genes with $q \leq 0.35$ and $ar \leq 1.2$ (1030 genes). The $q$-value is the false discovery rate (FDR) of the $p$-value, and was adjusted with Storey’s $q$-procedure [83].

The threshold for analysis was set to $q \leq 0.35$ since this value detected paternally expressed genes at an arithmetic ratio $(|ar|) \leq 0.8$. A functional classification was done at http://www.arabidopsis.org/tools/bulk/go/ using the GO-Slim Molecular Function classification system. For the detailed transcription factor analysis we used the Transcription Factor (TF) classification from the Arabidopsis transcription factor database (AtTFDB) hosted on the Arabidopsis Gene Regulatory Information Server (AGRIS, http://arabidopsis.med.oioh-state.edu/AtTFDB). The MADS TFs were sub-grouped as in de Folter et al [45]. We compared our microarray data with seed expression data generated by the Goldberg & Harada laboratories, available at http://seedgenenetwork.net/analyze2 project=Arabidopsis.

For data comparison a reference set of genes was used that contained all genes covered by the Operon chip used in our study (Arabidopsis thaliana 34K NARC serie 8; GEO Platform GPL11051GPL) and the Affimetrix chip used by Goldberg & Harada (Ath1, GEO Platform GPL198). For the Ath1 chip we used the annotation provided by Goldberg & Harada available at http://seedgenenetwork.net/media/Arab_Final_Annotations_09-07-07_completed.txt. For the operon chip we used the current TAIR 9.0 annotation. From these annotations all AGIs for nuclear genes were extracted and the overlap was calculated. This reference set contained 22130 genes.

We used the reference set overlap of the following Goldberg/Harada datasets for comparison: GH seed = call all present and experiment in Arabidopsis ATH1 Array/Arabidopsis/Globular Stage/Seed; GH seed coat = call all present and experiment in Arabidopsis ATH1 Array/Arabidopsis/Globular Stage/Chalazal Seed Coat or Arabidopsis ATH1 Array/Arabidopsis/Globular Stage/General Seed Coat; GH endosperm = call all present and experiment in Arabidopsis ATH1 Array/Arabidopsis/Globular Stage/Chalazal Endosperm or Arabidopsis ATH1 Array/Arabidopsis/Globular Stage/Microzygote Endosperm or Arabidopsis ATH1 Array/Arabidopsis/Globular Stage/Peripheral Endosperm; GH embryo = call all present and experiment in Arabidopsis ATH1 Array/Arabidopsis/Globular Stage/Embryo Proper.

Venn diagrams were generated using the VENN diagram generator designed by Tim Hulsen at http://www.cmbi.ru.nl/cdd/biovenn/ [86]. The test for statistical significance of the exclusive and specific AGL36 amplification.

ACTIN11 (ACT11), a housekeeping gene that is strongly expressed in the developing ovules [89], was shown in a preliminary analysis not to be affected by our experimental conditions (data not shown), and was therefore selected as a reference gene. GLYCERALDEHYDE-3-P DEHYDROGENASE A-SUBUNIT (GAP4) was used as an additional reference gene. The oligo sequences, their amplicons and appropriate UPL probes are shown in Table S2.

Real-time PCR of all samples and reference controls were performed in two independent biological replicates and repeated at least two times (technical replicas) unless otherwise stated. The PCR efficiency was determined independently for all replicates (biological and technical) by series of dilutions (100 ng, 50 ng, 20 ng, 5 ng template/rxn) for each experiment. This allowed us to obtain the efficiency for each single reaction. Calculations of relative expression ratios were performed according to a model described by Pfaffl [90] with minor exceptions. Since we had efficiency for all reactions (four values for each calculation corresponding to $E_{\text{target-sample}}$, $E_{\text{target-standard}}$, $E_{\text{reference-sample}}$ and $E_{\text{reference-standard}}$), we calculated the average $E_{\text{target}}$ and $E_{\text{reference}}$ values from the standards and the samples, ending up with two $E$-values that we could use in the formula described by Pfaffl.

Single Nucleotide Polymorphism analysis

RNA was isolated and cDNA synthesized and purified as described above. Polymorphisms between various ecotypes were identified using TAIR Genome Browser (www.arabidopsis.org) and/or the Arabidopsis SNP Sequence Viewer tool provided by the Salk Institute Genomic Analysis Laboratory (http://natural.salk.edu/cgi-bin/snp.cgi). A selected region spanning the SNP of interest was amplified by PCR using TaKaRa Ex Taq DNA polymerase applying 100 ng template per reaction, and the following PCR parameters in a 50 μl reaction: 94°C-3 min, 35x (94°C-1 min, 58°C-30 sec, 72°C-1 min/kb), 72°C-5 min, 4°C-∞. Parental-specific expression based on SNP was determined by setting up an appropriate restriction digest. For AGL36 SNP analysis, 20 μl of the SNP PCR reaction was digested with 15 U of

**Plasmid construction**

To generate the pAGL36:GUS construct we utilized the Gateway cloning technology (Gateway; Invitrogen). The promoter region (+1740–12) spanning the ATG start codon was amplified using the attB sequence containing primers attB1-pAGL36-AS7 and attB2-pAGL36-S4 (Table S2), and cloned into the pMD162 GUS-vector [67]. The resulting construct, after checking the DNA sequence, was introduced to Col ecotype by Agrobacterium tumefaciens mediated transformation using the floral-dip method [88].

**β-Glucuronidase expression analysis and histology**

Histochemical assays were performed after a modified protocol from Grini et al. (2002) by incubating the tissues in staining buffer (2 mM X-Glu; 50 mM NaPO₄, pH 7.2; 2 mM K₃Fe(CN)₆ × 3H₂O; 2 mM K₄Fe(CN)₆; 0.1% Triton) overnight at 37°C before the reaction was terminated using 50% EtOH. The tissues were cleared and mounted on slides according to Grini et al. (2002), and inspected using an Axioplan 2 Carl Zeiss Microscope. Images were acquired with an AxioCam HRC Carl Zeiss camera and processed with AxioVision V 4.5.0.0 software.
verifying that the difference could easily be detected using a DNA-1000-Lamplicon (373 bp versus 399 bp respectively). This amplifying AG34 70% and 91% sequence similarity between the primers and the AG34 unique restriction sites (these two amplicons, we took advantage of the presence of two ASP2-subcloning primers (Table S2)), and comparing the size differences obtained between the genomic PCR and cDNA PCR. Due to high sequence similarity, we suspected to amplify both AG90 and AG90. The presence of the intron in AG90 was confirmed by amplifying the intron-flanking region (AG90-SP1-subcloning and AG90-ASP2-subcloning primers (Table S2)), and comparing the size differences obtained between the genomic PCR and cDNA PCR. To rule out that the primers used for AG36 SNP PCR (AG36-SP7-SNP and AG36-ASPS-SNP) (Table S2) would amplify the highly similar AG90, we oriented the AG36-SP7-SNP primer such that it was located in a region that was annotated as intron in AG90 but not in AG36 (Figure S8). First, the presence of the intron in AG90 was confirmed by amplifying the intron-flanking region (AG90-SP1-subcloning and AG90-ASP2-subcloning primers (Table S2)), and comparing the size differences obtained between the genomic PCR and cDNA PCR.

Due to high sequence similarity, we suspected to amplify both AG36 and AG90 in these PCR reactions. To distinguish between these two amplicons, we took advantage of the presence of two unique restriction sites (MAl and BgII) in the amplified region of AG90 that are absent in AG90.

Sequence comparison between the abovementioned AG36- SNP primers and AG34 showed that there was approximately 70% and 91% sequence similarity between the primers and the AG34 gene. However, if these primers were functional in amplifying AG34, they would result in a smaller amplicon than AG36 amplicon (373 bp versus 399 bp respectively). This difference could easily be detected using a DNA-1000-LabOnChip. Our SNP data only showed the expected 399 bp band, verifying that AG34 was not amplified using the above primers. The paternally imprinted FWA gene was used as a positive control by utilizing primers FWA-RTf and FWA-dNheI (Table S2) for PCR amplification followed by NheI restriction digest.

Supporting Information

Figure S1 Genomic dissection of parental effects using cdka;1 as a tool. (A–D) Basic setup and hypothetical outcome of the Ler × Col vs. Ler × cdka;1 microarray screen. (A) Transcription profiles from Ler × Col seeds were compared to Ler × cdka;1. In the endosperm of Ler × cdka;1 no paternal genome is present. (B) Paternally expressed target genes will be absent in Ler × cdka;1 seeds and thus downregulated. (C) Target genes that are activated by a paternally expressed gene X will be silent without the activator present, and thus downregulated. (D) If repressed by a paternally expressed gene X, the target gene will be upregulated. (E) Previously identified imprinted genes display reduced expression in cdka;1 fertilized seeds with no paternal contribution to the endosperm. Paternally expressed genes are shown in the upper panel. Maternally expressed genes are shown in the lower panel. The q-value (1) is defined as the false discovery rate (FDR) of the p-value, and was adjusted with Storey’s q-procedure [85]. The seed expression profile (2), obtained from Genevestigator, is showing the level of gene expression in the embryo, the endosperm (micropylar, peripheral, and chalazal), the seed coat and the suspensor. The expression levels are shown in a range from low/none (white) to high (dark blue). The probe for the PHEI and PHE2 expression profile is not able to distinguish between these genes and is represented with **. Found at: doi:10.1371/journal.pgen.1001303.s001 (1.72 MB EPS)

Figure S2 Overlap between different seed compartment profiles and cdka;1 microarray expression data. (A) Venn diagrams representing overlap of genes expressed at globular stage in endosperm, seed coat or embryo of Arabidopsis Ws-0 plants (grey/white) and genes expressed in 3 DAP seeds from Ler plants pollinated with Col cdka;1 pollen (green). Genes significantly deregulated with respect to seeds from Ler plants pollinated with Col pollen are indicated in blue (downregulation, ar = ±0.8), red (upregulation, ar = ±1.5) and orange (upregulation ar = ±1.2). Gene numbers refer to the reference set of genes (see material and methods). GH endosperm represents expression in chalazal or micropylar or peripheral endosperm, GH seed coat represents expression in chalazal or general seed coat (www.seedgenenet work.net). GH: Goldberg & Harada laboratories. (B) Two groups of genes are compared and found to have x genes in common. A representation factor (rf) and the probability (p) of finding an overlap of x genes are calculated at http://elegans.uky.edu/MA/ progs/overlap_stats.html. The representation factor is the number of overlapping genes divided by the expected number of overlapping genes drawn from two independent groups. A representation factor >1 indicates more overlap than expected of two independent groups, a representation factor <1 indicates less overlap than expected. The overlap of the Ler x cdka;1 seed dataset (green) with each of the GH datasets (grey/white) is always r f = 1.3 with a p-value of <1.0 e -99 (highly significant, below calculation limit of the software). For all other comparisons see table: ar: arithmetic ratio of expression values (Ler x cdka;1/Ler x Col).

Found at: doi:10.1371/journal.pgen.1001303.s002 (0.39 MB EPS)

Figure S3 Differences in deregulation in Ler x cdka;1 seeds among the different transcription factor families. On the X-axis the different transcription factor (TF) families are listed, the Y-axis displays the amount of TFs as absolute numbers (A) or as percentage with respect to all seed-expressed TFs in each class (B). Green bars represent all TFs having a present call in our microarray experiment, blue bars indicate TFs that are downregulated (cut-off ≤0.8) and red and orange bars indicate the TFs that are upregulated in Ler x cdka;1 seeds at 3 DAP (cut-off ≥1.5 or ≥1.2 respectively). The MADS MTFs are the most strongly downregulated class in absolute numbers as well as in percentage. Found at: doi:10.1371/journal.pgen.1001303.s003 (0.44 MB EPS)

Figure S4 Downregulated AGAMOUS-LIKE (AGL) transcription factors. Upper panel: MADS-box transcription factors identified in genome-wide transcript profiling of endosperm without paternal contribution. The table shows all identified MADS-box genes with a cut-off at >0.8-fold downregulation. The logarithmic and arithmetic relative expression ratios are indicated. The Mz and the Mz Type-I subclass of MADS-box factors are indicated in blue and green respectively. The Type-II MADS-box factor is indicated in yellow. The p-value (1) (a score between 0 and 1) is the likelihood of an event. The q-value (2) is defined as the false discovery rate (FDR) of the p-value, and was adjusted with Storey’s q-procedure [85]. The seed expression profile (3), obtained from Genevestigator, is showing the level of gene expression in the embryo, the endosperm (micropylar, peripheral, and chalazal), the seed coat and the suspensor. The expression levels are shown in a range from low/none (white) to high (dark blue). The probe for the expression profile of AG36 and AG90 is...
not able to distinguish between these genes and is represented with *. The probe for PHE1 and PHE2 expression profile is also not able to distinguish between these genes and is represented with **. Lower panel: Map of interactions between selected AGL proteins, modified from de Folter, 2005 [46], and the Bio-Array Resource (BAR) Arabidopsis Interaction Viewer (http://bar.utoronto.ca/). Blue ring color indicates the M2 subclass while green ring color indicates M7 subclass. Genes identified in our microarray (pink fill) and their interacting partners (no fill) are visualized.

Found at: doi:10.1371/journal.pgen.1001303.s004 (0.41 MB EPS)

**Figure S5** Analysis of the agl36-1 T-DNA insertion line. (A) T-DNA insertion map of agl36-1. The KONCZ T-DNA is inserted 16 bp upstream of ATG. There is an 11 bp DNA filler positioned between the T-DNA insert and the genomic AGL36 sequence (asterisk). The position of the left border (LB) of the insertion was verified by PCR using primers HOOK1 and AGL36-AS2-KONCZ. (B) Real-time PCR of AGL36 expression levels in 3 DAP seeds of agl36-1 hemizygous and homozygous plants, representative to wild-type Col seeds. The graph represents the average relative expression values from two independent biological parallels where each gave rise to two independent technical replicas. STDEVs are derived from the independent biological parallels. The AGL36 transcript levels were normalized to ACT111 levels. (C) Phenotypic analysis of wild-type Col (upper panel) and agl36-1 T-DNA (lower panel) seeds. Samples are taken at 2, 4 and 6 DAP. There is no obvious mutant phenotype observed in the seed and the developing embryo. (D) Histochemical detection of GUS activity in agl36-1 T-DNA seeds expressing a maternal pFIS2::GUS construct. The division pattern and the nuclear migration is similar to that of wild-type seeds (not shown).

Found at: doi:10.1371/journal.pgen.1001303.s005 (1.65 MB TIF)

**Figure S6** The effect of maternal DNA methylation on AGL36 expression. SNP analyses of 3 DAP seeds from crosses with DNA methylation mutants. The cross is reciprocal of the results presented in Figure 5A. The amplified SNP containing a region of AGL36 cDNA was digested with AlwNI and analyzed in a Bioanalyzer. Homozygous cml3-7, hpy-2, and aga4-1 mutants in the Ler ecotype were pollinated with Col plants, while homozygous dmi1/dmi2 mutants in the Ws-2 ecotype were pollinated with Ler plants. No paternal AGL36 expression could be detected in these crosses.

Found at: doi:10.1371/journal.pgen.1001303.s006 (0.49 MB EPS)

**Figure S7** pAGL36::GUS in reciprocal crosses and mea background. (A) pAGL36::GUS used as the paternal pollen donor on mea−/−. Note: Weak expression in the 6 DAP embryo. (B) mea−/−; pAGL36::GUS plants crossed with wild-type pollen. An increased pAGL36::GUS expression is detected in the mea−/− background. (C) Wild-type plants are crossed with mea−/−; pAGL36::GUS pollen. No expression is detected. (D) Wild-type plants from Col ecotype was used to fertilize plants expressing pAGL36::GUS. Maternal expression is detected. (E) Plants expressing pAGL36::GUS were used as paternal partners in crosses with wild-type Col plants. No paternal expression is detected. Samples are taken at 3 DAP (left panel) and 6 DAP (right panel).

Found at: doi:10.1371/journal.pgen.1001303.s007 (6.64 MB TIF)

**Figure S8** AGL36 alignment with AGL34 and AGL90. Alignment of the transcribed and 3‘UTR regions of AGL34, AGL36 and AGL90. The ATG start and TAA stop codons are marked with red boxes. Sequence similarity between all three genes is shown in black (and shown with capital letters below the alignment), whereas similarity between two genes is indicated with gray (and shown in small letters below the alignment). Gaps are shown with dashed lines. The forward and reverse oligonucleotide sequences for AGL36 real-time PCR (AGL36-160-LP and AGL36-160-RP) are shown in red letters and indicated with (A) and red lines above the sequence. The corresponding UPL Probe #160 sequence is shown with orange letters and indicated with (B) and an orange line above the sequence. AGL30 intron is indicated with (C) and shown with black text and blue background color. The forward and reverse oligonucleotide sequences for AGL36 SNP analysis (AGL36-SP7-SNP and AGL36-ASP6-SNP) are shown in green letters and indicated with letter (D), and green lines above the sequence. Note: The reverse AGL36-ASP6-SNP primer is located in the 3‘UTR. The forward and reverse primers for AGL90 amplification flanking the introns are indicated with black text and yellow background (E).

Found at: doi:10.1371/journal.pgen.1001303.s008 (8.61 MB TIF)

**Figure S9** Description of mea-8, mea-9 and dme-6 T-DNA insertion lines. (A) T-DNA insertion map of mea-8 (SAIL_55_C04, [75]) and mea-9 (SAIL_724_E07) mutant lines. The T-DNA is inserted in the 4th and 6th intron respectively. A phenotypic characterization was performed on mea-9 mutant seeds (right) and compared to wild-type seeds (left) at 12 DAP. The mea-9 line displays the same arrested phenotype as previously described, and arrests at late heart stage of embryo development. The frequency of aborted seeds (see table) is similar to previously described mea alleles. (B) T-DNA insertion map of dme-6 (GK-252E03-014577) mutant line. The GABI-KAT T-DNA is inserted in the 2nd intron. A phenotypic characterization was performed on dme-6 mutant seeds (lower panel) and compared to wild-type seeds (upper) at 9 and 12 DAP. The dme-6 mutant displays a characteristic maternal gametophytic abortion phenotype with enlarged endosperm and an aborting embryo at late heart stage. The abortion rate of the mutant ovaules is approximately 50%. (C) Real-time PCR expression analysis in mutant lines. The expression level of PHE1, which is repressed by MEA4, is increased in both mea-6 and mea-9 mutant lines. The expression level of PHE1 in dme-6 increase at 6 DAP since DME activate MEA1 expression.

Found at: doi:10.1371/journal.pgen.1001303.s009 (7.57 MB EPS)

**Table S1** Segregation and reciprocal crosses of the agl36-1 mutant line. 1Number of hygromycin resistant and sensitive plants in self-fertilized and reciprocally crossed plants. 2Percent hygromycin resistant plants in self-fertilized and reciprocally crossed plants. Standard deviation is indicated in this field. 3Mean percent value for resistant plants in self-fertilized and reciprocally crossed plants. Median percent value for resistant plants in self-fertilized and reciprocally crossed plants. 4Percent hygromycin resistant plants in self-fertilized and reciprocally crossed plants. Standard deviation is indicated in this field. 5Mean percent value for resistant plants in self-fertilized and reciprocally crossed plants. Median percent value for resistant plants in self-fertilized and reciprocally crossed plants. 6Two-tail significance test: H0: 75% segregation in hemizygous self-fertilized plants or 50% segregation in hemizygous outcrossed plants. A P value of 0.05 with 1 degree of freedom was used, meaning that with χ² > 3.84, the hypothesis holds with 95% accuracy, and is not rejected.

Found at: doi:10.1371/journal.pgen.1001303.s010 (0.44 MB EPS)

**Table S2** Primer sequences and Real-time PCR probes. Oligo sequences are all given in 5’–3’ direction. See comments in the table for details.

Found at: doi:10.1371/journal.pgen.1001303.s011 (0.08 MB PDF)

**Table S3** Gene lists—evaluation of microarray expression data. The following gene sets extracted from our microarray experiments are shown: All expressed = all genes having a present call (17223 genes); Down 0.8 = in Ler x cdna1 downregulated genes with q=0.35 and arithmetic ratio (ar) ≤ 0.8 (602 genes); Up 1.5 = in Ler x cdna1 upregulated genes with q=0.35 and ar ≥ 1.5 (323 genes); Up 1.2 = in Ler x cdna1 upregulated genes with q ≤ 0.35
and ar $\geq$ 1.2 (1030 genes). The raw data for the microarray experiments can be downloaded from GEO (GEO Series GSE24809).

Found at: doi:10.1371/journal.pgen.1001303.s012 (3.18 MB XLS).

Acknowledgments

We would like to thank the staff at the NTNU Microarray Service facility and especially Torfinn Sparstad and Tommy Jorstad for conducting the microarray experiments and for their help with statistical analysis. We thank Roy Falleh, Charles Albin-Amiot, Solveig Hauge Engbretsen, and Ida Myrher Sto for technical assistance. We thank Abed Chaudhury for homozygous fs1 seeds and Eric Richards for ddm1-2 seeds.

Author Contributions

Conceived and designed the experiments: PEG. Performed the experiments: RS EDA KBMG ALU. Analyzed the data: RS EDA KBMG MH AU PW AS PEG. Contributed reagents/materials/analysis tools: CK RBA AS. Wrote the paper: RS AS PEG.

References

1. Nowack MK, Ungru A, Bjerkar KN, Grini PE, Schnittger A (2010) Reproductive cross-talk: seed development in flowering plants. Biochem Soc Trans 38: 604–612.
2. Feil R, Berger F (2007) Convergent evolution of genomic imprinting in plants and mammals. Trends Genet 23: 192–199.
3. Jullien PE, Berger F (2009) Gamete-specific epigenetic mechanisms shape genomic imprinting. Curr Opin Plant Biol 12: 657–642.
4. Liu BY (1986) Ploidy Barrier to Endosperm Development in Maize. Genetics 107: 103–115.
5. Scott RJ, Spielman M, Bailey J, Dickinson HG (1998) Parent-of-origin effects on seed development in Arabidopsis thaliana. Development 125: 3329–3341.
6. Horsch D, Westoby M (1991) Genomic Imprinting in Endosperm: Effect on Seed Development in Crosses between Species, and between Different Ploidies of the Same Species, and Its Implications for the Evolution of Apomixis. Philosophical Transactions: Biological Sciences 333: 1–13.
7. DeChiara TM, Efrat-Stadler A, Robertson EJ (1993) The role of a growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. Nature 345: 79–80.
8. Felson AJ, Losi E, Efrat-Stadler A, Robertson EJ (1995) Rescue of the T- associated maternal effect in mice carrying null mutations in Igf2 and Igf2r, two reciprocally imprinted genes. Development 118: 751–766.
9. Lau MM, Stewart CE, Liu Z, Bhatt H, Rotwein P, et al. (1994) Loss of the GPC endosperm by DNA methylation. Science 265: 81–83.
10. Leighton PA, Ingram RS, Eggenschwiler J, Efstratiadis A, Tilghman SM (1995) Repression of paternal gene expression in embryos and male gametes required for seed development. Embo J 22: 4804–4814.
11. Kohler C, Hennig L, Sippale C, Pien S, Grau wound P, et al. (2003) The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERESI. Genes Dev 17: 1540–1553.
12. Makarevich G, Leroy O, Alcini U, Schubert D, Clarens O, et al. (2006) Different Polycomb group complexes regulate common target genes in Arabidopsis. EMBO Rep 7: 947–952.
13. Scott RJ, Spielman M (2006) Genomic imprinting in plants and mammals: how life history constrains convergence. CytoGenet Genome Res 113: 55–67.
14. Kinoshita T, Saez H, Kinoshita T, Maura A, Suppe-VJ, et al. (2007) Control of FH41 gene silencing in Arabidopsis thaliana by SINE-related direct repeats. Plant J 49: 38–45.
15. Luo M, Bilodeau P, Dennis ES, Peacock WJ, Chaudiou A (2000) Expression and parent-of-origin effects for FIS2, FIS1, and FIS in the endosperm and embryo of developing Arabidopsis seeds. Proc Natl Acad Sci U S A 97: 10637–10642.
16. Tiwari S, Spielman M, Day RG, Scott RJ (2006) Proliferative phase endosperm promoters from Arabidopsis thaliana. Plant Biotechnol J 4: 395–407.
17. Gehring M, Bubb KL, Henikoff S (2009) DNA demethylation by DNA repair. Cell 110: 33–42.
18. Kohler C, Hennig L, Spillane C, Pien S, Grauw sorn P, et al. (2003) The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERESI. Genes Dev 17: 1540–1553.
19. Tiwari S, Spielman M, Day RG, Scott RJ (2006) Proliferative phase endosperm promoters from Arabidopsis thaliana. Plant Biotechnol J 4: 395–407.
20. Gehring M, Bubb KL, Henikoff S (2009) Extensive demethylation of repetitive DNA elements by DNA repair. Science 324: 1447–1451.
21. Iwakawa H, Shinmyo A, Sekine M (2006) Polycomb-group-dependent imprinting of the Polycomb group gene MEA in Arabidopsis. Proc Natl Acad Sci U S A 97: 10637–10642.
22. Luo M, Bilodeau P, Dennis ES, Peacock WJ, Chaudiou A (2000) Expression and parent-of-origin effects for FIS2, FIS1, and FIS in the endosperm and embryo of developing Arabidopsis seeds. Proc Natl Acad Sci U S A 97: 10637–10642.
23. Gehring M, Bubb KL, Henikoff S (2009) DNA demethylation by DNA repair. Cell 110: 33–42.
24. Roy Falleh, Charles Albin-Amiot, Solveig Hauge Engbretsen, and Ida Myrher Sto for technical assistance. We thank Abed Chaudhury for homozygous fs1 seeds and Eric Richards for ddm1-2 seeds.
25. Gehring M, Reik W, Henikoff S (2009) DNA demethylation by DNA repair. Trends Genet 25: 82–90.
26. Fitz Gerald JN, Hui PS, Berger F (2009) Polycomb group-dependent imprinting of the actin regulator ACT1 regulates morphogenesis in Arabidopsis thaliana. Development 136: 3399–3404.
27. Tiwari S, Spielman M, Day RG, Scott RJ (2006) Proliferative phase endosperm promoters from Arabidopsis thaliana. Plant Biotechnol J 4: 395–407.
28. Scott RJ, Spielman M (2006) Genomic imprinting in plants and mammals: how life history constrains convergence. CytoGenet Genome Res 113: 55–67.
29. Kinoshita T, Saez H, Kinoshita T, Maura A, Suppe-VJ, et al. (2007) Control of FH41 gene silencing in Arabidopsis thaliana by SINE-related direct repeats. Plant J 49: 38–45.
30. Luo M, Bilodeau P, Dennis ES, Peacock WJ, Chaudiou A (2000) Expression and parent-of-origin effects for FIS2, FIS1, and FIS in the endosperm and embryo of developing Arabidopsis seeds. Proc Natl Acad Sci U S A 97: 10637–10642.
31. Tiwari S, Spielman M, Day RG, Scott RJ (2006) Proliferative phase endosperm promoters from Arabidopsis thaliana. Plant Biotechnol J 4: 395–407.
32. Gehring M, Bubb KL, Henikoff S (2009) DNA demethylation by DNA repair. Cell 110: 33–42.
33. Luo M, Bilodeau P, Dennis ES, Peacock WJ, Chaudiou A (2000) Expression and parent-of-origin effects for FIS2, FIS1, and FIS in the endosperm and embryo of developing Arabidopsis seeds. Proc Natl Acad Sci U S A 97: 10637–10642.
34. Tiwari S, Spielman M, Day RG, Scott RJ (2006) Proliferative phase endosperm promoters from Arabidopsis thaliana. Plant Biotechnol J 4: 395–407.
factor family in Arabidopsis: new openings to the MADS world. Plant Cell 15: 1538–1551.

48. Rios G, Lousow A, Hertel B, Breuer F, Schaefer S, et al. (2002) Rapid identification of Arabidopsis insertion mutants by non-radioactive detection of T-DNA insertion. Plant J 28: 234–243.

49. Grini PE, Schnittger A, Schwarz H, Zimmermann I, Schwab B, et al. (1999) Isolation of ethyl methanesulfonate-induced gametophytic mutants in Arabidopsis thaliana by a segregation distortion assay using the multi-marker chromosome 1.

50. Cao X, Jacobsen SE (2002) Locus-specific control of asymmetric and CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature 416: 536–540.

51. Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, et al. (2004) RNA silencing genes control de novo DNA methylation. Science 303: 1336.

52. Rajabpour AM, Xie Z, Cokus S, Bernatavichute Y, Jacobsen SE (2006) Two-step recruitment of RNA-directed DNA methylation to tandem repeats. PLoS Biol 4: e363. doi:10.1371/journal.pbio.0040363.

53. Mosher RA, Melnyk CW, Kelly KA, Dunn RM, Studholme DJ, et al. (2009) Unparental expression of PolV-dependent siRNAs in developing endosperm of Arabidopsis. Nature 460: 203–206.

54. Hulsen T, de Vlieg J, Alkema W (2008) BioVenn - a web application for the BioVenn diagrams. BMC Genomics 9: 488.

55. Barlow DP (1993) Methylation and imprinting: from host defense to gene regulation? Science 260: 309–310.

56. Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Jacobsen SE (2007) Genome-wide demethylation of Arabidopsis endosperm. Science. pp 1337–1340.

57. Tajbakhsh S, Joffre G, Köhl D, Casu B, Grotz G, et al. (2001) Evidence for widespread maternal chromosome segregation in Arabidopsis. Nature 410: 85–89.

58. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29: e45.