THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Hybridization alters maternal and paternal genome contributions to early plant embryogenesis

Citation for published version:
Alaniz-Fabián, J, Orozco-Nieto, A, Abreu-Goodger, C & Gillmor, CS 2022, 'Hybridization alters maternal and paternal genome contributions to early plant embryogenesis', *Development*.
https://doi.org/10.1242/dev.201025

Digital Object Identifier (DOI):
10.1242/dev.201025

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Development

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 01. Dec. 2022

Title: Hybridization alters maternal and paternal genome contributions to early plant embryogenesis

Running title: Asymmetric Col/L*er* zygotes

Jaime Alaniz-Fabián[1], Axel Orozco-Nieto[1], Cei Abreu-Goodger[2] and C. Stewart Gillmor[1]*

[1]Langebio, Unidad de Genómica Avanzada, CINVESTAV-IPN, Irapuato, México

[2]Institute of Ecology and Evolution, School of Biological Sciences, University of Edinburgh, Scotland, UK

*Correspondence: stewart.gillmor@cinvestav.mx

Key Words Zygotic genome activation, maternal effect, parent-of-origin, hybrid, Arabidopsis

Summary statement

Parent-of-origin transcriptomes from hybrid Arabidopsis zygotes are more complex than previously reported and do not accurately reflect maternal effects seen in early isogenic embryos.

Abstract

After fertilization, zygotic genome activation (ZGA) results in a transcriptionally competent embryo. Hybrid transcriptome experiments in Arabidopsis have concluded that the maternal and paternal genomes make equal contributions to zygotes and embryos, yet *embryo defective (emb)* mutants in the Columbia (Col) ecotype display early maternal effects. Here we show that hybridization of Col with Landsberg *erecta* (L*er*) or Cape Verde Islands (Cvi) ecotypes decreases maternal effects of *emb* mutants. Reanalysis of Col/L*er* and Col/Cvi transcriptomes confirmed equal parental contributions in Col/Cvi early embryos. By contrast, thousands of genes in Col/L*er* zygotes and 1-cell embryos were biallelic in one cross and monoallelic in the reciprocal cross, with analysis of intron reads pointing to active transcription as responsible for this parent-of-origin bias. Our analysis shows that, contrary to previous conclusions, the maternal and paternal genomes in Col/L*er* zygotes are activated in an asymmetric manner. The decrease in maternal effects in hybrid embryos compared to isogenic Col plus differences in genome activation between Col/Cvi and Col/L*er* suggest that neither of these hybrids accurately reflects general trends of parent-of-origin regulation in Arabidopsis embryogenesis.

49 **Introduction**

50 In animals, early embryonic development is supported by maternally deposited products
51 and does not depend on zygotic transcription (Vastenhouw, Cao, and Lipshitz 2019;
52 O'Farrell 2015). In plants, initial experiments to study the onset of transcription after
53 fertilization in wheat, maize and tobacco found novel transcripts in cDNA libraries of late
54 zygotes compared to the egg cell (Sauter et al.,1998; Okamoto et al., 2005; Sprunck et
55 al., 2005; Ning et al., 2006; Zhao et al., 2011). RNA sequencing experiments comparing
56 gametes with isogenic embryos in maize, rice and Arabidopsis have provided more
57 evidence for transcription of a portion of the zygotic genome within hours after fertilization
58 (haf). In maize zygotes at 12 hours after pollination (hap), 3605 genes were found to be
59 upregulated compared to the egg and sperm (Chen et al., 2017). In rice, comparison of
60 zygote and egg cell transcriptomes found 499 genes upregulated at least two-fold in
61 zygotes at 2.5 hap, 1981 genes at 5 hap, and 2485 genes at 9 hap (Anderson et al.,
62 2017). Arabidopsis zygotes at 14 hap (approximately 6 haf) showed 2625 genes
63 upregulated at least two-fold compared to egg cells, while at 24 hap 2951 genes were
64 upregulated compared to egg cells (Zhao et al., 2019). Thus, transcriptomic analyses of
65 isogenic zygotes of several species show transcriptional activity for about 10% of the
66 genome. Immunostaining experiments against the active form of Pol II have detected
67 signal for active transcription in some experiments on plant zygotes (Niedojadło et al.
68 2012; Kao and Nodine 2019), though not in others (Pillot et al., 2010), perhaps because
69 of differences in the species studied, the developmental stage of the zygotes analyzed,
70 or immunological techniques. In summary, compared to animals, most evidence in plants
71 indicates transcriptional activation of at least a small portion of the genome occurs in the
72 zygote.

73 Initial experiments on maternal and paternal contributions to early isogenic embryos relied
74 on reporter genes and functional analysis of early-acting *embryo defective (emb)* mutants.
75 Assays of maternally and paternally-inherited reporter lines in Arabidopsis found almost
76 30 genes expressed primarily from the maternal allele (Vielle-Calzada et al., 2000; Autran
77 et al., 2011; Del Toro-De León et al., 2014), as well as 10 genes with more equal maternal
78 and paternal expression or function (Weijers et al., 2001; Lukowitz et al., 2004; Xu et al.,

79 2005; Andreuzza et al., 2010; Aw et al., 2010; Ueda et al., 2011; Guo et al., 2016; Yang
80 et al., 2017). In Arabidopsis, the functions of hundreds of *EMBRYO DEFECTIVE (EMB)*
81 genes are required in early embryogenesis (Meinke, 2019). Crosses of 49 *emb/+* mothers
82 to isogenic wild type plants showed that paternal alleles for 9 *EMB* genes immediately
83 complemented lack of maternal function, while paternal alleles for 40 *EMB* genes did not
84 fully complement until 3-5 dap (Del Toro-De León et al., 2014). Consistent with the RNA
85 sequencing experiments mentioned above, these marker gene and functional genetic
86 experiments in isogenic embryos show that zygotic activation varies between genes, with
87 many paternal alleles displaying less activity than maternal alleles in the first 2-3 dap.

88 Maternal and paternal transcripts in embryos can be quantified using single nucleotide
89 polymorphisms generated by crossing polymorphic ecotypes. Initial allele-specific
90 expression (ASE) RNA sequencing (RNAseq) experiments in Arabidopsis found evidence
91 for a transient maternal bias in early Landsberg *erecta* (L*er*) x Columbia-0 (Col) embryos
92 (the 'x' denotes a cross in only one direction was used) (Autran et al., 2011), and for early
93 and equal maternal and paternal contributions in Col/Cape Verde Islands-0 (Cvi) embryos
94 (the '/' denotes an experiment with crosses in both directions) (Nodine and Bartel, 2012).
95 Recent ASE RNAseq experiments of hybrid zygotes have also included data for gametes,
96 allowing for comparison of the egg, sperm, and zygote transcriptomes. In rice, more than
97 97% of 14,049 genes had a strong maternal bias at both 2.5 and 9 hap, with maternal
98 transcripts for most genes found in both the egg and zygote (Anderson et al., 2017). A
99 second experiment with Arabidopsis Col/L*er* hybrid zygotes had a moderate maternal
100 bias at 14 hap (9.9% of 12,746 genes showed maternally biased transcripts), but no
101 maternal bias by 24 hap (Zhao et al., 2019). 60% of genes upregulated at 24 hap were
102 reported as biallelic, suggesting that maternal and paternal alleles of many genes are
103 equally transcribed in late Col/L*er* zygotes (Zhao et al., 2019). The different conclusions
104 of these parent-of-origin studies have been ascribed to profiling of mRNA vs. total RNA
105 (Baroux et al., 2013), different hybrid combinations used for ASE experiments (Baroux et
106 al., 2013; Del Toro-De León et al., 2014 and 2016), and sporophytic maternal
107 contamination (Schon and Nodine, 2017). Differences between Arabidopsis and rice may
108 also reflect variability in the timing of parental contributions between different species. In
109 any case, the extent of maternal and zygotic regulation of early embryogenesis requires

110 further clarification (Baroux and Grossniklaus, 2015; Armenta-Medina and Gillmor, 2019;
111 Dresselhaus and Jürgens, 2021).

112 To determine if varying effects of hybridization between different combinations of
113 ecotypes might explain the divergent conclusions on zygotic genome activation previously
114 reached in Arabidopsis, we explored whether hybridization might alter maternal effects of
115 *emb* mutants. 20 *emb/*+ mutants in Col were crossed with fathers of the Cvi, L*er*,
116 Tsushima-1 (Tsu) and Burren-0 (Bur) ecotypes. *emb/*+ x Cvi and *emb/*+ x L*er* hybrids
117 showed decreased maternal effects compared to isogenic *emb/*+ x Col crosses, while
118 maternal effects in *emb/*+ x Tsu and *emb/*+ x Bur hybrids were indistinguishable from
119 isogenic *emb/*+ x Col. We then used consistent criteria to reanalyze published
120 transcriptome data for Col/Cvi (Nodine and Bartel 2012) and Col/L*er* (Zhao et al. 2019;
121 Zhao et al.2020) zygotes and embryos. Whereas our conclusions were consistent with
122 the original analysis of the Col/Cvi hybrid (Nodine and Bartel, 2012), for the Col/L*er* hybrid
123 we found that plotting of gene reads as the Log_2 fold change (Log_2FC) greatly diminished
124 representation of monoallelic expression in the zygote and basal cell lineage samples,
125 resulting in an incomplete picture of parent-of-origin expression. Genes in the zygote and
126 basal lineage previously proposed to exhibit allele dominance (Zhao et al., 2019 and
127 2020) were more accurately interpreted as showing monoallelic bias in one direction of
128 the cross, and biallelic expression in the other direction. A similar uniparental distribution
129 was observed among reads mapping to intronic regions, strongly suggesting that
130 transcription of the maternal and paternal alleles in the zygote and basal lineages of
131 Col/L*er* zygotes is unequal. Taken together, our findings show that the relative parental
132 contributions to zygotic genome activation can be altered by hybridization, and that
133 zygotic genome activation in the Col/L*er* hybrid occurs in an asymmetric manner.

134 Results

135 Hybridization of Col with Cvi or L*er* decreases maternal effects of *emb* mutants

136 Maternal and paternal gametophytic effects occur when the phenotype of the embryo
137 depends on the genotype of the egg or sperm, respectively (reviewed in Armenta-Medina
138 and Gillmor, 2019). We previously showed that 11 *emb* mutants exhibited transient
139 maternal effects during early embryogenesis, but no paternal effects, and that crossing

140 these *emb* mutants in the Col background with the Cvi and L*er* ecotypes decreased their
141 maternal effects (Del Toro-De León et al., 2014). To extend this analysis, we tested 20
142 additional *emb* mutants for maternal and paternal effects in early embryogenesis (Figure
143 1). Phenotypes of embryos derived from hand-selfed Col *emb/*+ plants, from *emb/*+
144 mothers crossed with wt Col fathers, from wt Col mothers crossed with *emb/*+ fathers,
145 and control crosses of Col mothers with Col fathers, were scored at 2, 3, 5 and 14 dap.
146
147 For all 20 *emb/*+ x Col crosses, maternal effects were observed beginning at 2 or 3 dap,
148 typically decreased by 5 dap, and by 14 dap were absent or almost absent (Figure 1B,
149 Table S1, Supplemental Data 1). To determine if any of the 20 *emb* mutants also had a
150 paternal effect, we used wt Col plants as mothers in crosses with *emb/*+ fathers. Paternal
151 effects were less common than maternal effects (Figure 1B, Table S1). At 2 and/or 3 dap,
152 *nse3*, *pect1*, *iyo* and *miro1* showed transient paternal effects statistically indistinguishable
153 from their transient maternal effect, while *nse1*, *zyg3*, *qqt2*, *gex1* and *fac1* showed
154 transient paternal effects weaker than their transient maternal effects. The remaining 11
155 *emb* mutants showed no paternal effect at all (Figure 1B, Table S1). In summary, maternal
156 effects were seen for all 20 *emb* mutants; maternal and paternal effects of a similar degree
157 were seen for 4 out of 20 *emb* mutants; and a paternal effect weaker than the maternal
158 effect was seen for 5 out of 20 *emb* mutants. These results show that for 16 out of 20
159 *EMB* genes, the maternal allele plays a more important functional role in early
160 embryogenesis than the paternal allele.
161
162 To explore the effect of hybridization on maternal effects of these 20 *emb* mutants, we
163 crossed *emb/*+ mothers in Col with the Cvi, L*er*, Tsu and Bur ecotypes as fathers, and
164 scored mutant phenotypes at 2, 3 and 5 dap (Figure 1C; Table S1). As a control, Col
165 mothers were crossed with Cvi, L*er*, Tsu and Bur and scored for morphologically
166 abnormal phenotypes (Table S1). The percentage of phenotypically abnormal embryos
167 scored for each *emb/*+ hybrid cross at each timepoint, normalized to the percentage of
168 abnormal embryos for the corresponding *emb/*+ x Col cross, is shown in Figure 1C.
169 Compared to isogenic (*emb/*+ x Col) crosses, *emb/*+ x Cvi and *emb/*+ x L*er* hybrids
170 showed significantly smaller maternal effects at 2, 3 and 5 dap, while maternal effects in

171 *emb/+* x Tsu and *emb/+* x Bur hybrids were statistically indistinguishable from isogenic
172 crosses. The decreased maternal effects observed in progeny of *emb/+* x Cvi and *emb/+*
173 x L*er* demonstrate that paternal alleles for many of these 20 *EMB* genes are active earlier
174 in these two hybrids compared to isogenic Col.
175
176 **Reanalysis of hybrid embryo transcriptome datasets reveals previously**
177 **unappreciated parent-of-origin and cell type specific trends in Col/L*er* hybrids** The
178 results above show that maternal effects of 20 *emb* mutants are less in Col x Cvi and Col
179 x L*er* hybrids than in isogenic Col. Nevertheless, the hypothesis that genomic parent-of-
180 origin contributions to zygotes and early embryos can differ between isogenic Col, Col/Cvi
181 and Col/L*er* hybrids needs to be explored further (Del Toro De León et al., 2016; Armenta-
182 Medina and Gillmor, 2019; Dresselhaus and Jürgens, 2021). To compare Col/Cvi and
183 Col/L*er* ASE RNAseq datasets as directly as possible, we comprehensively remapped
184 and reanalyzed data from Nodine and Bartel (2012), Zhao et al., (2019) and Zhao et al.,
185 (2020).

186 ASE RNAseq data from Nodine and Bartel (2012), Zhao et al., (2019) and Zhao et al.,
187 (2020) were mapped to their parent of origin (Table S2) and are plotted in Figure 2 as line
188 graphs in two different ways: as probability density vs $\log_2$FC ratio of the parental
189 ecotypes (as originally plotted in these three studies) (Figure 2A-D); and as probability
190 density vs maternal fraction of reads per gene (see Methods) (Figure 2E-H). As previously
191 shown, $\log_2$FC representation of Col/Cvi datasets produced peaks centered around 0
192 (biallelic) for all stages (Figure 2A) (Nodine and Bartel, 2012), while $\log_2$FC representation
193 of Col/L*er* datasets showed peaks with slight maternal biases in 14 hap zygotes which
194 shifted to 0 (biallelic) in 24 hap zygotes (Figure 2B) and stayed centered around 0 in
195 subsequent cell types and developmental stages (Figure 2C and D) (Zhao et al., 2019;
196 Zhao et al., 2020). Based on $\log_2$FC ratios, Nodine and Bartel (2012) concluded that 1-2,
197 8 and 32-cell Col/Cvi embryos had equal parental contributions to the transcriptome, and
198 Zhao et al., (2019 and 2020) concluded that Col/L*er* zygotes had a maternal transcript
199 bias at 14 hap, shifting to equal parental contributions in 24 hap zygotes and at later
200 stages.

201 Figure 2E-H shows the same datasets, represented as probability density vs maternal
202 fraction of reads per gene. To avoid confounding effects by low-count genes that could
203 easily accumulate in extreme peaks, strong filters were employed and the estimated
204 maternal fraction for each gene was calculated across biological replicates using edgeR
205 (see Methods). For the Col/Cvi dataset, plotting as maternal fraction of reads led to a
206 similar interpretation as for $\log_2$FC plotting: small peaks near 0 and 1 emerged,
207 representing genes that are exclusively paternal or exclusively maternal, but the majority
208 of reads continued to be clustered around 0.5, i.e. genes with equal maternal and paternal
209 contributions (Figure 2E). By contrast, representation of Col/L*er* datasets as maternal
210 fraction of reads had a drastic effect on their interpretation, as a trimodal distribution
211 emerged. 14 hap zygote samples had central peaks located at about 0.65 (Figure 2F),
212 while at 24 hap, this central peak was centered around 0.5 (Figure 2F). At 24 hap, the
213 Col x L*er* central peak was considerably smaller than the corresponding L*er* x Col peak,
214 demonstrating the asymmetric effect of hybridization of these ecotypes. In addition to the
215 central peaks, both Col x L*er* and L*er* x Col zygotes showed large peaks near 0
216 (completely paternal) and 1 (completely maternal) (Figure 2F). In the Col x L*er* cross, 13%
217 (n=12,342) of genes at 14 hap and 24% (n=10,994) of genes at 24 hap showed more
218 than 95% bias towards either parent. In the L*er* x Col cross, 16% (n=12,122) of genes at
219 14 hap and 7% (n=11,000) of genes at 24 hap showed more than 95% bias towards either
220 parent (values derived from Table S3). The strong filters used in our analysis indicated
221 that the maternal and paternal peaks were not caused by-sampling biases of genes with
222 low counts.

223 The trimodal distribution observed in Col/L*er* zygotes and embryos continued in the basal
224 cell produced by asymmetric division of the zygote (Figure 2G), and to a lesser extent in
225 the suspensor of the globular embryo, which is produced by the basal cell (Figure 2H).
226 Interestingly, the apical cell resulting from the asymmetric division of the zygote had an
227 almost monomodal distribution, with most reads clustered around 0.5, and only small
228 peaks at 0 and 1 (Figure 2G). This trend continued in the globular embryo, where the
229 peaks at 0 and 1 had essentially disappeared and almost all reads were clustered around
230 0.5 (equal parental contributions) (Figure 2H). Thus, after the first division of the zygote,
231 the asymmetric transcriptional state of the Col/L*er* zygote is reset in the apical cell, yet

232 continues in the basal cell lineage, decreasing during subsequent rounds of division of
233 the suspensor lineage.

234 The apparent scarcity of genes with strong maternal and paternal transcript bias in the
235 analyses by Zhao et al., (2019 and 2020) can be explained by handling of biological
236 replicates and log_2FC representation of data. Previous analyses of this data (Zhao et al.,
237 2019 and 2020) apparently relied on adding the counts across biological replicates and
238 then calculating log_2FC without properly handling zeroes (Zhao et al., 2019 Table S5 and
239 Zhao et al., 2020 Table S1). Zeros frequently arise due to lack of sampling or sequencing
240 depth and are commonly dealt with by adding a small prior count to the numerator and
241 denominator before calculating the fold-change (Chen et al., 2016), although there are
242 more sophisticated statistical frameworks (Love et al., 2014; Erhard, 2018). The methods
243 of Zhao et al. resulted in a lack of consideration of biological variability across replicates
244 and the generation of infinite log_2FC values when the counts came only from one parent,
245 artificially excluding completely monoallelic genes in their log_2FC plots. Log_2FC
246 representation also causes genes that tend towards unbiased behavior to cluster around
247 0, increasing their apparent frequency; and genes with biased parent-of-origin behaviors
248 to spread out along the X axis, minimizing their apparent frequency (Figure 3). For
249 example, a gene with 1 paternal and 100 maternal reads, and a gene with 1 paternal and
250 1000 maternal reads would both show almost the same maternal fraction (0.99 vs. 0.999),
251 but highly divergent log_2FC values (-6.6 vs. -10). Thus, for the purpose of analyzing and
252 interpreting parent-of-origin contributions on a gene-by-gene basis, we believe that
253 representation of data as maternal fraction of reads is more useful than log_2FC
254 representation.

255 To determine if genes whose transcripts showed an allelic imbalance in Figure 2E-H were
256 due to ecotype-dominant effects, parent-of-origin effects, or a mixture of both, we used
257 the exact test for negative binomial distributions in edgeR to calculate statistically
258 significant deviations from the expected 1:1 ratio, for each direction of the cross, for all
259 samples (see Methods, Table S3). In agreement with Zhao et al., (2019), 16.3% of 13,540
260 genes in 14 hap zygotes showed maternal bias in both directions of the cross; by 24 hap
261 this category had almost completely disappeared (Figure 2M). In contrast to the analysis

262 of Zhao et al., (2019 and 2020) that concluded that 24 hap Col/L*er* zygotes showed equal
263 parental transcriptome contributions, in our analysis, thousands of genes showed
264 maternal or paternal bias in one direction of the cross, and biallelic expression in the other
265 direction, a category that we refer to as unidirectional bias. In 14 hap zygotes, 42.5% of
266 13,540 of genes showed unidirectional bias, while at 24 hap 38.6% of 12,389 genes
267 showed unidirectional bias, indicating that thousands of genes in Col/L*er* zygotes
268 contribute transcripts primarily from the maternal allele or paternal allele (Figure 2M,
269 Table S4). By contrast, less than 3.1% of genes showed Col or L*er* dominant behavior at
270 any stage (Figure 2M, Table S4). It had previously been noted that a significant portion
271 of genes showed monoallelic expression in the basal cell lineage of the Col/L*er* hybrid
272 (Zhao et al. 2020). These genes were previously described as ecotype dominant because
273 their maternal to paternal ratio was found to be negatively correlated between reciprocal
274 crosses. In contrast to the conclusion of Zhao et al. (2019) that 24hap Col/L*er* zygotes
275 show equal parental contributions, our analysis shows that monoallelic expression in
276 zygotes and the basal lineage is common and is primarily caused by genes that are not
277 reciprocally biased towards any ecotype or parent but are rather unidirectionally biased,
278 i.e., that parental contributions to the Col/L*er* zygotic transcriptome vary greatly between
279 genes and are thus unequal.

280 As shown in Figure 2F-H, for a given cell type, a similar proportion of maternally or
281 paternally biased genes exists in both directions of the cross in the zygote and basal cell
282 lineage. Most of these genes are unidirectionally biased (Figure 2M), i.e., they are only
283 significantly biased in one direction. This suggests that regulation of allelic expression in
284 the zygote and embryo is strongly affected by an interaction between ecotype and parent-
285 of-origin effects. Considering that a comparable number of unidirectionally biased genes
286 were discovered across zygote, basal cell and globular-suspensor samples (Table S4),
287 we wondered whether cell identity also plays a role in regulating allele specific expression.
288 To test this, we examined the overlap of unidirectional maternally and paternally biased
289 genes between 14 hap and 24 hap zygotes, the basal cell, and the suspensor, for each
290 direction of the cross (Figure 4A-D; Table S5). This analysis showed that allele bias for
291 most genes is not preserved through developmentally adjacent samples. For example, of
292 4,325 genes with maternal bias in the Col x L*er* 14 hap zygote, only 692 also show

293 maternal bias in the 24 hap zygote (Figure 4A). Of 4,989 genes with maternal bias in the
294 Ler x Col 14 hap zygote, only 506 also show maternal bias in the 24 hap zygote (Figure
295 4B). A similar lack of overlap of paternally biased genes was seen between 14 and 24
296 hap zygotes (Figure 4C and D). Interestingly, while the total number of genes with
297 maternal bias decreases from 14 to 24 hap, the number of paternally biased genes
298 increases from 14 to 24 hap, perhaps reflecting a general increase in transcription of the
299 paternal genome between 14 and 24 hap. Thus, during early development of zygotes and
300 the basal cell lineage of Col/L*er* embryos, allele-specific contributions are continuously
301 restructured in a gene-specific manner. Taken together, these analyses contradict the
302 conclusion that 24 hap zygotes and 1-cell embryos show equal parental transcript
303 contributions (Zhao et al, 2019 and 2020), and show that the biological interactions
304 between ecotypes that regulate parental contributions to Col x L*er* and L*er* x Col
305 transcriptomes have previously been oversimplified.

306 **The use of intronic reads as a proxy for *de novo* transcription suggests non-**
307 **equivalent transcription from parental genomes at early stages** After the first zygotic
308 division, a significant portion of the transcriptome of the basal cell lineage remained
309 monoallelic (Figure 2G and H), even though parent-of-origin transcript bias was not
310 maintained through cell divisions of the zygote and basal lineage (Figure 4A-B). To test
311 whether zygotic transcription of maternal or paternal alleles could explain the biases seen
312 for thousands of gene transcripts in the basal cell lineage (Figure 2F-H, M), we first
313 quantified parent-of-origin transcripts for genes present in the zygote but absent in the
314 egg cell. Because the vast majority of transcripts expressed in the egg cell are also found
315 in the zygote (Zhao et al., 2019), excluding genes with egg cell transcripts removed most
316 zygotic genes, limiting the set of *de novo* parent-of-origin transcripts identified with this
317 approach to a few hundred genes (Figure S1 and Table S6). In 14 hap Col x L*er* zygotes,
318 179 of 251 genes showed paternal bias, with 80 of these having a maternal fraction of
319 less than 0.1. In 24 hap Col x L*er* zygotes, 199 of 326 genes showed paternal bias, with
320 66 having a maternal fraction less than 0.1. While the paternal bias in *de novo* transcripts
321 might indicate a paternal bias in transcription in the zygote, the lack of maternally biased
322 genes could also be due to overlap in transcription of maternal alleles between the egg

323 cell and zygote, limiting the conclusions that can be made from identification of *de novo*
324 expressed genes.

325 As a more comprehensive method for characterizing *de novo* transcription, we then
326 quantified reads mapping to variants in introns. Transcripts carried over from gametes
327 should be spliced, while active *de novo* transcription leads to a small proportion of reads
328 mapping to introns of immature intron-containing transcripts, which serve as a proxy for
329 active transcription and can be detected in regular RNA-Seq experiments (Gaidatzis et
330 al. 2015; Lee et al. 2013). Though reads derived from introns were less abundant than
331 exon reads (as expected), our methods allowed us to identify statistically significant
332 parent-of-origin behavior for intron reads for thousands of genes (Figure 2I-L,N; Table
333 S7, Table S8, Table S9).

334 In the 14 hap zygote samples, intron reads were heavily biased towards maternal alleles
335 (Figure 2J), suggesting that maternal alleles of most genes gain transcriptional
336 competence before paternal alleles, and that active transcription is partially responsible
337 for the maternal bias seen in whole gene reads at this stage (Figure 2F). In the 24 hap
338 zygote, both parental genomes generated intron reads, but primarily at different loci,
339 resulting in strong peaks of maternal and paternal gene transcripts when Col is used as
340 a mother, and a more uniform distribution (with smaller peaks) when L*er* is the mother
341 (Figure 2J). After the first asymmetric division, transcription in the basal lineage followed
342 a pattern similar to the zygote, while transcription in the apical cell showed a strong central
343 peak and was essentially biallelic by the globular embryo stage (Figure 2K, L).

344 Compared to the Zhao et al., (2019 and 2020) datasets, the Nodine and Bartel (2012)
345 dataset had relatively few reads mapping to introns, but some trends were nonetheless
346 evident. Except for the octant stage Cvi x Col transcriptome, all other Col/Cvi datasets
347 showed large peaks for the maternal and paternal fractions, with about 30% of genes
348 showing equal maternal and paternal reads (Figure 2I). Intriguingly, this was in contrast
349 with the distribution seen using reads mapping to variants in whole genes, where these
350 monoallelic tendencies were much less frequent (Figure 2E). Plots for active transcription
351 in Col/L*er* zygotes also showed increased bias towards one parent or the other (Figure
352 2K-L) compared to total reads for the same stages (Figure 2G-H). The difference between

353 active transcription and steady-state transcripts suggests an important role for post-
354 transcriptional regulation in parental mRNA contributions during early embryogenesis. It
355 is also possible that gene expression occurs in waves of active transcription from each
356 genome with allelic ratios partially balancing in the cytoplasm. Disparity between active
357 transcription and cytoplasmic mRNA levels could also be affected by carryover of spliced
358 mRNA from progenitor cells.

359 We performed a statistical analysis to examine whether the trends in intron reads were a
360 mixture of parent of origin and ecotype effects (Table S8), as seen with whole gene reads.
361 In 14 hap zygotes, 60.9% of genes were biallelic, 34.6% showed unidirectional bias and
362 3% were maternal (n=4,804), while in 24 hap zygotes 47.3% were biallelic, 45% showed
363 unidirectional bias and 1% were maternal (n=3,113) (Figure 2N, Table S9). After the first
364 division of the zygote, 97.5% of genes in the apical cell were biallelic (n=1,756) and only
365 1.8% of genes showed unidirectional bias, while in the basal cell 85.1% of genes were
366 biallelic and 14.3% of genes (n=2,037) showed unidirectional bias. Except for maternally
367 biased genes in 14 hap zygotes, which were 16.3% using whole gene reads and 3% using
368 intron reads, most other classes of genes were very similar between whole gene and
369 intron reads, at all stages and in all cell types, suggesting that the maternal reads in the
370 14 hap zygote are primarily inherited from the egg cell, with a smaller contribution from
371 zygotic maternal transcription, and that active transcription is primarily responsible for
372 determining transcript populations in the 24 hap zygote and at later stages. In the Col/Cvi
373 hybrid, despite the marked bias towards uniparental expression (Figure 2I), no statistically
374 significant results were found (Tables S4 and S9). However, the lack of biological
375 replicates for the Col/Cvi dataset and the general scarcity of the reads mapping to
376 polymorphisms in introns makes it likely that this view is not complete. Taken together,
377 this evidence suggests that active transcription is not equivalent for both parental
378 genomes at early stages of embryogenesis, with the ecotypes used and the direction of
379 the cross as key factors that determine allelic transcription.

380 **Discussion**

381 **Maternal effects of *emb* mutants decrease in Col/L*er* and Col/Cvi hybrids** Maternal
382 effects occur when the phenotype of the zygote depends on the genotype of the mother

383 (Armenta-Medina and Gillmor, 2019). 40 out of 49 *embryo defective (emb)* mutants in the
384 Col ecotype were previously shown to display transient maternal effects during
385 embryogenesis, presumably due to early inactivity of the corresponding wt paternal allele
386 (Del Toro De León et al., 2014). In this work we confirmed transient maternal effects for
387 20 additional *emb* mutants in Col and in hybrids of Col/Cvi, Col/Ler, Col/Tsu and Col/Bur.

388 Comparison of maternal effect and transcriptome data for *EMB* genes in isogenic Col
389 (Table S10) shed light on the molecular nature of the maternal effect for *MONOPTEROS*
390 *(MP)* and *GAMETE EXPRESSED PROTEIN 1* (*GEX1)* genes. *MP* transcripts were not
391 detected in the egg cell but were present at low levels in isogenic embryos beginning at
392 the 1 cell stage. Thus, the maternal effects seen for the *mp* mutant are likely to be zygotic
393 maternal effects. By contrast, *GEX1* transcripts were present at very high levels in the
394 egg cell, decreased in the 14 hap zygote, and were almost absent later, suggesting that
395 the maternal effects observed for *gex1* are likely to be maternal gametophytic effects.
396 Transcripts for the other *EMB* genes we tested were present in both egg cells and
397 zygotes, so the maternal effects observed in these isogenic *emb* x Col crosses could be
398 either gametophytic or zygotic.

399 For *emb/*+ x Cvi and *emb/*+ x L*er* crosses, maternal effects were decreased compared to
400 *emb/*+ x Col, suggesting that activation of paternal alleles for these *EMB* genes occurs
401 earlier in Col x Cvi and Col x L*er* hybrids than in isogenic Col. Looking at parent-of-origin
402 transcripts for individual *EMB* genes in Col x L*er* zygotes and 1-cell embryos and Col x
403 Cvi 1-2, 8 and 32 cell embryos (Table S10), all *EMB* genes for which there are data are
404 bi-allelic at all stages, with the exception of *ZYG3* which is paternal at 24 hap and *AtLA1*
405 and *IYO* which are maternal at 24 hap (Table S10). Bi-allelic zygotic and embryo
406 transcriptomes in Col x L*er* and Col x Cvi make sense with the reduced maternal effects
407 for *emb* mutants in these hybrids compared to isogenic Col and are consistent with the
408 conclusions of Nodine and Bartel (2012) and Zhao et al. (2019) that transcriptional
409 activation of paternal alleles in Col/Cvi occurs by the 1-2 cell embryo stage and in Col/L*er*
410 hybrids by the 24 hap zygote stage.

411 It should be noted that *emb/*+ x L*er* and *emb/*+ x Cvi crosses do indeed show transient
412 maternal effects, but of less magnitude than isogenic *emb/*+ x Col and hybrid *emb/*+ x

413 Tsu and *emb/+* x Bur crosses. These maternal effects could originate from maternal
414 transcripts carried over from the egg cell to the zygote and/or zygotic maternal transcripts.
415 At the genome level, Col/L*er* zygotes provide an example of how transient zygotic
416 maternal effects could occur. As seen in Figure 2J, 14 hap Col/L*er* zygotes have a strong
417 maternal bias in intron transcripts. This maternal zygotic transcriptional bias is short lived:
418 after the first division of the zygote, the apical cell shows more equal maternal and
419 paternal intron transcripts while in the basal cell most genes show either a strong maternal
420 or paternal bias (Figure 2K). The increased maternal effects of *emb* mutants in *emb/+* x
421 Col compared to *emb/+* x L*er* suggests that a maternal bias in zygotic transcription may
422 last longer in isogenic Col than in Col/L*er* zygotes. Another possibility is that gametophytic
423 maternal effects are stronger in isogenic Col than in Col/L*er* hybrids. In either case,
424 variation in gene expression and transcript stability between isogenic and hybrid embryos
425 may be due to differences in the timing of similar developmental mechanisms.

426 **Col/L*er* hybrids show strong non-reciprocal parent-of-origin biases in zygotes and**
427 **the basal lineage** Previous studies with the Col/L*er* hybrid found that 66% of informative
428 reads in 14 hap zygotes mapped to the maternal genome, while in 24 hap zygotes and in
429 both apical and basal cells after the first zygotic division, maternal and paternal genome
430 reads were balanced at about 50% each (Zhao et al. 2019; 2020). Genes upregulated at
431 14 hap in the zygote compared to the egg cell were primarily biallelic, while downregulated
432 genes were most often maternally biased (Zhao et al. 2019). Zhao et al. (2019) concluded
433 that a maternal effect exerted through mRNA carryover from the egg cell was present at
434 14 hap but was quickly abolished via mRNA degradation. Despite similar amounts of total
435 reads mapping to the maternal and paternal genomes, 33% of genes in the apical cell
436 and 20% of genes in the basal cell were interpreted as monoallelic (Zhao et al. 2020).
437 The allelic bias for these genes was interpreted as negatively correlated with their bias in
438 the reciprocal cross and thus these imbalances were attributed to allele dominant effects,
439 i.e. genes which are primarily expressed from the Col or L*er* allele, regardless of the
440 parent-of-origin.

441 In our analysis of parent-of-origin contributions in Col/L*er* hybrids, it became clear that
442 while the contributions of the maternal and paternal genomes considered as a whole are

443 indeed similar, the parent-of-origin contributions of thousands of genes vary widely. For
444 example, in 14 hap zygotes 42.5% of genes showed parental bias in only one direction
445 of the cross (Figure 2M), while at subsequent developmental stages most of these genes
446 showed varying parent-of-origin contributions (Figure 4A-D). A similar trend was observed
447 in 24 hap zygotes (Figure 2M and Figure 4A-D). Thus, the assertion that the maternal
448 and paternal genomes make equal contributions to 24 hap zygotes and the subsequent
449 basal cell lineage (Zhao et al., 2019 and 2020) applies only when reads for the entire
450 genome are considered. These observations strongly suggest that, at least in the Col/L*er*
451 hybrid, parental contributions to the early Arabidopsis embryo are dynamic and shaped
452 by several factors, including parent-of-origin, the combination of parental ecotypes, and
453 cell identity.

454 Given the unique epigenetic state of gametes and their dynamic remodeling after
455 fertilization (Okada et al., 2005; Borg et al., 2020; Ingouff et al., 2017; Niedojadło et al.,
456 2012; Ingouff et al., 2010), it would not be unexpected for the maternal and paternal
457 genomes to undergo different or asynchronous reprogramming in the zygote. In isogenic
458 Arabidopsis, maternally encoded histone 3 lysine 9 di-methylation ($H3K9me^2$) as well as
459 various components of the RNA dependent DNA Methylation (RdDM) pathway were
460 shown to repress several paternal marker genes, while maternally encoded components
461 of the chromatin assembly factor complex (CAF1) acted to promote the expression of
462 paternal marker genes (Autran et al., 2011). In addition, POLYCOMB REPRESSIVE
463 COMPLEX 2 (PRC2) has been shown to regulate imprinted expression of several genes
464 in Arabidopsis embryos (Raissig et al., 2013). In the case of hybrids, some Arabidopsis
465 ecotype combinations may have divergent epigenetic configurations in their gametes that
466 could lead to changes in DNA methylation and gene expression due to siRNA-mediated
467 interaction between the two genomes, as has been found in vegetative tissues of
468 Arabidopsis (Groszmann et al., 2011; Greaves et al., 2012). Rice zygotes undergo
469 significant changes in siRNA populations compared to gametes (Li et al., 2020; 2022),
470 and thus varying epigenetic configurations between gametes could potentially alter the
471 epigenetic pathways that may regulate parent-of-origin expression in the embryo, priming
472 or silencing a different set of genes in each ecotype. Likewise, differences in parent-of-
473 origin transcription between developmental stages (Figure 4A-D) could be a consequence

474 of an interaction between reconciliation of epigenetic marks in the gametes and zygote
475 and developmental programs in the embryo.

476 **Parent-of-origin transcript differences between the suspensor and the embryo**
477 **proper in Col/L*er* hybrids** After the first division of the zygote establishes the apical and
478 basal lineages, parent-of-origin and ecotype effects almost disappear in the apical lineage
479 (apical cell and globular embryo proper), which transitions to essentially equal maternal
480 and paternal transcripts for most genes (Figure 2G,H,M). However, strong maternal and
481 paternal transcript peaks persist in the basal cell (Figure 2G) and to a lesser extent in the
482 suspensor of the globular embryo (Figure 2H). In agreement with the analysis of gene
483 reads, analysis of active transcription via intron reads suggests that many genes in the
484 basal lineage are being transcribed mostly from either the maternal or paternal allele,
485 whereas active transcription is mostly biallelic in the apical cell and globular embryo
486 proper (Figure 2K,L). Although thousands of genes in the suspensor showed an allelic
487 ratio far from 1:1 or marked peaks of monoallelic expression, our analysis identified few
488 genes in the suspensor with a statistically significant bias (Figure 2M,N). This may be due
489 to the increased variability in relative parental contributions for suspensor samples
490 between biological replicates in comparison with other tissues (Figure S2), a drawback of
491 the conservative statistical approach we used to call parentally biased genes.

492 Why might asymmetric parent-of-origin expression persist longer in the basal lineage than
493 in the apical lineage? One possibility is that inherent biological differences between the
494 functions of the suspensor and the embryo cause differential parental genome activity in
495 these two tissues. For example, the suspensor is a transient tissue that moves nutrients
496 and hormones from the maternal tissues to the embryo proper (Peng and Sun, 2018),
497 and thus the parental conflict theory would predict a high pressure to imprint genes in the
498 suspensor (Pires and Grossniklaus, 2014). What is clear is that the basal cell and its
499 lineage show much more widespread parental bias in transcript abundance than the
500 apical cell lineage, and that the direction of the cross and ecotypes used strongly affect
501 parental contributions to the embryo transcriptome.

502 **Col/Cvi hybrids lack parent-of-origin and ecotype dominant effects** Col/Cvi
503 transcriptomes showed almost equal parental contributions starting at the 1-2 cell embryo

504 stage, with only about 100 genes that show monoallelic behavior, suggesting quick
505 degradation of maternally inherited products and synchronous activation of both parental
506 genomes (this work; Nodine and Bartel, 2012). Thus, 1-2 cell and later embryos in the
507 Col/Cvi hybrid almost completely lack the variable pattern of parental contributions seen
508 in Col/L*er* zygotes and 1-cell embryos, such as unidirectionally biased transcripts in the
509 zygote and monoallelic behaviors seen in the basal cell and globular suspensor. Though
510 it is possible that sampling of whole 1-2 cell Col/Cvi embryos diluted any monoallelic
511 patterns that might have existed in the basal lineage, the dramatic contrast between
512 Col/Cvi and Col/L*er* hybrids points to intrinsic differences between hybrid combinations in
513 terms of the dynamics of zygotic genome activation (Baroux et al., 2013; Del Toro-De
514 León et al., 2016).

515 One possibility to explain variation in parent-of-origin genome activation in Col/L*er* and
516 Col/Cvi hybrids might be the epigenetic differences between Cvi, L*er* and Col gametes.
517 Throughout male gametogenesis in Col, CpG methylation patterns remain stable (Calarco
518 et al. 2012; Ingouff et al. 2017). However, the major CpG-DNA maintenance
519 methyltransferase *MET1* is barely detected in the mature egg cell of Col (Jullien et al.
520 2012), and the signal of a CpG-methylation fluorescent reporter sharply decreased in the
521 mature egg cell (Ingouff et al. 2017). Taking this into account, it is tempting to speculate
522 that differences in CpG methylation in parental genomes might be partly responsible for
523 the differential remodeling and transcriptional activation after fertilization, which could also
524 explain how an accession like Cvi behaves so differently in these terms, considering that
525 the Cvi ecotype has significantly lower CpG methylation in genes compared with Col and
526 L*er*(Schmitz et al., 2013; Pignatta et al., 2014; Kawakatsu et al., 2016). Compared to Col
527 and L*er*, Cvi also lacks *HISTONE DEACETYLASE 6 (HDA6)* function, shows
528 decondensed chromatin and reduced DNA and H3K9 methylation at chromocenters
529 (Tessadori et al., 2009), all of which are characteristic of active transcription and could
530 impact genome activation in the zygote (Baroux et al., 2013).

531 In summary, our analysis shows that Col/L*er* and Col/Cvi hybrids decrease maternal
532 effects of *emb* mutants compared to isogenic Col, pointing to Col/Tsu and Col/Bur hybrids
533 as better proxies for understanding parent-of-origin behavior in Arabidopsis. Contrary to

534 previous reports, the Col/L*er* hybrid exhibits parent-of-origin transcriptome contributions
535 to the zygote and basal lineage that vary widely between genes and between crosses,
536 demonstrating that on a gene-by-gene basis the maternal and paternal genomes in
537 Col/L*er* hybrids make unequal contributions to the zygote at both 14 and 24 hap. By
538 contrast, 1-2, 8 and 32 cell Col/Cvi hybrid transcriptomes show almost no parent-of-origin
539 behavior. The differing parent-of-origin transcriptome patterns and altered maternal
540 effects of *emb* mutants in Col/Cvi and Col/L*er* hybrids suggest that neither of these hybrid
541 transcriptomes paints an accurate picture of parent-of-origin gene regulation in isogenic
542 Columbia embryos. In the future, differences in zygotic genome activation between
543 hybrids might be leveraged to better understand transcriptional regulation in plant zygotes
544 and embryos. Though mechanisms that regulate parent-of-origin expression in the
545 Arabidopsis embryo at the genome level remain to be described, studies of gene
546 expression in the Arabidopsis endosperm have uncovered multiple epigenetic pathways
547 regulating imprinting (Batista and Köhler, 2020). In embryos of the liverwort *Marchantia*
548 *polymorpha,* Polycomb-mediated histone 3 lysine 27 trimethylation is widely deposited on
549 paternal chromosomes and is responsible for transcriptional silencing of hundreds of
550 paternal alleles during the brief sporophytic phase of this lower plant (Montgomery et al.,
551 2022). It will be interesting to determine whether epigenetic mechanisms for imprinting in
552 the Arabidopsis endosperm or Marchantia embryo also regulate parent-of-origin gene
553 expression in the Arabidopsis embryo.

554 **Methods**

555 **Maternal, paternal and hybridization effects of *emb* mutants** In notation of crosses,
556 the mother is always listed first. In referring to hybrid embryos, 'Col x L*er*' refers to a hybrid
557 cross in one direction, while 'Col/L*er*' refers to hybrid crosses in both directions, i.e. Col x
558 L*er* and L*er* x Col. *emb/*+ crosses were conducted as in Del Toro-De León et al. (2014).
559 Seed stocks used for all crosses are listed in Table S1. For each cross at each timepoint
560 for each *emb* mutant, seed from at least three different siliques from three different plants
561 were examined. Statistical significance was calculated using Fisher's Two-Tailed Exact
562 Test, for $p < 0.05$. An *emb* mutant was considered to have a maternal effect when the
563 proportion of mutants in the *emb/*+ x Col cross was significantly different from the Col x

564 Col control, and a paternal effect was called when the portion of mutants in the Col x *emb*
565 cross was significantly different from the Col x Col control. Maternal and paternal effects
566 for a particular *emb* mutant were labeled as indistinguishable when there was no
567 statistically significant difference between the proportion of mutants observed between
568 the *emb/*+ x Col and Col x *emb/*+crosses (Table S1). For comparison of *emb* maternal
569 effects between isogenic Col and hybrid crosses, mutant frequency for each gene was
570 normalized by adding 0.02 and dividing by the frequency in isogenic Col for that gene
571 plus 0.02. This addition was to stabilize mutants with low penetrance and was based on
572 the mutant frequency observed in wild type Col. A Wilcoxon signed rank test was then
573 used with the null hypothesis that these values were centered at 1 (equivalent to isogenic
574 Col) for each ecotype used for each developmental stage (2, 3, and 5 dap).

575 **Parent-of-origin-specific mapping and counting** Custom genomes for L*er* and Cvi
576 were generated by replacing the variants from each ecotype (1001 Genomes Consortium,
577 2016) in the reference sequence of Col. The fasta files were then each joined with the
578 fasta file from Col. These were used to map the reads. Reads were mapped using STAR
579 V2.7.8.a (Dobin et al. 2013). The following flags were used: --alignIntronMax 900 --
580 alignMatesGapMax 900 --outFilterMultimapNmax 1 --outFilterMultimapScoreRange 0.
581 Two-pass mode was used with all libraries from each study used together for the first
582 pass, to better define the relevant splice junctions.

583 For the parent-of-origin-specific counting, for each hybrid, an annotation file that included
584 only the variants in genes was generated by adjusting the Araport 11 annotation for indels,
585 then each gene span was reduced to only the variant regions generating an entry per
586 variant that overlapped a gene. These entries were then split into 2: one annotated as
587 located in the reference (Col) chromosome and one in the alternative (Cvi or L*er*)
588 chromosome. This annotation was put in the SAF format for use with featureCounts (see
589 documentation for featureCounts).

590 Reads overlapping variants were counted using featureCounts V.2.0.1 with the following
591 flags: -O -p --fraction -F SAF. The fractionated counts option was added to avoid counting
592 multiple times reads that overlapped more than 1 variant.

593 **Parent-of-origin-specific intron read counting** To analyze parental bias in intron reads, 594 an annotation file was generated where only the segments of the genome that contained 595 sequence variants and that are spliced out in all annotated isoforms in Araport 11 were 596 considered. Reads were counted using this file with the same parameters as reads 597 mapping to variants in whole genes.

598 **Transcriptome statistical analysis** To be included in differential expression analysis, 599 genes were required to have at least 20 counts and at least 3 counts per million in all but 600 one of the replicates available or in the only library for each Col/Cvi sample. For each 601 library the sum of parental counts was used to calculate library depths and normalization 602 factors for edgeR. Then, for each library maternal vs paternal counts were used to test 603 for significant deviation of the expected 1:1 ratio using the exact test implemented in 604 edgeR. Dispersion was calculated only among biological replicates with the tagwise 605 option. Since the Col/Cvi samples had no biological replicates, dispersion was set to 0.58 606 (the median value among samples). Fold changes were calculated using edgeR and were 607 converted to maternal read fractions per gene. Estimated maternal fractions, FDR and 608 total variant counts are available in Table S3. The same data for intronic reads is available 609 in Table S8.

610 **Representation of parental contributions** To generate the logarithmic fold change plots 611 from Figure 2, raw allelic counts for genes were filtered in the same way as for differential 612 expression analysis. To replicate results from Zhao et al. (2019 & 2020), logarithmic fold 613 change was then calculated for the sum of counts across replicates. The maternal fraction 614 of reads estimated using edgeR was used for the rest of the plots in Figures 2, 3 and 4. 615 To allow for comparison across samples in Figure 2, the range of the Y axis was set to 616 the maximum in each row. Probability densities for Figure 2 were estimated with the 617 geom_density() function in R, which produces a smoothed version of a frequency 618 histogram, so that areas under all curves are equal to 1. Nodine and Bartel (2012) and 619 Zhao et al., (2019 & 2020) likely used a similar calculation for their density curves, though 620 the method used to produce their curves, and the exact definition, were not specified.

621 **Acknowledgements**

622 Seed of *emb* lines were obtained from the Arabidopsis Biological Resource Center. We
623 are grateful to David Meinke and colleagues for their years of effort assembling and
624 maintaining the *emb* mutant collection, which served as a foundation of this work. We
625 thank Daphné Autran, Sean Cutler and Raju Datla for critical reading and comments.
626

627 **Competing interests**

628 No competing interests declared.
629

630 **Funding**

631 This research was supported by Mexico CONACyT Ciencia Básica [A1-S-34956 to CSG],
632 SEP-CINVESTAV [173 to CSG], and CINVESTAV [institutional funds to CSG].
633

634 **Data availability**

635 All RNAseq data analyzed had been previously published. Col/Cvi Illumina data was from
636 Nodine and Bartel (2012), GEO accession GSE33713; Col/L*er* Illumina data for zygote
637 samples was from Zhao et al., (2019), GSE121003; Col/L*er* Illumina data for one-cell and
638 globular stage embryos was from Zhao et al., (2020), GSE107700.
639

640 **Author Contributions**

641 JA-F and CSG designed research; JA-F and AO-N performed research; CSG and CA-G
642 supervised research; JA-F, AO-N, CA-G and CSG analyzed data; JA-F and CSG wrote
643 the paper; CA-G edited the paper; all authors read and approved the manuscript.
644

645 **References**

646 **The 1001 Genomes Consortium.** (2016). 1,135 Genomes Reveal the Global Pattern of
647 Polymorphism in Arabidopsis thaliana. *Cell* **166**, 481–491.

648 **Anderson, S. N., Johnson, C. S., Chesnut, J., Jones, D. S., Khanday, I.,**
649 **Woodhouse, M., Li, C., Conrad, L. J., Russell, S. D. and Sundaresan, V.** (2017). The
650 Zygotic Transition Is Initiated in Unicellular Plant Zygotes with Asymmetric Activation of
651 Parental Genomes. *Dev Cell* **43**, 349-358.e4.

652 **Andreuzza, S., Li, J., Guitton, A.-E., Faure, J.-E., Casanova, S., Park, J.-S., Choi,**
653 **Y., Chen, Z. and Berger, F.** (2009). DNA LIGASE I exerts a maternal effect on seed
654 development in Arabidopsis thaliana. *Development* **137**, 73–81.

655 **Armenta-Medina, A. and Gillmor, C. S.** (2019). Genetic, molecular and parent-of-
656 origin regulation of early embryogenesis in flowering plants. *Curr Top Dev Biol* **131**,
657 497–543.

658 **Autran, D., Baroux, C., Raissig, M. T., Lenormand, T., Wittig, M., Grob, S., Steimer,**
659 **A., Barann, M., Klostermeier, U. C., Leblanc, O., et al.** (2011). Maternal Epigenetic
660 Pathways Control Parental Contributions to Arabidopsis Early Embryogenesis. *Cell* **145**,
661 707–719.

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

665 **Baroux, C., Autran, D., Raissig, M. T., Grimanelli, D. and Grossniklaus, U.** (2013).
666 Parental contributions to the transcriptome of early plant embryos. *Curr Opin Genet Dev*
667 **23**, 72–74.

668 **Baroux, C. and Grossniklaus, U.** (2015). Chapter Ten The Maternal-to-Zygotic
669 Transition in Flowering Plants Evidence, Mechanisms, and Plasticity. *Curr Top Dev Biol*
670 **113**, 351–371.

671 **Batista, R. A. and Köhler, C.** (2020). Genomic imprinting in plants—revisiting existing
672 models. *Gene Dev* **34**, 24–36.

673 **Borg, M., Jacob, Y., Susaki, D., LeBlanc, C., Buendía, D., Axelsson, E.,**
674 **Kawashima, T., Voigt, P., Boavida, L., Becker, J., et al.** (2020). Targeted
675 reprogramming of H3K27me3 resets epigenetic memory in plant paternal chromatin.
676 *Nat Cell Biol* **22**, 621–629.

677 **Calarco, J. P., Borges, F., Donoghue, M. T. A., Van Ex, F., Jullien, P. E., Lopes, T.,**
678 **Gardner, R., Berger, F., Feijó, J. A., Becker, J. D., et al.** (2012). Reprogramming of

679 DNA Methylation in Pollen Guides Epigenetic Inheritance via Small RNA. *Cell* **151**, 194–
680 205.

681 **Chen, Y., Lun, A. T. L. and Smyth, G. K.** (2016). From reads to genes to pathways:
682 differential expression analysis of RNA-Seq experiments using Rsubread and the
683 edgeR quasi-likelihood pipeline. *F1000research* **5**, 1438.

684 **Chen, J., Strieder, N., Krohn, N. G., Cyprys, P., Sprunck, S., Engelmann, J. C. and**
685 **Dresselhaus, T.** (2017). Zygotic Genome Activation Occurs Shortly after Fertilization in
686 Maize. *Plant Cell* **29**, 2106–2125.

687 **Del Toro-De León, G., García-Aguilar, M. and Gillmor, C. S.** (2014). Non-equivalent
688 contributions of maternal and paternal genomes to early plant embryogenesis. *Nature*
689 **514**, 624–627.

690 **Del Toro-De León, G., Lepe-Soltero, D. and Gillmor, C. S.** (2016). Zygotic genome
691 activation in isogenic and hybrid plant embryos. *Curr Opin Plant Biol* **29**, 148–153.

692 **Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P.,**
693 **Chaisson, M. and Gingeras, T. R.** (2013). STAR: ultrafast universal RNA-seq aligner.
694 *Bioinformatics* **29**, 15–21.

695 **Dresselhaus, T. and Jürgens, G.** (2021). Comparative Embryogenesis in
696 Angiosperms: Activation and Patterning of Embryonic Cell Lineages. *Annu Rev Plant*
697 *Biol* **72**, 1–36.

698 **Erhard, F.** (2018). Estimating pseudocounts and fold changes for digital expression
699 measurements. *Bioinformatics* **34**, 4054–4063.

700 **Gaidatzis, D., Burger, L., Florescu, M. and Stadler, M. B.** (2015). Analysis of intronic
701 and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional
702 regulation. *Nat Biotechnol* **33**, 722–729.

703 **Greaves, I. K., Groszmann, M., Ying, H., Taylor, J. M., Peacock, W. J. and Dennis,**
704 **E. S.** (2012). Trans Chromosomal Methylation in Arabidopsis hybrids. *Proc National*
705 *Acad Sci* **109**, 3570–3575.

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

709 **Guo, L., Jiang, L., Zhang, Y., Lu, X., Xie, Q., Weijers, D. and Liu, C.** (2016). The
710 anaphase-promoting complex initiates zygote division in Arabidopsis through
711 degradation of cyclin B1. *Plant J* **86**, 161–174.

712 **Ingouff, M., Rademacher, S., Holec, S., Šoljić, L., Xin, N., Readshaw, A., Foo, S. H.,**
713 **Lahouze, B., Sprunck, S. and Berger, F.** (2010). Zygotic Resetting of the HISTONE 3
714 Variant Repertoire Participates in Epigenetic Reprogramming in Arabidopsis. *Curr Biol*
715 **20**, 2137–2143.

716 **Ingouff, M., Selles, B., Michaud, C., Vu, T. M., Berger, F., Schorn, A. J., Autran, D.,**
717 **Durme, M. V., Nowack, M. K., Martienssen, R. A., et al.** (2017). Live-cell analysis of
718 DNA methylation during sexual reproduction in Arabidopsis reveals context and sex-
719 specific dynamics controlled by noncanonical RdDM. *Gene Dev* **31**, 72–83.

720 **Jullien, P. E., Susaki, D., Yelagandula, R., Higashiyama, T. and Berger, F.** (2012).
721 DNA Methylation Dynamics during Sexual Reproduction in Arabidopsis thaliana. *Curr*
722 *Biol* **22**, 1825–1830.

723 **Kao, P. and Nodine, M. D.** (2019). Transcriptional Activation of Arabidopsis Zygotes Is
724 Required for Initial Cell Divisions. *Sci Rep* **9**, 17159.

725 **Kawakatsu, T., Huang, S. C., Jupe, F., Sasaki, E., Schmitz, R. J., Urich, M. A.,**
726 **Castanon, R., Nery, J. R., Barragan, C., He, Y., et al.** (2016). Epigenomic Diversity in
727 a Global Collection of Arabidopsis thaliana Accessions. *Cell* **166**, 492–505.

728 **Lee, M. T., Bonneau, A. R., Takacs, C. M., Bazzini, A. A., DiVito, K. R., Fleming, E.**
729 **S. and Giraldez, A. J.** (2013). Nanog, Pou5f1 and SoxB1 activate zygotic gene
730 expression during the maternal-to-zygotic transition. *Nature* **503**, 360–364.

731 **Li, C., Gent, J. I., Xu, H., Fu, H., Russell, S. D. and Sundaresan, V.** (2022). Resetting
732 of the 24-nt siRNA landscape in rice zygotes. *Genome Res* **32**, 309–323.

733 **Li, C., Xu, H., Fu, F.-F., Russell, S. D., Sundaresan, V. and Gent, J. I.** (2020).
734 Genome-wide redistribution of 24-nt siRNAs in rice gametes. *Genome Res* **30**, 173–
735 184.

736 **Love, M. I., Huber, W. and Anders, S.** (2014). Moderated estimation of fold change
737 and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550.

738 **Lukowitz, W., Roeder, A., Parmenter, D. and Somerville, C.** (2004). A MAPKK
739 Kinase Gene Regulates Extra-Embryonic Cell Fate in Arabidopsis. *Cell* **116**, 109–119.

740 **Meinke, D. W.** (2020). Genome-wide identification of EMBRYO-DEFECTIVE (EMB)
741 genes required for growth and development in Arabidopsis. *New Phytol* **226**, 306–325.

742 **Montgomery, S. A., Hisanaga, T., Wang, N., Axelsson, E., Akimcheva, S., Sramek,**
743 **M., Liu, C. and Berger, F.** (2022). Polycomb-mediated repression of paternal
744 chromosomes maintains haploid dosage in diploid embryos of Marchantia. *Elife* **11**,
745 e79258.
746

747 **Niedojadło, K., Pięciński, S., Smoliński, D. J. and Bednarska-Kozakiewicz, E.**
748 (2012). Transcriptional activity of Hyacinthus orientalis L. female gametophyte cells
749 before and after fertilization. *Planta* **236**, 153–169.

750 **Ning, J., Peng, X.-B., Qu, L.-H., Xin, H.-P., Yan, T.-T. and Sun, M.** (2006). Differential
751 gene expression in egg cells and zygotes suggests that the transcriptome is restructed
752 before the first zygotic division in tobacco. *Febs Lett* **580**, 1747–1752.

753 **Nodine, M. D. and Bartel, D. P.** (2012). Maternal and paternal genomes contribute
754 equally to the transcriptome of early plant embryos. *Nature* **482**, 94–97.

755 **O'Farrell, P. H.** (2015). Growing an Embryo from a Single Cell: A Hurdle in Animal Life.
756 *Csh Perspect Biol* **7**, a019042.

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

760 **Okamoto, T., Scholten, S., Lörz, H. and Kranz, E.** (2005). Identification of Genes that
761 are Up- or Down-regulated in the Apical or Basal Cell of Maize Two-celled Embryos and
762 Monitoring their Expression During Zygote Development by a Cell Manipulation- and
763 PCR-based Approach. *Plant Cell Physiol* **46**, 332–338.

764 **Peng, X. and Sun, M.-X.** (2018). The suspensor as a model system to study the
765 mechanism of cell fate specification during early embryogenesis. *Plant Reprod* **31**, 59–
766 65.

767 **Pignatta, D., Erdmann, R. M., Scheer, E., Picard, C. L., Bell, G. W. and Gehring, M.**
768 (2014). Natural epigenetic polymorphisms lead to intraspecific variation in Arabidopsis
769 gene imprinting. *Elife* **3**, e03198.

770 **Pillot, M., Baroux, C., Vazquez, M. A., Autran, D., Leblanc, O., Vielle-Calzada, J. P.,**
771 **Grossniklaus, U. and Grimanelli, D.** (2010). Embryo and Endosperm Inherit Distinct
772 Chromatin and Transcriptional States from the Female Gametes in Arabidopsis. *Plant*
773 *Cell* **22**, 307–320.

774 **Pires, N. D. and Grossniklaus, U.** (2014). Different yet similar: evolution of imprinting
775 in flowering plants and mammals. *F1000prime Reports* **6**, 63.

776 **Raissig, M. T., Bemer, M., Baroux, C. and Grossniklaus, U.** (2013). Genomic
777 Imprinting in the Arabidopsis Embryo Is Partly Regulated by PRC2. *Plos Genet* **9**,
778 e1003862.

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

782 **Schmitz, R. J., Schultz, M. D., Urich, M. A., Nery, J. R., Pelizzola, M., Libiger, O.,**
783 **Alix, A., McCosh, R. B., Chen, H., Schork, N. J., et al.** (2013). Patterns of population
784 epigenomic diversity. *Nature* **495**, 193–198.

785 **Schon, M. A. and Nodine, M. D.** (2017). Widespread Contamination of Arabidopsis
786 Embryo and Endosperm Transcriptome Data Sets. *Plant Cell* **29**, 608–617.

787 **Sprunck, S., Baumann, U., Edwards, K., Langridge, P. and Dresselhaus, T.** (2005).
788 The transcript composition of egg cells changes significantly following fertilization in
789 wheat (*Triticum aestivum L.*). *Plant J* **41**, 660–672.

790 **Tessadori, F., Zanten, M. van, Pavlova, P., Clifton, R., Pontvianne, F., Snoek, L. B.,**
791 **Millenaar, F. F., Schulkes, R. K., Driel, R. van, Voesenek, L. A. C. J., et al.** (2009).
792 PHYTOCHROME B and HISTONE DEACETYLASE 6 Control Light-Induced Chromatin
793 Compaction in Arabidopsis thaliana. *Plos Genet* **5**, e1000638.

794 **Ueda, M., Zhang, Z. and Laux, T.** (2011). Transcriptional Activation of Arabidopsis Axis
795 Patterning Genes WOX8/9 Links Zygote Polarity to Embryo Development. *Dev Cell* **20**,
796 264–270.

797 **Vastenhouw, N. L., Cao, W. X. and Lipshitz, H. D.** (2019). The maternal-to-zygotic
798 transition revisited. *Development***146**, dev161471.

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

801 **Weijers, D., Geldner, N., Offringa, R. and Jürgens, G.** (2001). Early paternal gene
802 activity in Arabidopsis. *Nature* **414**, 709–710.

803 **Xu, J., Zhang, H., Xie, C., Xue, H., Dijkhuis, P. and Liu, C.** (2005). EMBRYONIC
804 FACTOR 1 encodes an AMP deaminase and is essential for the zygote to embryo
805 transition in Arabidopsis. *Plant J* **42**, 743–758.

806 **Yang, K., Guo, L., Hou, X., Gong, H. and Liu, C.** (2017). ZYGOTE-ARREST 3 that
807 encodes the tRNA ligase is essential for zygote division in Arabidopsis. *J Integr Plant*
808 *Biol* **59**, 680–692.

809 **Zhao, J., Xin, H., Qu, L., Ning, J., Peng, X., Yan, T., Ma, L., Li, S. and Sun, M.**
810 (2011). Dynamic changes of transcript profiles after fertilization are associated with de
811 novo transcription and maternal elimination in tobacco zygote, and mark the onset of
812 the maternal-to-zygotic transition. *Plant J* **65**, 131–145.

813 **Zhao, P., Zhou, X., Shen, K., Liu, Z., Cheng, T., Liu, D., Cheng, Y., Peng, X. and**
814 **Sun, M.** (2019). Two-Step Maternal-to-Zygotic Transition with Two-Phase Parental
815 Genome Contributions. *Dev Cell* **49**, 882-893.e5.

816 **Zhao, P., Zhou, X., Zheng, Y., Ren, Y. and Sun, M.** (2020). Equal parental contribution
817 to the transcriptome is not equal control of embryogenesis. *Nat Plants* **6**, 1354–1364.

818

819

820

821

822

823

824

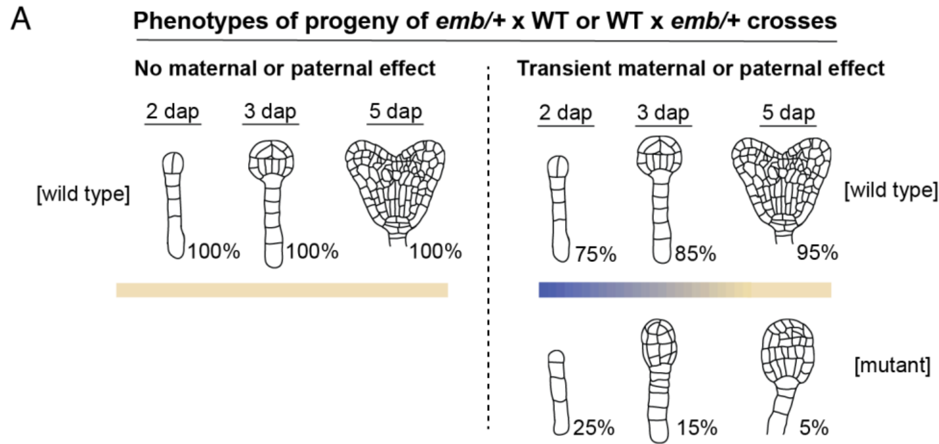
825

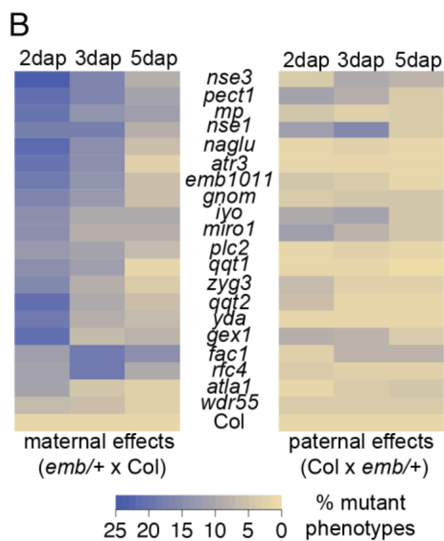
826

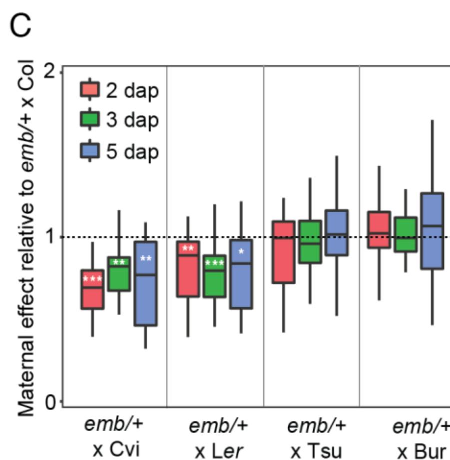
827

828

829

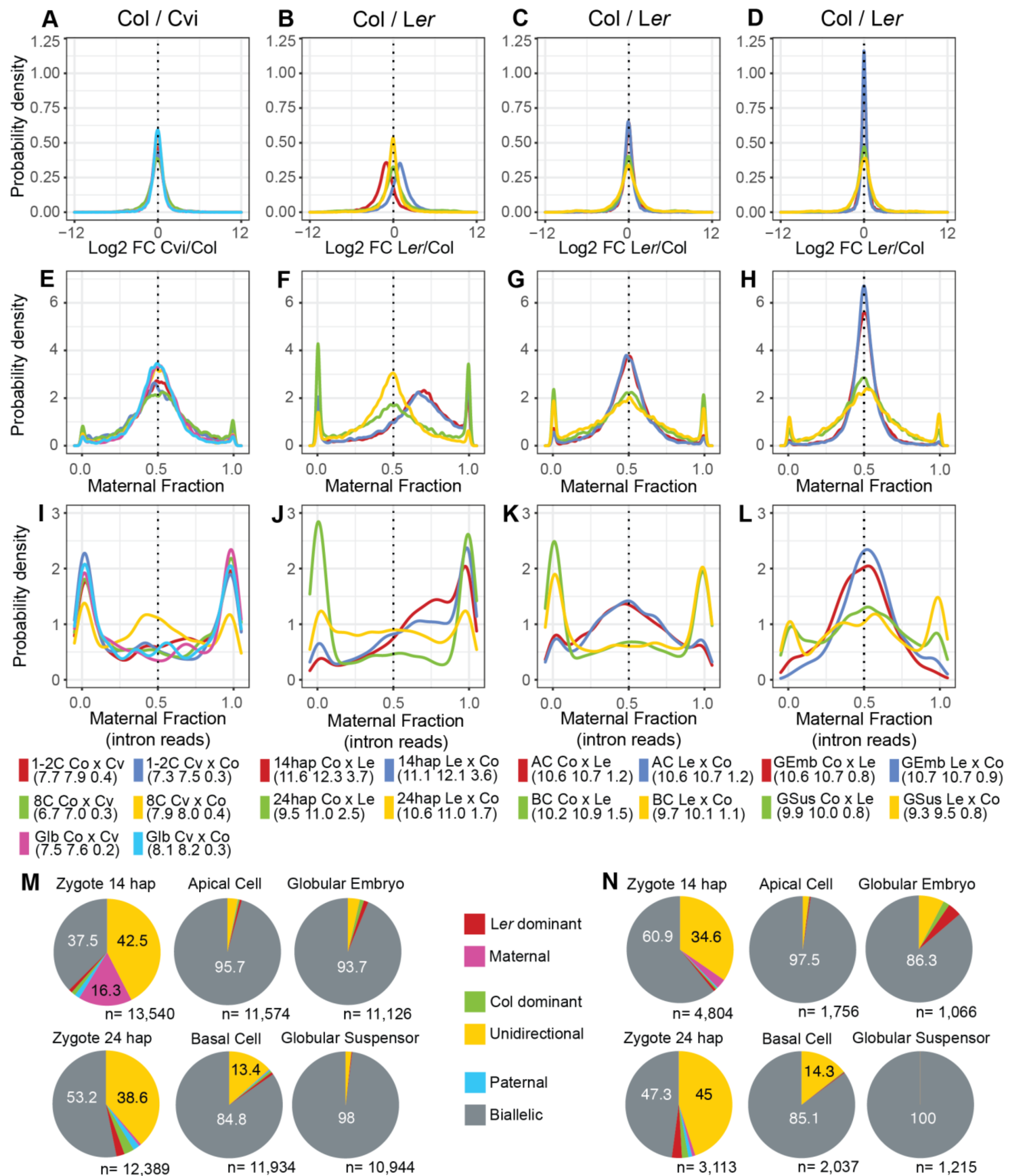






830

831 **Figure 1 Maternal effects of *embryo defective (emb)* mutants in Col are diminished by hybridization**
832 **with Cvi and L*er*** (A) Schematic representations of results for crosses to test maternal and paternal effects
833 of *emb* mutants. Left, in the absence of maternal and paternal effects, all progeny from *emb/+* x WT and
834 WT x *emb/+* crosses have wild type phenotypes (beige bar indicates no mutant phenotypes were observed).
835 Right, example of penetrance of mutant phenotypes when maternal or paternal effects are seen (bar that
836 is purple on the left and gradually turns to beige indicates that mutant phenotypes were initially observed).
837 (B) Maternal and paternal effects observed in embryos from *emb/+* x Col and Col x *emb/+* crosses at 2, 3,
838 and 5 dap. (C) Maternal effects of *emb/+* crosses to Cvi, L*er*, Tsu and Bur ecotypes (hybrid crosses),
839 normalized to the maternal effect of the corresponding *emb/+* x Col cross (isogenic cross). Upper and lower
840 lines of box plots represent the first and third quartiles with the middle line indicating median. Cross data
841 are derived from at least three different siliques from three different plants, with n>90 embryos scored for
842 each cross and timepoint. Hybrid crosses which show a statistical difference for the Wilcoxon signed rank
843 test compared to isogenic crosses are marked with an asterisk; *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$. See
844 Table S1 for full information and Supplemental Data 1 for phenotypes scored for *emb* mutants.



845

846

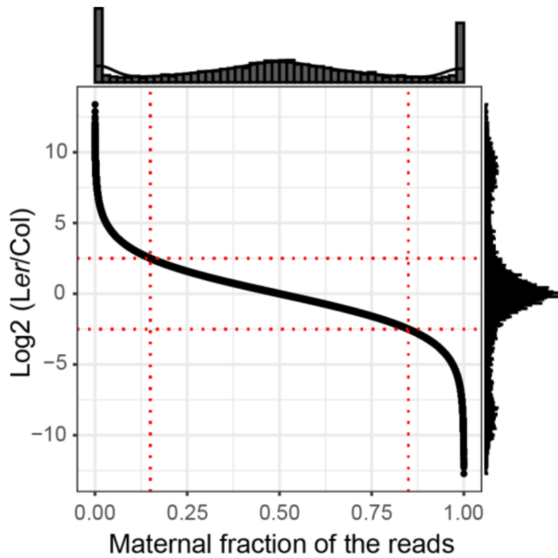
847 **Figure 2 Asymmetric zygotic genome activation in Col/L*er* hybrids** Probability density
848 plots of parent-of-origin reads from Col/Cvi (Nodine and Bartel, 2012) (A, E, I) and Col/L*er*

849 (Zhao et al., 2019 and 2020) (B-D; F-H; J-L) datasets. Probability density is the probability
850 of a gene to have a particular range of values on the x-axis and is calculated so that the
851 total area under each curve is 1 (see Methods). (A-D) data for gene reads represented
852 as probability density vs. Log_2fold change (Log_2FC). (E-H) data for gene reads represented
853 as probability density vs maternal fraction of reads. (I-L) data for intron reads represented
854 as probability density vs maternal fraction of reads. Y-axis values for E-H and I-L are
855 equivalent, while the range of the Y axis in E-H and I-L was determined according to the
856 maximum value in the row. Cross directions and embryo stages are listed below the
857 probability density plots. Ecotypes are abbreviated as Co, Col; Cv, Cvi; Le, L*er*. Embryo
858 stages are abbreviated as 1-2C, 1-2 cell embryo; 8C, 8-cell embryo; Glb, globular embryo;
859 14hap, 14 hap zygote; 24hap, 24 hap zygote; AC, apical cell of 1-cell embryo; BC, basal
860 cell of 1-cell embryo; GEmb, globular embryo proper; GSus, suspensor of globular
861 embryo. The number of genes represented in each line plot for each cross for each
862 dataset is listed in parentheses below panels I-L. The first number refers to the Log_2FC
863 plot, the second to maternal fraction of reads, and the third to maternal fraction of intron
864 reads (all numbers are x1000). (M and N) percent of genes in each statistically significant
865 category of parental bias using reads mapping to genes (M) and introns (N) for Col/L*er*
866 datasets. Statistically significant parental bias categories were calculated using the exact
867 test from edgeR with a cutoff of FDR<0.05. The number of genes represented in each pie
868 chart is represented below the chart. L*er* dominant, genes expressed predominantly from
869 L*er* allele; Col dominant, genes expressed predominantly from Col allele; Paternal, genes
870 expressed primarily from paternal allele; Maternal, genes expressed primarily from
871 maternal allele; Unidirectional, genes that are not statistically significant in one direction
872 of the cross but show maternal or paternal bias in the other direction; Biallelic, genes that
873 show no statistically significant parent-of-origin bias.

874

875
876
877
878
879

880
881
882
883



884
885

886 **Figure 3 Parent-of-origin trends are masked by Log_2 Fold Change representation**

887 Parental bias for genes in 24 hap zygote Col x L*er* transcriptome, represented as
888 maternal fraction of reads (x axis) and Log_2FC (L*er*/Col) (y axis). Compared to maternal
889 fraction of reads representation, Log_2FC (L*er*/Col) concentrates the biallelic observations
890 (here considered as genes with 0.15 to 0.85 maternal fraction, red vertical dotted lines),
891 within a range of -2.5 to 2.5 (red horizontal dotted lines), while spreading the distribution
892 of biased observations (here considered as genes with less than 0.15 or more than 0.85
893 maternal fraction) over a much wider range (± 2.5 to ± ∞) (y axis on right). When plotting
894 the frequency distribution using marginal histograms, biased behavior appears much
895 less frequent using logarithmic representation (right) than when using maternal fraction
896 (top). Maternal fraction of the reads and Log_2FC were calculated using edgeR, for
897 n=10.994 genes.

898

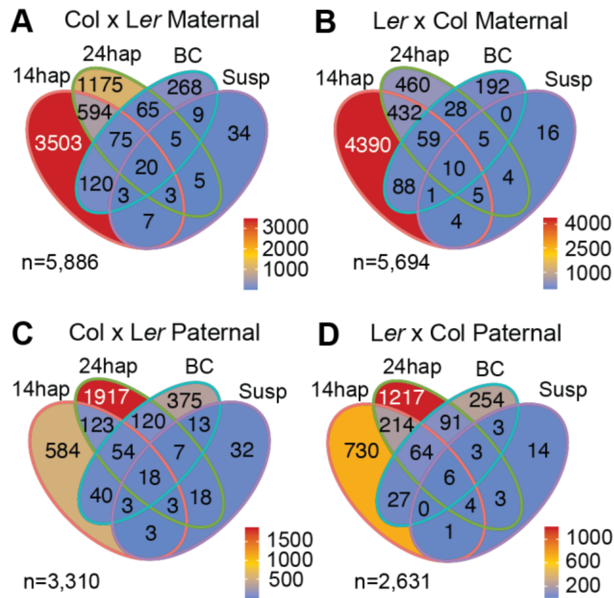


Figure 4 Parent-of-origin behavior at successive developmental stages for maternally and paternally biased genes Venn diagrams showing data for maternally biased genes in Col x L*er* (A) and L*er* x Col (B) crosses, and for paternally biased genes in Col x L*er* (C) and L*er* x Col (D) crosses, for 14 hap zygotes, 14 hap; 24 hap zygotes, 24 hap; basal cells after the first division of the zygote, BC; and globular stage suspensors, Susp. The number of genes in each category is show within the diagram. The color key for the number of genes in each category, as well as the total number of genes represented in each Venn diagram, are shown below each diagram. The exact test from edgeR was used to call parental bias with an FDR of <0.05. See Table S5 for complete gene lists and p-values.