Epigenetic signatures associated with imprinted paternally-expressed genes in the *Arabidopsis* endosperm

Jordi Moreno-Romero[1], Gerardo Del Toro-De León, Vikash Kumar Yadav, Juan Santos-González and Claudia Köhler*

Department of Plant Biology, Uppsala BioCenter, Swedish University of Agricultural Sciences and Linnean Center for Plant Biology, Uppsala, Sweden

[1]Present address: Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UAB-UB, Campus UAB, Bellaterra, Barcelona, Spain.

*Corresponding Author (claudia.kohler@slu.se)

Running title: Epigenetic signatures of imprinted genes

ABSTRACT

Imprinted genes are epigenetically modified during gametogenesis and maintain the established epigenetic signatures after fertilization, causing parental-specific gene expression. In this study, we show that imprinted paternally-expressed genes (PEGs) in the *Arabidopsis* endosperm are marked by an epigenetic signature of Polycomb-mediated H3K27me3 together with heterochromatic H3K9me2 and CHG methylation, which specifically mark the silenced maternal alleles of PEGs. Based on the presence of this epigenetic signature on maternal alleles we were able to predict novel PEGs and could confirm several of those using published and novel datasets generated in this study. The combined activity of the otherwise largely exclusive H3K27me3 together with the heterochromatic H3K9me3 has also been reported to silence the maternal *Xist* locus in mammalian preimplantation embryos, suggesting an evolutionary conserved employment of both pathways during reproduction.

AUTHOR SUMMARY

Genomic imprinting is an epigenetic phenomenon causing parental alleles to be active depending on their parent-of-origin. In plants, imprinted genes are mainly confined to the endosperm, an ephemeral tissue supporting embryo development. Differential methylation of histone H3 on lysine 27 (H3K27me3) established by the Polycomb Repressive Complex 2 (PRC2) is a major regulatory mechanism determining activity of paternally expressed imprinted genes (PEGs) in animals and plants. Here, we show that the coding region of many PEGs is marked by an epigenetic signature of H3K27me3, H3K9me2 and CHG methylation and that the combination of these three modifications correlates with paternally-biased gene expression in the endosperm. The maternal alleles of PEGs are marked by CHG methylation in the central cell, indicating that the repressive epigenetic signature of PEGs is established before fertilization. We use the presence of the three modifications to predict novel PEGs and propose that genomic imprinting is substantially more common than previously estimated based on expression data.

INTRODUCTION

Genomic imprinting is an epigenetic phenomenon causing maternal and paternal alleles to be differentially expressed after fertilization. In plants, genomic imprinting is mainly confined to the endosperm, an ephemeral nutritive tissue supporting embryo growth, similar to the placenta in mammals [1]. The endosperm is the product of a double fertilization event, where one of the haploid sperm cells fertilizes the haploid egg cell giving rise to the diploid embryo, while the second sperm cell fertilizes the diploid central cell to give rise to the triploid endosperm [2]. Imprinted genes are epigenetically modified during gamete formation and the established epigenetic asymmetry is maintained after fertilization. Differential DNA methylation is established by the DNA glycosylase DEMETER (DME) that removes methylated cytosine residues in the central cell of the female gametophyte [3]. DME is not active in sperm, leading to differential DNA methylation between male and female genomes in the endosperm. DME acts on small transposable elements (TEs) in the vicinity of genes [4] and its activity has been connected to the expression of maternally-expressed imprinted genes (MEGs). Hypomethylation can furthermore expose binding sites for the Fertilization Independent Seed (FIS)-Polycomb Repressive Complex 2 (PRC2) [5, 6], an evolutionary conserved chromatin modifying complex that applies a trimethylation mark on histone H3 at lysine 27 (H3K27me3) [7]. The Arabidopsis FIS-PRC2 consists of the subunits MEDEA (MEA), FIS2, FERTILIZATION INDEPENDENT ENDOSPERM (FIE), and MULTICOPY SUPPRESSOR OF IRA1 (MSI1) [8] and is specifically active in the central cell of the female gametophyte and in the endosperm [9]. Repression of the maternal alleles of PEGs is mediated by the activity of the FIS-PRC2 [10], consistent with maternal PEG

79 alleles being marked by H3K27me3 [6]. In this manuscript, we addressed the
80 mechanism of maternal allele repression in PEGs. We surprisingly found that PEGs are
81 regulated by two otherwise largely exclusive epigenetic repressive pathways, the FIS2-
82 PRC2 and the pathway establishing the heterochromatin-localized H3K9me2
83 modification [11]. We demonstrate that both modifications, H3K27me3 and H3K9me2,
84 overlap on the maternal alleles of the majority of PEGs. Our data suggest that most
85 likely FIS-PRC2 acts first and is required to establish H3K9me2. Furthermore, we find
86 maternal alleles of PEGs to be marked by CHG methylation in the central cell, indicating
87 that repressive pathways establishing H3K27me3, H3K9me2 and CHG methylation act
88 in the central cell of the female gametophyte. Finally, we use the presence of the three
89 modifications to predict novel PEGs and propose that the number of PEGs predicted
90 based on expression data strongly underestimates the real number of PEGs.

91

92 **RESULTS**

93 **The maternal alleles of PEGs are marked by H3K27me3, H3K9me2, and CHG**
94 **methylation**

95 In a previous study, we revealed that H3K27me3 and H3K9me2 overlapped at paternal
96 heterochromatic regions in the endosperm, suggesting a partial functional redundancy
97 of both modifications [6]. We now addressed the question whether this redundancy
98 extends to other regions of the genome. We found that about one-third of genes marked
99 by H3K27me3 on the maternal alleles were also marked by H3K9me2 (Fig 1A,
100 hypergeometric test, P=0; S1 Fig, P=0). Genes containing both modifications on their

101 maternal alleles had significantly higher levels of H3K27me3 compared to those only
102 marked by H3K27me3 (Fig 1B, S1 Fig). The majority of double-marked genes contained
103 both modifications specifically on the maternal but not the paternal alleles (Fig 1A, S1
104 Fig) and increasing levels of H3K27me3 on maternal alleles correlated with increasing
105 levels of H3K9me2 on the maternal but not the paternal alleles (S2 Fig). We thus
106 proposed that the presence of both modifications on the maternal alleles correlates with
107 paternally-biased expression. Consistent with this notion we found that genes previously
108 identified as PEGs [12] were substantially enriched for both modifications on their
109 maternal alleles, while MEGs did not show enrichment of both marks on either maternal
110 or paternal alleles (Fig 1C and 1D, S1 Fig). This trend was independent of the direction
111 of the cross and similarly observed in Col x L*er* and L*er* x Col crosses (Fig 1 and S1
112 Fig). The presence of H3K27me3 is generally confined to gene bodies [13], consistent
113 with the observed enrichment of H3K27me3 in gene bodies of PEG maternal alleles
114 (Fig 1E and S1 Fig). Interestingly, H3K9me2 was similarly restricted to gene bodies of
115 PEGs (Fig 1E and S1 Fig), contrasting the exclusion of H3K9me2 from genic regions in
116 sporophytic tissues [14]. The maternal alleles of PEGs were also strongly enriched for
117 CHG methylation (Fig 1F, S1 and S2 Figs), consistent with the chromomethylase 3
118 (CMT3) acting in a positive feedback loop with the histone methyltransferase proteins
119 KYP/SUVH4, SUVH5, and SUVH6 [11]. Levels of CHH methylation were generally
120 higher on maternal than on paternal alleles (Fig 1F, S1 Fig), consistent with depletion of
121 CHH methylation in sperm [4, 15].

122 **Maternal-specific CHG methylation is established in the central cell and depends**
123 **on FIS-PRC2**

124 PEGs had increased levels of CHG methylation in the central cell, but low CHG
125 methylation in sperm and vegetative cells of pollen (Fig 2A), revealing that differences in
126 CHG methylation are established before fertilization. The FIS-PRC2 had been proposed
127 to promote non-CG methylation [4], raising the hypothesis that increased CHG
128 methylation on the maternal alleles of PEGs depends on the activity of FIS-PRC2 in the
129 central cell. Consistently, we observed strongly decreased CHG methylation on the
130 maternal alleles of PEGs in seeds lacking maternal FIE activity (Fig 2B), suggesting that
131 FIS-PRC2 activity is required to recruit CHG methylation. The DNA glycosylase DME is
132 required for the activation of *MEA* and *FIS2*, both encoding subunits of the FIS-PRC2
133 [16, 17]. Maternal PEG alleles had reduced levels of CHG methylation in seeds lacking
134 maternal DME activity (Fig 2B), supporting the notion that FIS-PRC2 activity is required
135 for CHG methylation establishment on maternal PEG alleles.

136 Previous work revealed that loss of FIS-PRC2 function caused activation of maternal
137 PEG alleles [10]. We addressed the question whether loss of CMT3 and the
138 redundantly acting SUVH4,5,6 [18] similarly affect silencing of maternal PEG alleles.
139 Based on published expression studies, *CMT3*, *SUVH4*, and *SUVH6* are expressed in
140 the central cell of the female gametophyte [19]. The maternal alleles of seven tested
141 PEGs remained silenced in reciprocal crosses of wild type with *cmt3* and *suvh4,5,6*
142 triple mutants (S3 Fig), indicating that stable silencing of the maternal alleles of PEGs
143 does not depend on CMT3 and SUVH4,5,6 activity before fertilization. We tested the
144 requirement of CMT3 and SUVH4,5,6 for PEG regulation after fertilization by monitoring
145 PEG expression in homozygous mutant *cmt3* and *suvh4,5,6* seeds. None of the seven
146 tested genes was significantly upregulated in seeds of *cmt3* or *suvh4,5,6* mutants (S3

147 Fig), indicating that CMT3 and SUVH4,5,6 are not required for repression of maternal
148 PEG alleles after fertilization. In the *suvh4,5,6* triple mutant, H3K9me2 is genomewide
149 eliminated in vegetative tissues [20]; however, whether H3K9me2 is similarly eliminated
150 in the central cell and endosperm remains to be shown. Therefore, deciphering the role
151 of H3K9me2 in repressing the maternal alleles of PEGs remains to be subject of future
152 investigations.

153 **Paternally-biased expression coincides with the combination of H3K27me3,**
154 **H3K9me2, and CHG methylation**

155 We addressed the question whether the presence of H3K9me2 and CHG methylation is
156 functionally relevant for maternal allele repression by testing whether the presence of
157 both modifications either alone or in combination with H3K27me3 correlates with
158 maternal allele repression. Strikingly, while the presence of either CHG methylation,
159 H3K9me2, or H3K27me3 on maternal alleles did not shift the allelic balance, the
160 combination of more than one modification shifted the balance towards preferential
161 paternal allele expression (Fig 3A, S5 Fig), suggesting that the combined presence of
162 H3K27me3, H3K9me2, and CHG methylation is a hallmark for maternal allele
163 repression. Consistently, increased bias towards paternal allele expression correlated
164 with increased levels of H3K27me3, H3K9me2, and CHG methylation on maternal
165 alleles (Fig 3B, S5 Fig).

166 We tested which combination of marks would most reliably allow predicting paternally-
167 biased genes. We assigned scores for the allele-specific presence of the three
168 modifications (see S1 Table) and tested whether highly scoring genes were enriched for

169 previously predicted PEGs [12]. The combination of CHG methylation in the central cell
170 together with maternal-specific H3K27me3 and H3K9me2 allowed to predict the highest
171 number of previously described PEGs in Col and L*er* accessions (24 out of 42 (57%),
172 [12]) in relation to the number of genes in the category with the highest score (Fig 4A,
173 S1 Table) and was chosen for further analysis. The category with the highest score was
174 significantly enriched for PEGs (hypergeometric test, *P*=1.0e-32); while categories with
175 lower scores contained only few PEGs (Fig 4A). Similarly, out of 64 PEGs that had been
176 predicted by a recent study re-evaluating previously published imprintome datasets [21],
177 40 (62%) PEGs were present in the highest score category (Fig 4A). Nearly half (96
178 genes, 46.4%) of those genes in the highest score category were significantly paternally
179 biased (Chi-square <0.05, Bonferroni corrected, S2 Table), which was significantly more
180 than the 8% paternally-biased genes identified among all genes tested (Hypergeometric
181 test, *P*= 8.9e-52). We thus conclude that the presence of the three modifications, CHG
182 in the central cell, H3K27me3 and H3K9me2 on maternal alleles in the endosperm
183 allows to predict genes with paternally-biased expression. Paternally biased genes were
184 particularly involved in transcriptional regulation (*P*=7.54e-5) and chromatin organization
185 (*P*=1.57e-3), consistent with previous reports on the functional role of PEGs [12, 22].

186 **Maternal seed coat contamination restricts the identification of PEGs**

187 Published endosperm transcriptome data contain a substantial fraction of transcripts
188 from the maternal seed coat, which may limit the correct prediction of paternally-biased
189 genes [21]. We hypothesized that there are several genes that based on their
190 epigenetic modifications (score 12, Fig 4A and S2 Table) are likely to be PEGs but
191 based on available expression data are not correctly classified. A prediction of this

192 hypothesis is that genes in the highest score category that have been classified as
193 biallelically expressed or maternally-biased are more highly expressed in the seed coat
194 than those with paternally-biased expression. We tested this hypothesis using available
195 expression data of seed coat tissue [23] at a similar developmental stage as published
196 endosperm expression data [12]. Indeed, genes with paternally-biased expression were
197 significantly less expressed in the seed coat compared to genes that based on available
198 expression data are either predicted to be biallelically expressed or maternally-biased
199 (Wilcoxon rank sum test, $P<0.005$; Fig 4B), strongly suggesting that maternal tissue
200 contamination in available expression data limits the prediction of PEGs. To further test
201 this hypothesis, we generated transgenic reporter lines containing the promoter (≃2kb)
202 and downstream genic regions fused to the green fluorescent protein (GFP) of seven
203 genes belonging to the highest score category but predicted to be maternally
204 (*AT2G33620*, *AT1G43580*, *AT1G47530*, *AT1G64660*, *AT2G30590*, *AT4G15390*) or
205 biallelically expressed (*AT5G53160*). We detected a GFP signal in the endosperm only
206 for construct *AT1G64660* (Fig 5, S3Table). For construct *AT1G47530* a signal was
207 detected in the seed coat, while for the other constructs no GFP signal was detected in
208 seeds, indicating that the regulatory elements required for the expression of those
209 genes are located outside the promoter and genic regions used to generate the reporter
210 lines. Reciprocal crosses using the *AT1G64660* reporter lines revealed that this gene is
211 indeed a PEG and strongly expressed in the endosperm when paternally, but not when
212 maternally inherited (Fig 5). This data support the hypothesis that seed coat
213 contamination limits the transcriptome-based identification of PEGs.

214 To test whether we could confirm additional PEGs that have been predicted based on
215 their epigenetic signatures, we applied INTACT (isolation of nuclei tagged in specific cell
216 types) to purify endosperm nuclei at 4 days after pollination (DAP) from Col × L*er* and
217 L*er* × Col reciprocal crosses. Isolated RNA was sequenced and profiled for allele-
218 specific gene expression. By analyzing the maternal to total reads ratio in each
219 epigenetic category, we confirmed that the genes in the highest score category (group
220 with score 12) indeed showed a clear trend towards paternally-biased expression (Fig
221 6A, S4 Table). Following previously established criteria [12], we predicted 107 PEGs
222 that were reciprocally imprinted in both directions of the crosses. There was a
223 significantly higher number of PEGs present in the highest score category compared to
224 a representative random sample of genes with informative reads (Fig 6B). Furthermore,
225 the highest score category had the highest frequency of PEGs, with other categories
226 having significantly fewer PEGs (Fig 6C). Of the 107 genes that we predicted as PEGs
227 based on our RNA sequencing data, 38 were present in the highest score category,
228 which is significantly more than expected by chance (p=3.6 e-59, Fig 6D). Out of those,
229 20 were previously predicted based on published data [10, 12, 24, 25], while 18 genes
230 are likely new high-confidence PEGs, revealing that PEGs are more common than
231 previously estimated.

232 **DISCUSSION**

233 In this study, we identified the concomitant presence of maternal-specific CHG
234 methylation, H3K27me3, and H3K9me2 as an epigenetic signature for paternally-biased
235 expression in the endosperm. We furthermore predict that there are substantially more
236 PEGs than previously reported in Arabidopsis [10, 12, 24], suggesting that in

237 Arabidopsis the number of PEGs exceeds the number of MEGs. Recent re-evaluation of
238 published imprintome data of Arabidopsis revealed that a large number of previously
239 predicted MEGs were seed coat expressed genes, while the number of PEGs was
240 underestimated [21]. Our study supports and extends this notion by showing that there
241 is a large number of paternally-biased genes that likely failed to be identified in previous
242 studies because of maternal seed coat contamination or early stage-specific
243 expression.

244 A previous study reported that the maternal alleles of PEGs in *A. lyrata* are marked by
245 CHG methylation and implicated that closely related species use different mechanisms
246 to regulate imprinted gene expression [26]. Our study reveals that similar to *A. lyrata* the
247 maternal alleles of many PEGs are also marked by CHG methylation in *A. thaliana*,
248 highlighting that epigenetic mechanisms employed to maintain monoallelic expression
249 are rather conserved between related species. Interestingly, while the maternal alleles
250 of PEGs in maize are marked by H3K27me3, they are not marked by CHG methylation
251 [27], indicating that diverged species may use different mechanism in maintaining
252 maternal allele repression.

253 How H3K9me2 and CHG methylation are established at PRC2 target genes remains to
254 be studied; however, the strong activation of maternal PEG alleles upon loss of FIS-
255 PRC2 function [10] suggests that H3K9me2 and CHG methylation require FIS-PRC2
256 function. The PRC2 is generally targeting genes with specific roles during development
257 [28], while complexes establishing H3K9me2 are mainly targeting TEs localized in
258 heterochromatic regions of the genome [29]. This functional division of PRC2 and
259 machineries establishing heterochromatic marks is conserved in plants as well as in

260 mammals [30]; however, there are notable exceptions to this rule in both groups of
261 organisms. In rice seedlings, about one-third of H3K27me3 marked genes are also
262 marked by CHG and CHH methylation [31]. Similar to the findings reported in our study,
263 higher levels of H3K27me3 correlate with higher CHG and CHH methylation in gene
264 bodies of rice [31]. The rice H3K27me3 methyltransferase SDG711 physically interacts
265 with the CHH methyltransferase OsDMR2 and the SRA-domain containing SUVH
266 protein SDG703, uncovering a mechanistic connection between PRC2 and non-CG
267 methylation. In mammalian cells, H3K9 methyltransferases colocalize with PRC2 [32-
268 34], revealing crosstalk between these two major epigenetic silencing pathways that
269 likely is required for stable gene silencing. The imprinted maternal *Xist* locus encoding
270 an X-linked long-noncoding RNA is covered by H3K27me3 and H3K9me3 in
271 preimplantation embryos [35, 36]. Importantly, loss of maternal H3K27me3 induces *Xist*
272 activation, indicating that maternal H3K27me3 is the major imprinting mark of *Xist* [36].
273 This is strikingly similar to findings made in this study revealing H3K27me3 as the major
274 repressive mark for PEGs. Recent work revealed that maternal H3K27me3 controls
275 DNA methylation-independent imprinting in mammalian preimplantation embryos [37].
276 While imprinted expression of most genes is lost in the embryonic cell lineage, few
277 genes maintain their imprinted expression in the extra-embryonic cell lineage [37]. The
278 *Xist* locus that is marked by H3K27me3 and H3K9me3 is among those loci that remain
279 imprinted in extra-embryonic tissues [36]. Whether the presence of both marks
280 distinguishes those genes that maintain their imprinted expression from those that
281 become biallelically expressed remains to be tested, but we consider this a very
282 attractive hypothesis. We speculate that the presence of both marks in certain tissue

283 types of mammals and flowering plants is a conserved epigenetic signature marking
284 stably repressed genes.

285

286 **MATERIAL AND METHODS**

287 **Plant Material and Growth Conditions**

288 All seeds were surface sterilized (5% sodium hypochlorite and 0.01% Triton X-100),
289 stratified for 2 days at 4°C and germinated on half-strength Murashige and Skoog
290 medium containing 1% sucrose under long-day conditions (16 h light/8 h darkness;
291 21°C). Plants were transferred to soil after 10 to 12 days and grown under long day
292 conditions. The *cmt3-11* (SALK_148381; [38]) and *suvh456* mutants (kindly provided by
293 Judith Bender) used in this study are in the Col-0 background.

294

295 **Imprinting Assays and Expression Analysis**

296 To generate siliques of indicated crosses, three to five flowers were emasculated, hand-
297 pollinated, and harvested at 4 DAP. RNA extraction was performed using the MagJET
298 Plant RNA Purification Kit (Thermo Scientific) following the manufacturer´s instructions.
299 Residual DNA was removed using Invitrogen DNase I (Amplification Grade), and cDNA
300 was synthesized using the Fermentas first strand cDNA synthesis kit according to the
301 manufacturer's instructions. Quantitative PCR was performed using a MyiQ5 real-time
302 PCR detection system (Bio-Rad) and Solis BioDyne-5x Hot FIREPol EvaGreen qPCR
303 Mix Plus (ROX, Solis BioDyne). For the imprinting-by-sequencing assay, the PCR

304 products were purified and analyzed by Sanger sequencing. For the imprinting-by-
305 restriction enzyme digestion assay, the PCR products were purified and digested.
306 Restriction enzymes and primers used are listed in the S5 Table.

307 **Data analysis**

308 We made use of endosperm-specific ChIP-seq data that have been previously
309 generated in our group [6]. Data correspond to pooled biological triplicates with ChIP
310 signals being normalized with H3 ChIP data by calculating the log2 ratio in 150-bp bins
311 across the genome. Data were standardized and normalized with a z-score
312 transformation [39]. Metagene plots over genes were constructed between -2 kb and
313 +2 kb by calculating mean levels of methylation signals in 100-bp bins in the flanks of
314 the genes and in 40 equally long bins between the transcriptional start and stop. Gene
315 z-scores were calculated as an average of z-scores over the gene body. DNA
316 methylation data of *fie* and *dme* and their corresponding wild types are from [4]. DNA
317 methylation data of the central cell, sperm, and vegetative cells are from [3] and [15].
318 Endosperm expression data are from [12] and seed coat expression data from [23].

319 **Endosperm nuclei isolation and RNA sequencing**

320 We performed Col × L*er* reciprocal crosses using *Arabidopsis* lines expressing
321 *PHE1::NTF* and *PHE1::BirA* (lines referred as INT hereafter) [40]. To facilitate the
322 crosses, we used the male sterile mutants *pistillata* (*pi-1*, in L*er* accession) and *dde2* (in
323 Col accession containing INT) as female parents and pollinated them with the INT line
324 (Col accession) and L*er* wild type, respectively. A total of 250mg of siliques per sample
325 was collected at 4DAP. Tissue homogenization and nuclei purification were performed

326 from three biological replicates as previously described [40]. Total RNA was extracted
327 from purified nuclei using the mirVana Isolation Kit Protocol (Ambion). mRNA extraction
328 was performed using NEBNext Poly(A) mRNA Magnetic Isolation and the Libraries were
329 prepared with the NEBNext Ultra II RNA Library Prep Kit from Illumina. Samples were
330 sequenced at the National Genomic Infrastructure (NGI) from SciLife Laboratory
331 (Uppsala, Sweden) on an Illumina HiSeq2500 in paired-end 125bp read length. Reads
332 were trimmed and mapped in single-end mode to the Arabidopsis (TAIR10) genome
333 previously masked for *rRNA* genes and for SNP positions between Col and
334 L*er*accessions, using TopHat v2.1. [41] (parameters adjusted as -g 1 -a 10 -i 40 -I 5000
335 -F 0 -r 130). Gene expression was normalized to reads per kilobase per million mapped
336 reads (RPKM) using GFOLD40 [42]. For discrimination of maternal and paternal
337 transcripts, reads were assigned to the Col or L*er* genomes using single nucleotide
338 polymorphisms (SNPs) between the accessions using SNPsplit [43]. We calculated
339 contamination levels based on the deviation of the read counts from the expected 2:1
340 maternal/paternal genome ratio in the endosperm [40] (S6 Table). Based on this
341 analysis, one of the triplicates of the cross direction Col × L*er* was not included in the
342 downstream analyses. To increase the statistical power to detect parentally-biased
343 genes, we merged libraries from two replicates of Col × L*er* and three replicates of L*er* ×
344 Col for downstream analysis.

345 **Allele-specific expression analysis**

346 We defined a minimum threshold of 20 and 30 informative reads for Col × L*er* (2
347 replicates) and L*er* × Col (3 replicates) crosses, respectively. Statistical differences
348 between maternal and paternal read counts for each gene were calculated using a Chi-

349 square test, considering genes with a false discovery rate adjusted *P*-value of less than
350 0.01. Additionally, MEGs required to have at least 85% maternal informative reads in
351 both directions of the reciprocal cross and PEGs to have at least 50% paternal
352 informative reads in both directions of the reciprocal cross, following previously defined
353 conditions [12].

354 **Generation of reporter constructs and transgenic lines**

355 For the generation of reporter constructs, we used the ClonExpress® MultiS One Step
356 Cloning kit. Promoters (≃2kb) and genic sequences of *AT2G33620*, *AT1G43580*,
357 *AT1G47530*, *AT1G64660*, *AT2G30590*, *AT4G15390*, and *AT5G53160* were amplified
358 from WT Col-0 genomic DNA using primers specified in S5 Table and cloned into vector
359 pB7FWG.0. Constructs were transformed into *Agrobacterium tumefaciens* strain
360 GV3101 and *Arabidopsis* plants were transformed using the floral dip method [44]. Ten
361 transgenic lines per construct were generated and analyzed.

362

363 **Imprinting analyses**

364 For reciprocal crosses, designated female partners were emasculated at 1–2 days prior
365 to anthesis. Two days after emasculation pistils were hand-pollinated with respective
366 pollen donors. Seeds were dissected from the siliques and mounted on a microscope
367 slide for imaging and counting at 2 and 4 DAP. For fluorescence analyses, seeds were
368 stained with 0.1 mg/mL propidium iodide (PI) solution in 7% glucose. Seeds of
369 reciprocal crosses of reporter lines were analyzed under confocal microscopy on a
370 Zeiss 800 Inverted Axio Observer with a supersensitive GaASp detector. Images were

371 acquired, analyzed and exported using Zeiss ZEN software. For each reporter construct
372 50-60 seeds per line were analyzed.
373

374 **Acknowledgements**

375 This research was supported by a European Research Council Starting Independent
376 Researcher grant (to C.K.), a grant from the Swedish Science Foundation (to C.K.), a
377 grant from the Knut and Alice Wallenberg Foundation (to C.K.), and an EMBO
378 fellowship (to G.D.T.D.L). Sequencing was performed by the SNP&SEQ Technology
379 Platform, Science for Life Laboratory at Uppsala University, a national infrastructure
380 supported by the Swedish Research Council (VRRFI) and the Knut and Alice
381 Wallenberg Foundation.
382

383 **References:**

384 1. Rodrigues JA, Zilberman D: Evolution and function of genomic imprinting in
385 plants. Genes Dev 2015;29: 2517-2531.

386 2. Li J, Berger F: Endosperm: food for humankind and fodder for scientific
387 discoveries. New Phytol 2012;195: 290-305.

388 3. Park K, Kim MY, Vickers M, Park JS, Hyun Y, Okamoto T, Zilberman D, Fischer
389 RL, Feng X, Choi Y et al: DNA demethylation is initiated in the central cells of
390 Arabidopsis and rice. Proc Natl Acad Sci USA 2016;113: 15138-15143.

391 4. Ibarra CA, Feng X, Schoft VK, Hsieh TF, Uzawa R, Rodrigues JA, Zemach A,
392 Chumak N, Machlicova A, Nishimura T et al: Active DNA demethylation in plant
393 companion cells reinforces transposon methylation in gametes. Science
394 2012;337: 1360-1364.

395 5. Weinhofer I, Hehenberger E, Roszak P, Hennig L, Köhler C: H3K27me3 profiling
396 of the endosperm implies exclusion of Polycomb group protein targeting by DNA
397 methylation. PLoS Genet 2010;6: e1001152.

398 6. Moreno-Romero J, Jiang H, Santos-Gonzalez J, Kohler C: Parental epigenetic
399 asymmetry of PRC2-mediated histone modifications in the Arabidopsis
400 endosperm. EMBO J 2016;35: 1298-1311.

401 7. Simon JA, Kingston RE: Occupying chromatin: Polycomb mechanisms for getting
402 to genomic targets, stopping transcriptional traffic, and staying put. Mol Cell
403 2013;49: 808-824.

404 8. Mozgova I, Hennig L: The Polycomb Group protein regulatory network. Annu Rev
405 Plant Biol 2015;66: 269-296.

406 9. Luo M, Bilodeau P, Dennis ES, Peacock WJ, Chaudhury A: Expression and
407 parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of
408 developing Arabidopsis seeds. Proc Natl Acad Sci USA 2000;97: 10637-10642.

409 10. Hsieh TF, Shin J, Uzawa R, Silva P, Cohen S, Bauer MJ, Hashimoto M, Kirkbride
410 RC, Harada JJ, Zilberman D et al: Regulation of imprinted gene expression in
411 Arabidopsis endosperm. Proc Natl Acad Sci USA 2011;108: 1755-1762.

412 11. Law JA, Jacobsen SE: Establishing, maintaining and modifying DNA methylation
413 patterns in plants and animals. Nat Rev Genet 2010;11: 204-220.

414 12. Pignatta D, Erdmann RM, Scheer E, Picard CL, Bell GW, Gehring M: Natural
415 epigenetic polymorphisms lead to intraspecific variation in Arabidopsis gene
416 imprinting. Elife 2014;3: e03198.

417 13. Shu H, Nakamura M, Siretskiy A, Borghi L, Moraes I, Wildhaber T, Gruissem W,
418 Hennig L: Arabidopsis replacement histone variant H3.3 occupies promoters of
419 regulated genes. Genome Biol 2014;15: R62.

420 14. Roudier F, Ahmed I, Berard C, Sarazin A, Mary-Huard T, Cortijo S, Bouyer D,
421 Caillieux E, Duvernois-Berthet E, Al-Shikhley L et al: Integrative epigenomic
422 mapping defines four main chromatin states in Arabidopsis. EMBO J 2011;30:
423 1928-1938.

424 15. Calarco JP, Borges F, Donoghue MT, Van Ex F, Jullien PE, Lopes T, Gardner R,
425 Berger F, Feijo JA, Becker JD et al: Reprogramming of DNA methylation in
426 pollen guides epigenetic inheritance via small RNA. Cell 2012;151: 194-205.

427 16. Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen
428 SE, Fischer RL: DEMETER, a DNA glycosylase domain protein, is required for
429 endosperm gene imprinting and seed viability in Arabidopsis. Cell 2002;110: 33-
430 42.

431 17. Jullien PE, Kinoshita T, Ohad N, Berger F: Maintenance of DNA methylation
432 during the Arabidopsis life cycle is essential for parental imprinting. Plant Cell
433 2006;18: 1360-1372.

434 18. Ebbs ML, Bender J: Locus-specific control of DNA methylation by the
435 Arabidopsis SUVH5 histone methyltransferase. Plant Cell 2006;18 1166-1176.

436 19. Wuest SE, Vijverberg K, Schmidt A, Weiss M, Gheyselinck J, Lohr M, Wellmer F,
437 Rahnenfuhrer J, von Mering C, Grossniklaus U et al: Arabidopsis female
438 gametophyte gene expression map reveals similarities between plant and animal
439 gametes. Curr Biol 2010;20: 506-512.

440 20. Stroud H, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE:
441 Comprehensive analysis of silencing mutants reveals complex regulation of the
442 Arabidopsis methylome. Cell 2013;152: 352-64.

443 21. Schon MA, Nodine MD: Widespread Contamination of Arabidopsis Embryo and
444 Endosperm Transcriptome Data Sets. Plant Cell 2017;29: 608-617.

445 22. Waters AJ, Bilinski P, Eichten SR, Vaughn MW, Ross-Ibarra J, Gehring M,
446 Springer NM: Comprehensive analysis of imprinted genes in maize reveals allelic
447 variation for imprinting and limited conservation with other species. Proc Natl
448 Acad Sci USA 2013;110: 19639-19644.

449 23. Belmonte MF, Kirkbride RC, Stone SL, Pelletier JM, Bui AQ, Yeung EC,
450 Hashimoto M, Fei J, Harada CM, Munoz MD et al: Comprehensive
451 developmental profiles of gene activity in regions and subregions of the
452 Arabidopsis seed. Proc Natl Acad Sci USA 2013;110: 435-444.

453 24. Wolff P, Weinhofer I, Seguin J, Roszak P, Beisel C, Donoghue MT, Spillane C,
454 Nordborg M, Rehmsmeier M, Köhler C et al: High-Resolution Analysis of Parent-
455 of-Origin Allelic Expression in the Arabidopsis Endosperm. PLoS Genet 2011;7:
456 e1002126.

457 25. Gehring M, Missirian V, Henikoff S: Genomic Analysis of Parent-of-Origin Allelic
458 Expression in Arabidopsis thaliana Seeds. PLoS One 2011;6: e23687.

459 26. Klosinska M, Picard CL, Gehring M: Conserved imprinting associated with unique
460 epigenetic signatures in the Arabidopsis genus. Nat Plants 2016;2: 16145.

461 27. Zhang M, Xie S, Dong X, Zhao X, Zeng B, Chen J, Li H, Yang W, Zhao H, Wang
462 G et al: Genome-wide high resolution parental-specific DNA and histone
463 methylation maps uncover patterns of imprinting regulation in maize. Genome
464 Res 2014;24: 167-176.

465 28. Mozgova I, Kohler C, Hennig L: Keeping the gate closed: functions of the
466 polycomb repressive complex PRC2 in development. Plant J 2015;83: 121-132.

467 29. Du J, Johnson LM, Jacobsen SE, Patel DJ: DNA methylation pathways and their
468 crosstalk with histone methylation. Nat Rev Mol Cell Biol 2015;16: 519-532.

469 30. Alabert C, Groth A: Chromatin replication and epigenome maintenance. Nat Rev
470 Mol Cell Biol 2012;13: 153-167.

471 31. Zhou S, AUID- ORCID: ht, Liu X, Zhou C, Zhou Q, Zhao Y, Li G, Zhou DX:
472 Cooperation between the H3K27me3 Chromatin Mark and Non-CG Methylation
473 in Epigenetic Regulation. Plant Physiol 2016;172: 1131-1141.

474 32. Bilodeau S, Kagey MH, Frampton GM, Rahl PB, Young RA: SetDB1 contributes
475 to repression of genes encoding developmental regulators and maintenance of
476 ES cell state. Genes Dev 2009;23: 2484-2489.

477 33. Yuan P, Han J, Guo G, Orlov YL, Huss M, Loh YH, Yaw LP, Robson P, Lim B,
478 Ng HH et al: Eset partners with Oct4 to restrict extraembryonic trophoblast
479 lineage potential in embryonic stem cells. Genes Dev 2009;23: 2507-2520.

480 34. Mozzetta C, Pontis J, Fritsch L, Robin P, Portoso M, Proux C, Margueron R, Ait-
481 Si-Ali S: The histone H3 lysine 9 methyltransferases G9a and GLP regulate
482 polycomb repressive complex 2-mediated gene silencing. Mol Cell 2014;53: 277-
483 289.

484 35. Fukuda A, Tomikawa J, Miura T, Hata K, Nakabayashi K, Eggan K, Akutsu H,
485 Umezawa A: The role of maternal-specific H3K9me3 modification in establishing
486 imprinted X-chromosome inactivation and embryogenesis in mice. Nat Commun
487 2014;5: 5464.

488 36. Inoue A, Jiang L, Lu F, Zhang Y: Genomic imprinting of Xist by maternal
489 H3K27me3. Genes Dev 2017;31: 1927-1932.

490 37. Inoue A, Jiang L, Lu F, Suzuki T, Zhang Y: Maternal H3K27me3 controls DNA
491 methylation-independent imprinting. Nature 2017;547: 419-424.

492 38. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR,
493 Shinn P, Pellegrini M, Jacobsen SE et al: Genome-wide high-resolution mapping
494 and functional analysis of DNA methylation in arabidopsis. Cell 2006;126: 1189-
495 1201.

496 39. Cheadle C, Vawter MP, Freed WJ, Becker KG: Analysis of microarray data using
497 Z score transformation. J Mol Diagn 2003;5: 73-81.

498 40. Moreno-Romero J, Santos-Gonzalez J, Hennig L, Köhler C: Applying the
499 INTACT method to purify endosperm nuclei and to generate parental-specific
500 epigenome profiles. Nat Protoc 2017;12: 238-254.

501 41. Trapnell C, Pachter L, Salzberg SL: TopHat: discovering splice junctions with
502 RNA-Seq. Bioinformatics 2009;25: 1105-1111.

503 42. Feng J, Meyer CA, Wang Q, Liu JS, Shirley Liu X, Zhang Y: GFOLD: a
504 generalized fold change for ranking differentially expressed genes from RNA-seq
505 data. Bioinformatics 2012;28: 2782-2788.

506 43. Krueger F, Andrews SR: SNPsplit: Allele-specific splitting of alignments between
507 genomes with known SNP genotypes. F1000Res 2016;5: 1479.

508 44. Clough SJ, Bent AF: Floral dip: a simplified method for agrobacterium-mediated
509 transformation of Arabidopsis thaliana. Plant J 1998;16: 735-743.

510

511

512

513 **Figure Legends**

514 **Fig 1. H3K27me3, H3K9me2, and CHG methylation overlap on maternal alleles of**
515 **PEGs.**

516 **(A)** Overlap of genes marked by H3K27me3 and H3K9me2 on the maternal alleles.
517 Dark red inset marks genes that have H3K27me3 and H3K9me2 exclusively on their
518 maternal alleles (excluding genes also marked on the paternal alleles).

519 **(B)** Genes marked by H3K27me3 and H3K9me2 have higher levels of H3K27me3. Box
520 plots show mean values of z-scored H3K27me3 and H3K9me2 of genes where the
521 maternal alleles are either marked by H3K27me3 alone or by a combination of
522 H3K27me3 and H3K9me2. Genes containing maternal-specific H3K27me3 and
523 H3K9me2 are particularly enriched for H3K27me3. Boxes show medians and the
524 interquartile range, error bars show the full range excluding outliers. Notches show the
525 95% confidence interval for the median. Asterisks mark significance (Wilcoxon rank sum
526 test, P <1.0E-3). Presence of the mark corresponds to the mark highest score defined in
527 S1 Table.

528

529 **(C)** The majority of PEGs are marked by H3K27me3 and H3K9me2 on their maternal
530 alleles. Overlap of PEGs [12] with maternal and paternal H3K27me3 and H3K9me2.
531 **(D)** Levels of H3K27me3 and H3K9me2 on PEGs, MEGs, and all genes. Box plots (as
532 specified in panel B) show mean values of z-scored H3K27me3 and H3K9me2.
533 Asterisks mark significance (Wilcoxon rank sum test, *P* <1.0E-3).
534 **(E)** Metagene plots showing average distribution of z-score normalized H3K27me3 and
535 H3K9me2 levels on maternal and paternal alleles of PEGs [12].
536 **(F)** Levels of DNA methylation in each sequence context (mCG, mCHG and mCHH) on
537 PEGs, MEGs, and all genes. Box plots (as specified in panel B) show mean values of
538 relative DNA methylation. Asterisks mark significance (Wilcoxon rank sum test, *P*<5.0E-
539 3). Data shown correspond to cross Col × L*er* [6], for reciprocal cross direction see S1
540 Fig.
541

542 **Fig 2. Maternal CHG methylation is established in the central cell and depends on**
543 **FIS-PRC2.**
544 **(A)** Box plots showing mean values of relative DNA methylation in the central cell,
545 sperm cell, and vegetative cell for PEGs, MEGs, and all genes. Boxes show medians
546 and the interquartile range, error bars show the full range excluding outliers. Notches
547 show the 95% confidence interval for the median. Asterisk marks significance (Wilcoxon
548 rank sum test, *P* <5.0E-3).

549 **(B)** Box plots showing mean values of CHG methylation for PEGs and MEGs on
550 maternal and paternal alleles in wild-type, *fie* (left panels) and *dme* (right panels)
551 endosperm. Asterisk marks significance (Wilcoxon rank sum test, *P* <5.0E-2).
552
553 **Fig 3. Paternally-biased expression coincides with the combination of H3K27me3,**
554 **H3K9me2, and CHG methylation.**
555 **(A)** Box plots show mean values of maternal to total reads for genes marked by one or
556 several of the following modifications: CHG methylation in the central cell, difference of
557 CHG methylation, H3K27me3, or H3K9me2 levels on maternal and paternal alleles in
558 the endosperm. For each modification only genes present in the highest scoring
559 category were included in the analysis (see S1 Table). Boxes show medians and the
560 interquartile range, error bars show the full range excluding outliers. Widths of boxes
561 are proportional to square-root of the number of genes in each category. Blue line
562 indicates the expected ratio for biallelically expressed genes.
563 **(B)** Paternal expression bias correlates with high levels of CHG methylation in the
564 central cell and H3K27me3 and H3K9me2 on maternal alleles in the endosperm.
565 Plotted are mean values of CHG methylation, H3K27me3, and H3K9me2 on maternal
566 and paternal alleles for genes with deviating parental expression levels from the
567 expected two maternal to one paternal ratio.
568 Data shown correspond to cross Col × L*er*, for reciprocal cross direction see S4 Fig.
569
570 **Fig 4. Epigenetic signatures allow predicting paternally-expressed imprinted**
571 **genes (PEGs).**

572 **(A)** The combination of CHG methylation in the central cell together with maternal-
573 specific H3K27me3 and H3K9me2 allows predicting previously described PEGs [12,
574 21]. Scores are calculated as described in the S1 Table. High and low scores
575 correspond to high or low levels of CHG methylation, H3K27me3, and H3K9me2,
576 respectively. Number of PEGs corresponds to PEGs previously predicted [12] (i) and
577 [21] (ii) in Col and L*er* accessions. Total genes correspond to all genes present in the
578 corresponding categories. Box plots show mean values of maternal to total reads in the
579 indicated categories. Boxes show medians and the interquartile range, error bars show
580 the full range excluding outliers. Blue line indicates the expected ratio for biallelically
581 expressed genes.
582 **(B)** Box plots show relative expression level of genes in the seed coat of torpedo stage
583 embryos [23]. Genes present in the highest score category (S2 Table) were grouped in
584 paternally biased genes (Chi-square test, $P<0.05$ Bonferroni corrected) and not-
585 paternally biased genes. Asterisks mark significance (Wilcoxon rank sum test, P <5.0E-
586 3).
587
588 **Fig 5. Reporter lines for *AT1G64660* confirm imprinted expression.**
589 Representative images of seeds derived after reciprocal crosses of the *AT1G64660*
590 reporter line (fusion with green fluorescent protein (GFP)) with wild-type (WT) plants.
591 GFP fluorescence was detected in the seed coat when the *AT1G64660* reporter was
592 maternally inherited, but endosperm-specific expression was only detected when the
593 *AT1G64660* reporter was paternally inherited. Seeds at 2 DAP were used for imaging.
594 Red stain is propidium iodide. Scale bars correspond to 50 μm.

595 **Fig 6. Confirmation of predicted PEGs in RNA samples of INTACT-purified**
596 **endosperm nuclei.**
597 **(A)** Distribution of maternal to total read values for genes in each epigenetic category.
598 The distribution for the total population of genes with informative reads (U) is included.
599 Genes with the highest levels of CHG methylation, H3K27me3, and H3K9me2 show a
600 clear trend towards paternally-biased expression (around 0.5). Values correspond to
601 combined values of Col × L*er* and L*er* × Col crosses. Boxes show medians and the
602 interquartile range, lines show the full range of outliers. Blue line indicates the expected
603 ratio for biallelically expressed genes (0.66).
604 **(B)** Parent-of-origin distribution of the group of genes with score 12, compared with a
605 random group of the same size. Significantly more PEGs are detected in the group of
606 genes with score 12, compared with the random sample (Wilcoxon test, *P*-values
607 correspond to the comparison with the total population of PEGs in the whole RNAseq
608 data). The scatter plots show maternal to total read values for Col × L*er* and L*er* × Col
609 reciprocal crosses. Colors indicate genes with significant parent-of-origin biased
610 expressed: MEGs, 85% or more of maternal reads in each reciprocal cross; PEGs, 50%
611 or more of paternal reads in each reciprocal cross. Maternally and paternally biased
612 genes show significantly biased expression ($P <= 0.01$) in both reciprocal crosses.
613 **(C)** Percentage of PEGs, MEGs and biallelically expressed genes in each epigenetic
614 category. Color code indicates the percentage of genes within each category.
615 Categories with genes having lower levels of CHG methylation, H3K27me3, and
616 H3K9me2 marks have significantly fewer PEGs than the highest scoring category (Chi-

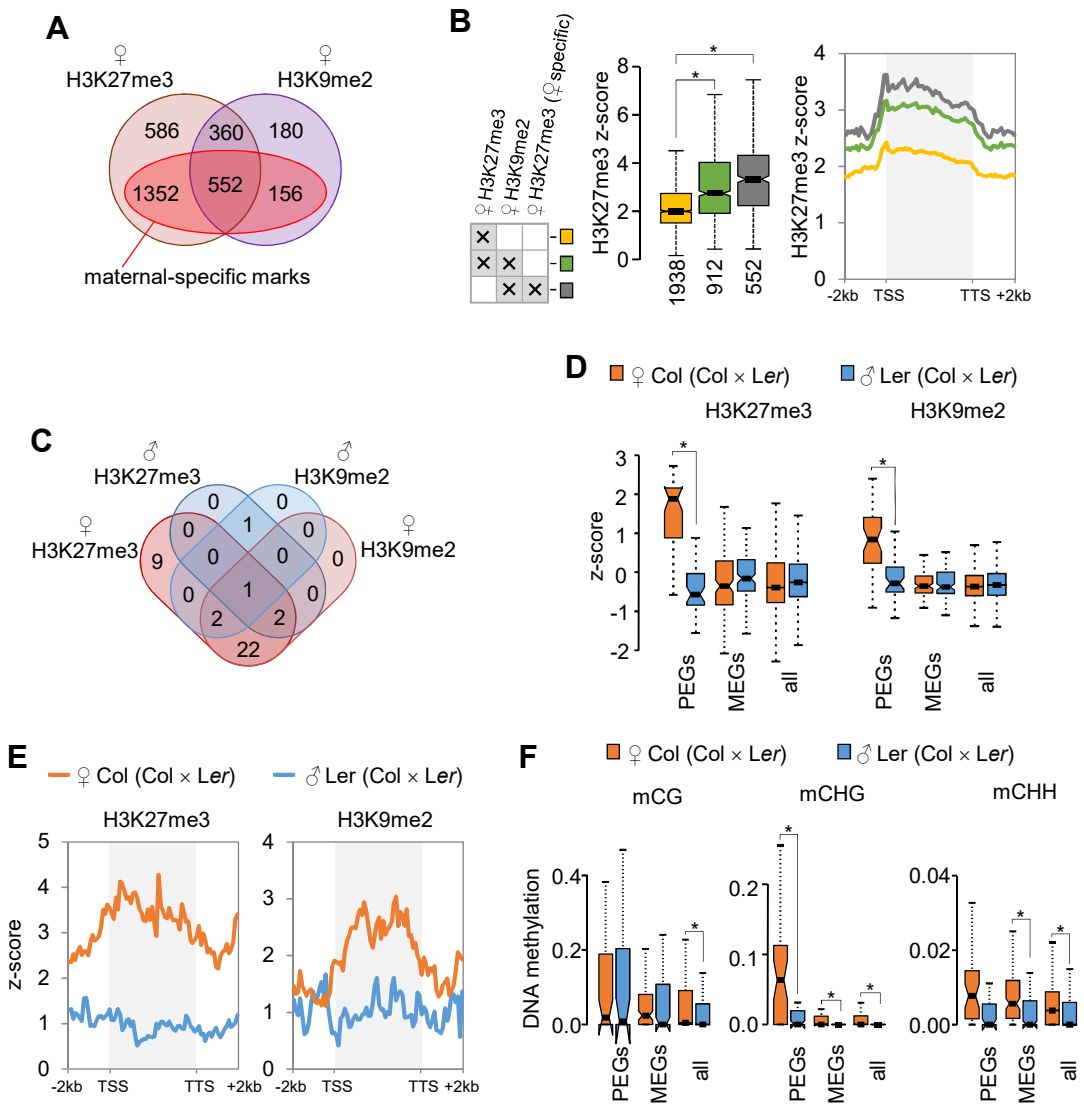
617 square test, *P*<1e-8 in categories 1-10; *P*=0.001 in category 11). Only PEGs, MEGs and
618 Biallelic categories were considered.
619 **(D)** Overlap of PEGs that were detected in our INTACT-based RNAseq samples (blue)
620 with genes predicted as PEGs based on their epigenetic score 12 (green) and
621 previously reported PEGs (yellow) by [10, 12, 24, 25]. New predicted and confirmed
622 PEGs in bold. Statistical significance was calculated using a hypergeometric test.
623

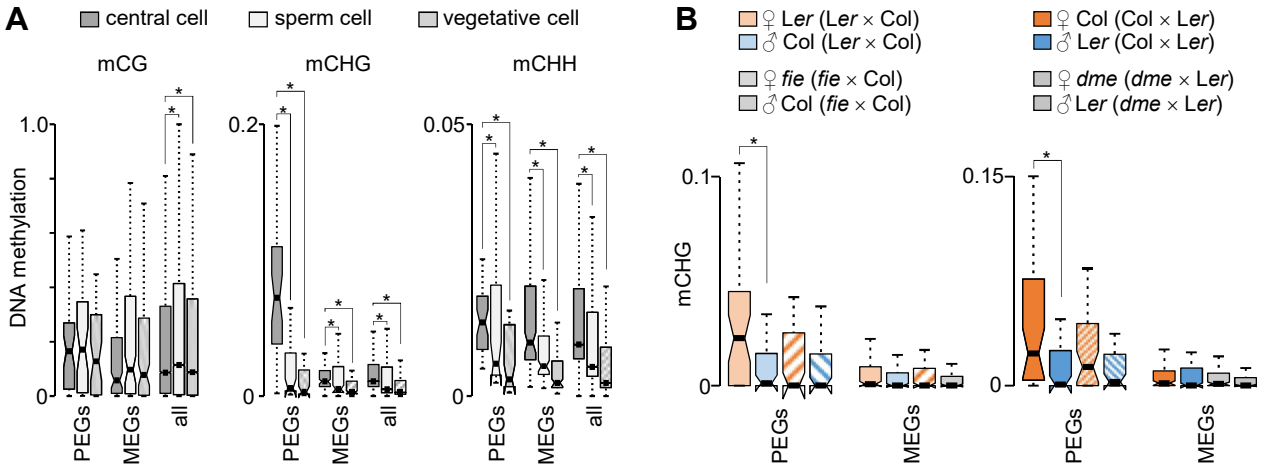
624 **Supporting information**

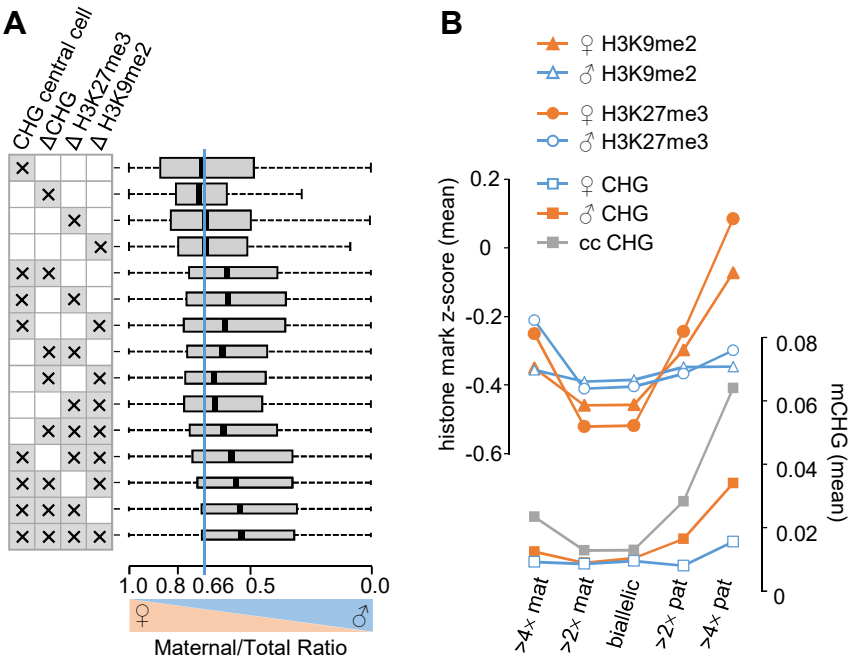
625 **S1 Table.** Scores assigned to genes based on the presence of CHG methylation in the
626 central cell as well as H3K9me2 and H3K27me3 on maternal alleles in the endosperm.
627 **S2 Table.** Seed coat expression of genes with maximum score (12) based on S1Table.
628 **S3 Table:** Fluorescence analysis of 4DAP seeds derived after indicated reciprocal
629 crosses.
630 **S4 Table.** Parent-of-origin RNAseq dataset of 4DAP INTACT-purified endosperm of Col
631 × L*er* reciprocal crosses.
632 **S5 Table.** List of primers used in this study.
633 **S6 Table.** Quality of sequencing samples.
634

635 **S1 Fig.** H3K27me3, H3K9me2, and CHG methylation overlap on maternal alleles of
636 PEGs.
637 **S2 Fig.** Total and parental-specific H3K27me3, H3K9me2, and CHG methylation levels
638 in the endosperm of previously described PEGs [12].
639 **S3 Fig.** Maternal alleles remain silenced in *suvh4,5,6* and *cmt3* mutants.

640 **S4 Fig.** Increasing levels of H3K27me3 on maternal alleles correlates with increasing
641 levels of H3K9me2.
642 **S5 Fig.** Paternally-biased expression coincides with the combination of H3K27me3,
643 H3K9me2, and CHG methylation.
644
645
646
647

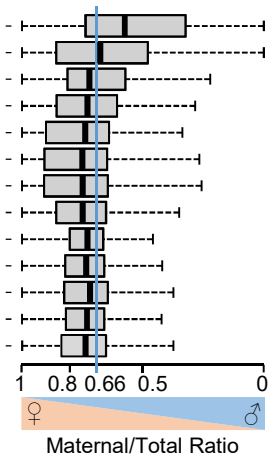

A
♀ H3K27me3
♀ H3K9me2
586
360
180
1352
552
156
maternal-specific marks
B
♀H3K27me3
♀H3K9me2
♀H3K27me3 (♀specific)
H3K27me3 z-score
1938
912
552
-2kb
TSS
TTS
+2kb
C
♂ H3K27me3
♂ H3K9me2
♀ H3K27me3
♀ H3K9me2
D
♀ Col (Col × Ler)
♂ Ler (Col × Ler)
H3K27me3
H3K9me2
z-score
PEGs
MEGs
all
E
♀ Col (Col × Ler)
♂ Ler (Col × Ler)
H3K27me3
H3K9me2
z-score
F
mCG
mCHG
mCHH
DNA methylation


A
central cell
sperm cell
vegetative cell
mCG
mCHG
mCHH
DNA methylation
1.0
0
0.2
0
0.05
0
PEGs
MEGs
all
B
♀ Ler (Ler × Col)
♂ Col (Ler × Col)
♀ Col (Col × Ler)
♂ Ler (Col × Ler)
♀ fie (fie × Col)
♂ Col (fie × Col)
♀ dme (dme × Ler)
♂ Ler (dme × Ler)
mCHG
0.1
0
0.15
0
PEGs
MEGs

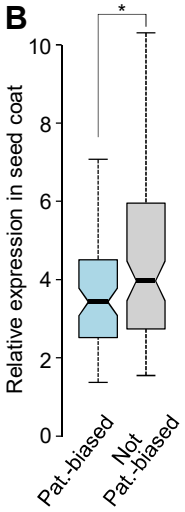

A
CHG central cell
ΔCHG
Δ H3K27me3
Δ H3K9me2
1.0 0.8 0.66 0.5 0.0
♀
♂
Maternal/Total Ratio
B
♀ H3K9me2
♂ H3K9me2
♀ H3K27me3
♂ H3K27me3
♀ CHG
♂ CHG
cc CHG
histone mark z-score (mean)
0.2
0
-0.2
-0.4
-0.6
mCHG (mean)
0.08
0.06
0.04
0.02
0
>4× mat
>2× mat
biallelic
>2× pat
>4× pat

A

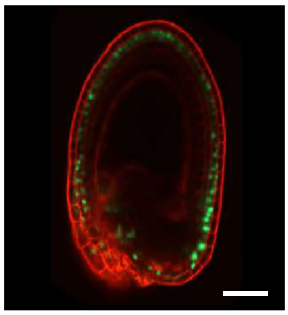
score	PEGs (i)	PEGs (ii)	PEGs (shared)	Total genes
12	24	40	17	207
11	3	3	2	80
10	3	3	2	114
9	1	0	0	108
8	0	1	0	103
7	1	0	0	206
6	1	3	1	405
5	0	4	0	787
4	2	1	0	4516
3	2	4	2	2490
2	1	1	1	351
1	1	1	1	972
0	3	2	2	1995



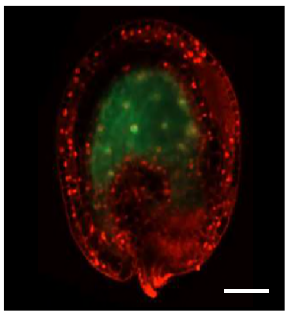
B

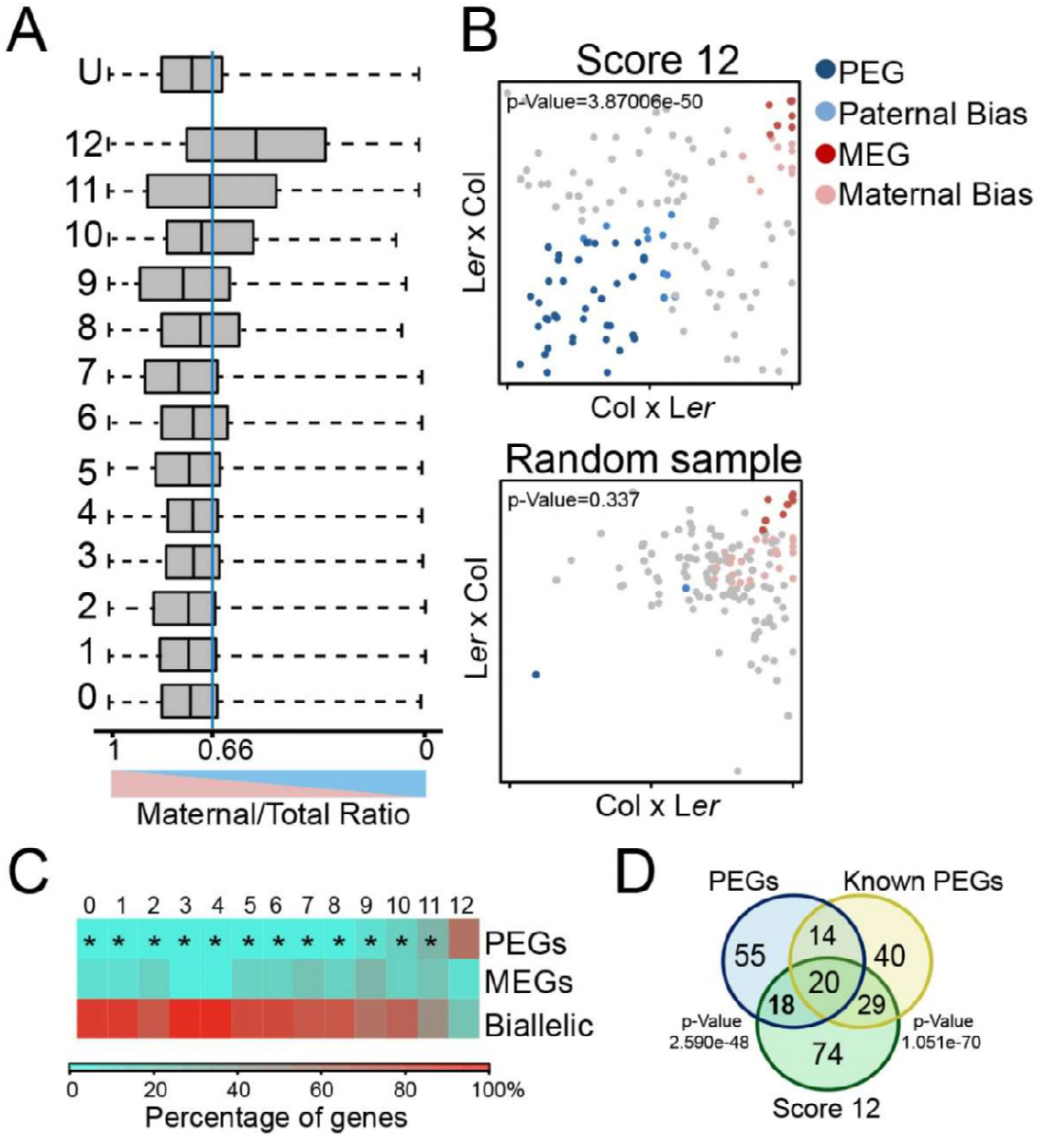


AT1G64660-GFP × WT



WT × *AT1G64660 -GFP*




A
U
12
11
10
9
8
7
6
5
4
3
2
1
0
1
0.66
0
Maternal/Total Ratio
B
Score 12
p-Value=3.87006e-50
Ler x Col
Col x Ler
PEG
Paternal Bias
MEG
Maternal Bias
Random sample
p-Value=0.337
Ler x Col
Col x Ler
C
0 1 2 3 4 5 6 7 8 9 10 11 12
PEGs
MEGs
Biallelic
0
20
40
60
80
100%
Percentage of genes
D
PEGs
Known PEGs
55
14
40
20
18
29
74
p-Value
2.590e-48
p-Value
1.051e-70
Score 12