ARTICLES
https://doi.org/10.1038/s41477-020-00793-x

nature plants

Equal parental contribution to the transcriptome is not equal control of embryogenesis

Peng Zhao[1,3]✉, Xuemei Zhou[1,2,3], Yifan Zheng[1], Yanru Ren[1] and Meng-xiang Sun[1]✉

In animals, early embryogenesis is maternally controlled, whereas in plants, parents contribute equally to the proembryo transcriptome. Thus, the question remains whether equivalent parental contribution to the transcriptome of the early proembryo means equal control of early embryogenesis. Here, on the basis of cell-lineage-specific and allele-specific transcriptome analysis, we reveal that paternal and maternal genomes contribute equally to the transcriptomes of both the apical cell lineage and the basal cell lineage of early proembryos. However, a strong maternal effect on basal cell lineage development was found, indicating that equal parental contribution to the transcriptome is not necessarily coupled with equivalent parental control of proembryonic development. Parental contributions to embryogenesis therefore cannot be concluded solely on the basis of the ratio of paternal/maternal transcripts. Furthermore, we demonstrate that parent-of-origin genes display developmental-stage-dependent and cell-lineage-dependent allelic expression patterns. These findings will facilitate the investigation of specific parental roles in specific processes of early embryogenesis.

The maternal-to-zygotic transition (MZT) is the first major transition in the life cycle of both animals and plants. The MZT consists of two integrated processes: maternal transcript clearance and zygotic genome activation (ZGA)[1–3]. In animals, early embryogenesis is controlled by maternal factors deposited in egg cells, whereas ZGA occurs much later[4,5]. However, there are conflicting reports on plant parental contributions to the transcriptome of early proembryos[6,7], which have been explained by the contamination of maternal tissues[8] or the effects of hybridization on gene expression[9,10]. We recently showed that both maternal and paternal genomes are activated at the zygotic stage and contribute equally to the transcriptome of the elongated zygote and proembryos[3]. However, whether equivalent parental contributions to the transcriptome of the proembryo represent equal functional contributions to proembryo development remains in question.

In animals, both daughter cells of the zygote contribute to the mature embryo. However, in higher plants, the two daughter cells of the zygote contribute differently to the embryo. The smaller apical cell (AC) divides to form the major parts of the mature embryo, including the apical–basal axis and the radial pattern. The larger basal cell (BC) usually undergoes limited divisions to form a suspensor composed of several cells. Only the uppermost suspensor cell differentiates into an embryo hypophysis, and the other suspensor cells degenerate via programmed cell death[11–13]. In our previous work, we found a maternal effect on the development of the BC lineage (BCL). The morphology and cell number of the hybrid embryo suspensor were similar to those of the maternal line[14], suggesting maternal control of suspensor development. We thus compared the transcriptomes and developmental processes between the AC lineage (ACL) and the BCL to determine the following: (1) whether the parental contributions to the transcriptomes of the BCL differ from those of the ACL, and (2) whether equivalent parental contributions to the transcriptome represent equal parental control of early embryogenesis.

In this study, we investigated the parental contributions to the transcriptome and the development of the ACL and BCL in early proembryos. We performed cell-lineage-specific and allele-specific transcriptome analyses and careful observations of the suspensor morphology and structure in hybrid embryos from reciprocal crosses between different *Arabidopsis* ecotypes. On the basis of these results, we demonstrated that the parental genome contributed equally to both ACL and BCL transcriptomes and that there was an obvious maternal effect on BCL development. This indicates that the equivalent parental contributions to the transcriptome are not necessarily coupled with equal parental control of ACL and BCL specification. In addition, we identified more parent-of-origin genes in the BCL than in the ACL at corresponding stages, which displayed developmental-stage-dependent and cell-lineage-dependent allelic expression patterns. These genes may be valuable clues to understanding the parental roles of cell fate determination and the regulation of distinct developmental pathways of the ACL and BCL.

Results

Construction of the allele-specific transcriptomes of the ACL and BCL in early proembryos. To investigate the parental contributions to the ACL and BCL, we sequenced the transcriptomes of the ACL and BCL in 1-cell and 32-cell embryos (Fig. 1a) derived from reciprocal crosses between *Arabidopsis* Columbia-0 (Col-0) and Landsberg erecta (Ler). Three independent biological replicates were carried out for each hybrid cell type from the reciprocal crosses. Thus, 24 sequencing libraries were constructed and sequenced at sufficient depth for the *Arabidopsis* transcriptome, generating over 18 million read pairs for each library (Supplementary Table 1). Overall, the transcriptomes from the same cell type derived from the reciprocal crosses and different biological replicates were highly correlated ($r \geq 0.91$; Extended Data Fig. 1). All transcriptome datasets were found to be free of RNA contamination from the endosperm and the seed coat using publicly available software (Extended Data Fig. 2) (https://github.com/Gregor-Mendel-Institute/tissue-enrichmenttest)[8]. After the RNA sequencing quality assessment, reads covering single-nucleotide polymorphisms (SNPs) were

[1]State Key Laboratory of Hybrid Rice, College of Life Sciences, Wuhan University, Wuhan, China. [2]College of Life Sciences, South-Central University for Nationalities, Wuhan, China. [3]These authors contributed equally: Peng Zhao, Xuemei Zhou. ✉e-mail: pzhao2000@whu.edu.cn; mxsun@whu.edu.cn

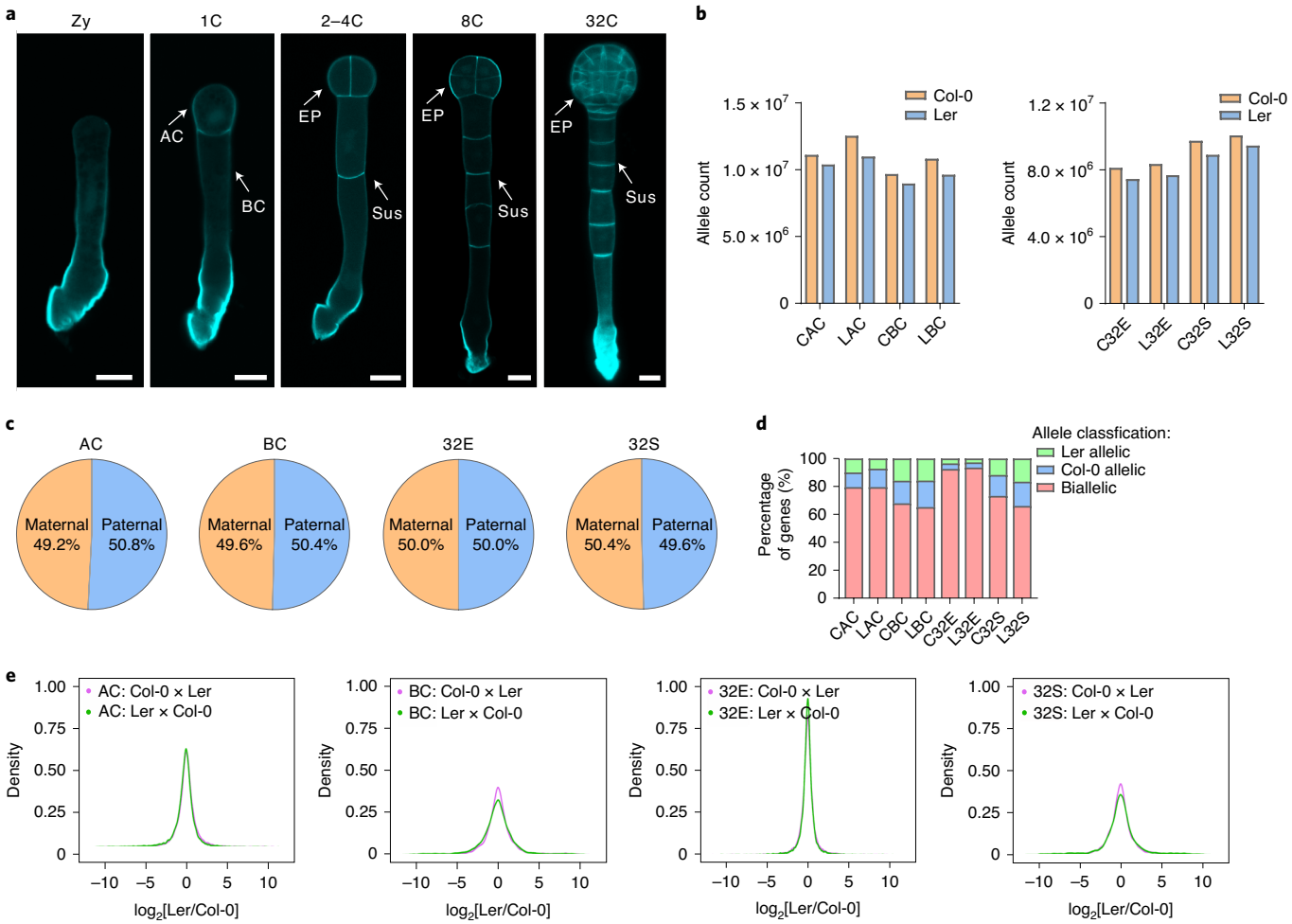

Fig. 1 | Parental contributions to the transcriptomes of ACLs and BCLs of early proembryos. a, A diagram showing early embryo development in *Arabidopsis* ($n=100$). Scale bars, 10 μm. **b**, Numbers of Col-0- and Ler-derived reads in the transcriptomes of ACLs and BCLs of early proembryos. **c**, Percentages of parental RNA-seq reads in the transcriptomes of ACLs and BCLs of early proembryos. The percentages of parental contributions to the transcriptomes are the comprehensive results of three independent biological replicates in the reciprocal crosses between Col-0 and Ler. **d**, Percentages of genes with Col-0 monoallelic, Ler monoallelic and biallelic expression patterns in each cell type derived from the reciprocal crosses. Each datum represents the mean of three independent biological replicates. **e**, Distribution of the ratio of Ler/Col-0 ($\log_2$ value) for genes with ≥10 SNP-overlapping reads in the transcriptomes of ACLs and BCLs. Zy, zygote; 1C, 1-cell embryo; 2–4C, 2- to 4-cell embryo; 8C, 8-cell embryo; 32C, 32-cell embryo; EP, embryo proper; Sus, suspensor; CAC, hybrid AC from the cross between Col-0 and Ler; LAC, hybrid AC from the cross between Ler and Col-0; similar abbreviations are used for CBC, LBC, C32E, L32E, C32S and L32S.

extracted, and transcripts with at least ten reads that overlapped SNPs in each cross were chosen for further analysis of the parental contribution. The number of detectable genes covering SNPs ranged from approximately 11,000 to 13,000 in each cell type (Supplementary Table 1), which could be used to determine the origin of the transcripts and was sufficient for genome-wide analysis of the parental contributions to the ACL and BCL.

Parental genomes contribute equally to the ACL and BCL transcriptomes. We next tested the differences between the maternal and paternal contributions to the transcriptomes of the ACL and BCL in early proembryos. Transcripts with at least ten SNPs that overlapped reads in each sample were extracted to quantify the parental contributions to the ACL and BCL. First, maternal- and paternal-derived reads were quantified in the progenitor cells of the ACL and BCL. We did not find a notable difference in the numbers of maternal and paternal reads in the ACL or BCL from the reciprocal crosses (Fig. 1b and Supplementary Table 2). We then analysed the transcriptomes of the ACL and BCL in 32-cell embryos. Similarly, we found near-equal reads derived from both maternal and paternal genomes in both reciprocal crosses (Fig. 1b). The reads from the paternal genome made up nearly 50% of the transcriptome of both the ACL and BCL (Fig. 1c) in the 1-cell and 32-cell embryos, indicating that the maternal and paternal genomes contribute equally to the transcriptomes of both the ACL and BCL of early proembryos.

Parent-of-origin transcripts in the ACL and BCL are due mainly to de novo monoallelic transcription. Since the parental contributions to the transcriptome of the ACL and BCL of early proembryos are equivalent, we analysed the allelic expression pattern of each gene in the ACL and BCL. We classified the expression patterns as biallelic, Col-0 monoallelic and Ler monoallelic according to the following criteria: fold change (Ler/Col-0) > 4 and false discovery rate < 0.05 (Supplementary Data 1). As expected, most genes were transcribed equally from the maternal and paternal alleles in both the ACL and BCL, even at the one-cell embryo stage (Fig. 1d,e). However, thousands of transcripts were transcribed from the

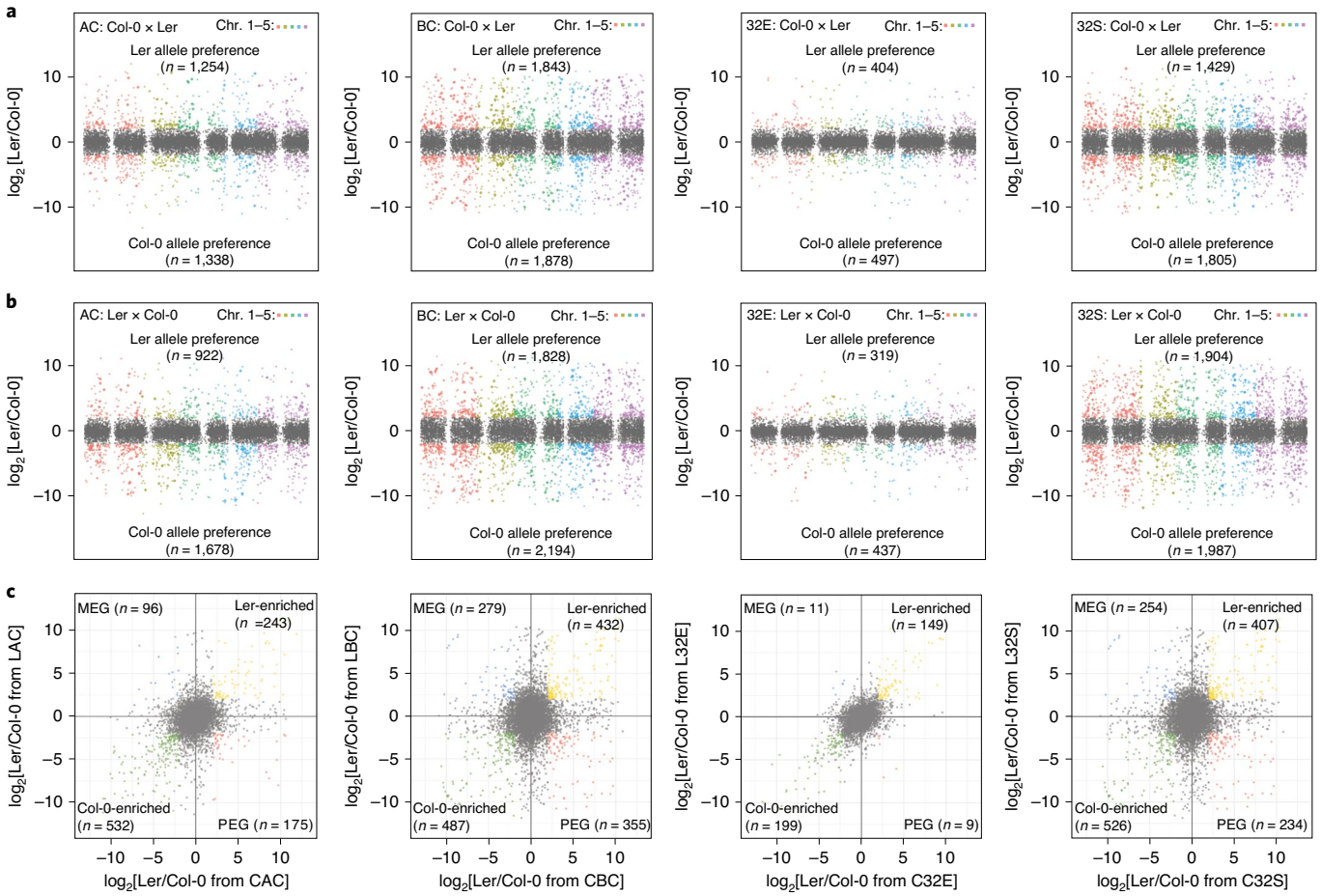

Fig. 2 | Identification of parent-of-origin genes in the transcriptomes of ACLs and BCLs. a,**b**, Scatter plots depicting the allelic ratio (Ler/Col-0) of each gene in the transcriptomes of ACLs and BCLs of hybrid proembryos from the crosses between Col-0 and Ler (**a**) or between Ler and Col-0 (**b**). The allelic expression patterns of the genes are ordered along the *x* axis and coloured according to chromosomal location (1–5); the ratios of Ler to Col-0 ($\log_2$ value) are shown on the *y* axis. The grey dots indicate the genes with biallelic expression. The coloured squares indicate MEGs and PEGs, and the coloured circles indicate other monoallelically expressed genes. **c**, Ler/Col-0 ratios for the transcripts covering SNP-overlapping reads (≥10) in ACLs and BCLs of hybrid proembryos from the reciprocal crosses between Ler and Col-0. The blue dots and red dots indicate MEGs and PEGs, respectively. The green dots and yellow dots indicate Col-0 monoallelically expressed genes (CEGs) and Ler monoallelically expressed genes (LEGs).

maternal or paternal allele in each cell type from the reciprocal crosses. Generally, the number of Col-0 monoallelic genes was comparable to that of the Ler monoallelic genes in each cell type (Figs. 1d and 2a,b). Even though monoallelic genes were found in both the ACL and BCL, the percentage of monoallelically expressed genes in the BCL (33.5%) was significantly higher than that in the ACL (20.4%). The ratio of monoallelically expressed genes in the ACL decreased sharply as the embryo developed (20.4% in 1-cell embryos versus 7.0% in 32-cell embryos). There were no substantial differences between the number of monoallelically expressed genes in the BCL of 1-cell embryos and that of 32-cell embryos (Figs. 1d and 2a,b). Then, the maternally or paternally expressed genes (MEGs or PEGs) were filtered from the monoallelically expressed genes according to the coherence of the allelic expression patterns in the reciprocal crosses. Surprisingly, the maternal/paternal ratio of most monoallelically expressed genes was negatively correlated with the respective reciprocal crosses, which were transcribed from the same *Arabidopsis* ecotype in the reciprocal crosses and were preferentially transcribed from the Col-0 or Ler allele irrespective of their parent of origin (Fig. 2c). The monoallelically expressed genes in the reciprocal cross are probably due to the differences between the Col-0 and Ler genetic backgrounds rather than a parent-of-origin effect on gene expression. After filtering out these genes, the remaining allele-enriched genes were classified as MEGs or PEGs (Fig. 2c). These PEGs and MEGs were scattered across all chromosomes (Fig. 2a,b), indicating that allele-specific transcription is not confined to specific chromosomes or chromosomal regions.

To further distinguish whether the MEGs and PEGs identified in the ACL and BCL are gamete-delivered transcripts or are generated by the prioritization of transcription from one parental allele, the expression patterns of these PEGs and MEGs in sperm cells, egg cells, zygotes and early embryos were analysed[3,15]. The expression levels of most PEGs and MEGs were significantly upregulated in the zygote (Extended Data Fig. 3), indicating that de novo transcription of these genes occurred shortly after fertilization and their allele-specific expression was due to the prioritization of maternal or paternal transcription. Several MEGs and PEGs (Extended Data Fig. 3) are likely to be the result of gamete delivery during fertilization since they showed higher expression levels in gametes, which decreased rapidly after fertilization. The expression levels of 20 selected MEGs and PEGs in gametes and early embryos were further confirmed by quantitative PCR with reverse transcription (RT–qPCR; Extended Data Fig. 4), indicating the high reliability of the transcriptome data.

To further validate the transcriptome analysis results, we selected 20 parent-of-origin genes identified in the ACL or BCL for RT–PCR amplification across SNPs, followed by Sanger sequencing. The allelic expression patterns of these 20 MEGs and PEGs in zygotes and the ACL and BCL of early proembryos at the 1-cell and 32-cell stages were highly consistent with our transcriptome analysis results (Fig. 3 and Extended Data Fig. 5), confirming the high quality of the allele-specific ACL and BCL transcriptomes in early proembryos.

ACs and BCs inherit distinct MEG and PEG allelic expression patterns upon asymmetric zygote division. Since most parent-of-origin genes in ACs or BCs were coexpressed in zygotes (Fig. 4a), we investigated the impact of asymmetric zygote division on the allelic expression of parent-of-origin genes in the two daughter cells. The allelic expression patterns of MEGs and PEGs in the hybrid zygotes were then analysed using hybrid zygote RNA-seq data at 24 h after pollination (Zy24)[3]. We found that a substantial number of PEGs in both ACs (18 of 31) and BCs (31 of 118) inherited the same allelic expression pattern from Zy24 (Fig. 4b,c). However, only a small proportion of MEGs in both ACs (2 of 20) and BCs (7 of 77) displayed the same allelic expression pattern as that of the zygote (Fig. 4b,c). This suggests that PEGs maintained a steady-state monoallelic expression pattern compared with MEGs during asymmetric zygote division, which may be required for both zygote development and subsequent cell lineage specification. Interestingly, we found that a set of biparentally expressed genes (BEGs) in Zy24 were converted into MEGs or PEGs in the BC but maintained a biparental expression pattern in ACs (Fig. 4d). The quantification of parental contributions to the expression of these BEGs revealed that the transition from BEGs in the zygote to MEGs in BCs after asymmetric zygote division mainly resulted from specific paternal transcript degradation (Fig. 4e, Group II) or both maternal expression activation and paternal transcript degradation (Fig. 4e, Group I) in BCs. Conversely, the transition from BEGs in the zygote to PEGs in BCs resulted from specific maternal transcript degradation (Fig. 4f, Group I) or both maternal expression degradation and paternal expression activation (Fig. 4f, Group II).

Parent-of-origin genes display a stage-dependent allelic expression pattern in the ACL and BCL. We then performed a comparative analysis of the AC and the embryo proper of the 32-cell embryo (32E) to investigate the dynamics of parental contributions during ACL development. Significantly fewer parent-of-origin genes were detected in the 32E than in the AC (20 in the 32E versus 271 in the AC) (Fig. 5a). The transition from the maternal class or paternal class to the biparental class indicates de novo transcription of both paternal and maternal alleles of these parent-of-origin genes in the AC during embryonic development (Fig. 5b–d). Unlike the ACL, there are still hundreds of parent-of-origin genes in the suspensor of the 32-cell embryo (32S) (488 in the 32S versus 634 in the BC) (Figs. 2c and 5a). However, only a small proportion of parent-of-origin genes maintain the same allelic expression pattern during BCL development (Fig. 5e). Most parent-of-origin genes in the BCL display a variable allelic expression pattern (Fig. 5f,g). Taken together, allele-specific transcriptome analysis revealed that the allelic expression pattern of parent-of-origin genes is stage-dependent and variable during BCL development.

Parent-of-origin genes display a cell-lineage-dependent allelic expression pattern. Although parent-of-origin genes were identified in both the ACL and BCL, there were more parent-of-origin genes in the BCL than in the ACL in both 1-cell (634 in the BC versus 271 in the AC) and 32-cell embryos (488 in the 32S versus 20 in the 32E) (Figs. 2c and 5a). We found that only a few of the parent-of-origin genes in the ACL and BCL overlapped (Fig. 5b), suggesting that the parent-of-origin genes are expressed in a cell-lineage-dependent manner. Most MEGs and PEGs identified in the BC showed differential maternal and paternal transcript levels in the AC (Fig. 5h,i). Similarly, MEGs and PEGs identified in the 32S were biparentally expressed in the 32E (Fig. 5j,k). Taken together, this analysis revealed that parent-of-origin expression was cell lineage dependent, which is a previously unknown characteristic of parent-of-origin genes. Parent-of-origin genes thus seem to play a more active role in BCL development than in ACL development.

Maternal effects on BCL development. The cell-lineage- and allele-specific transcriptome analysis supports the idea that parent-of-origin genes in the BCL affect the whole BCL developmental process. To test possible parent-of-origin effects on BCL development, different *Arabidopsis* accessions were collected to compare the morphology of BCs and their developmental processes. Generally, the development of BCs was similar among different *Arabidopsis* accessions. However, the lengths of BCs varied in different *Arabidopsis* accessions (Fig. 6a,b). According to morphological differences among accessions, a reciprocal cross between Col-0 and Ler was performed to investigate the parent-of-origin effects on BCL development (Fig. 6c,d). Although egg cell lengths in Col-0 and Ler are similar (Extended Data Fig. 6), the length of the BC in the hybrid proembryo was consistent with that of the maternal lines in the reciprocal crosses between Col-0 and Ler. This phenomenon was further confirmed by another reciprocal cross between Wassileskija (Ws) and Ler (Fig. 6c,e). Similarly, no notable maternal effect on the AC length was observed in the hybrid proembryo (Extended Data Fig. 7). However, BCL development was also highly consistent with that of the maternal line in both mature suspensor length and cell number (Fig. 6f–h), suggesting that the BCL must be guided by maternal machinery to regulate cell elongation and cell division.

Although a maternally biased transcriptome was not observed in BCL specification, hundreds of MEGs have been observed in the BCL in both 1-cell and 32-cell embryos, suggesting that these BCL-specific MEGs probably contributed to the maternal control of BCL development. To investigate the roles of these MEGs in the process of BCL specification, Gene Ontology (GO) analysis was performed for each MEG identified in the BCL (Supplementary Data 2). The numbers of MEGs and PEGs in each GO term were then quantified for statistical analysis. As expected, the GO enrichment analysis revealed that distinct molecular pathways were enriched in PEGs and MEGs identified in the BCL (Fig. 7a,b). The MEGs in the BCL at the one-cell embryo stage were enriched mainly in biological processes related to lipid metabolism ($P=7.88\times10^{-3}$), cell fate specification ($P=3.93\times10^{-3}$) and histone demethylation ($P=2.67\times10^{-3}$). The MEGs in the BCL at the 32-cell embryo stage were enriched in processes related to cell cycle regulation ($P=3.84\times10^{-3}$), protein phosphorylation ($P=1.35\times10^{-4}$) and cell communication ($P=1.17\times10^{-3}$), which may be involved in suspensor morphology and functional establishment. We then used the Kyoto Encyclopedia of Genes and Genomes (KEGG) database to investigate potential MEG and PEG pathways. MEGs and PEGs in the BCL were classified into known KEGG pathways (Supplementary Data 3). Consistent with the GO analysis results, MEGs and PEGs in the BCL were also grouped into distinct KEGG pathways (Fig. 7c). Pathways including MAPK signalling ($n=5$), plant hormone signal transduction ($n=5$) and carbon metabolism pathways ($n=4$) were annotated for MEGs in the BCL (Fig. 7c). Notably, 16 MEGs maintained their maternal monoallelic expression patterns during BCL development (Fig. 4b), involving several biological processes including transcription, protein phosphorylation and carbohydrate metabolic process (Supplementary Data 4).

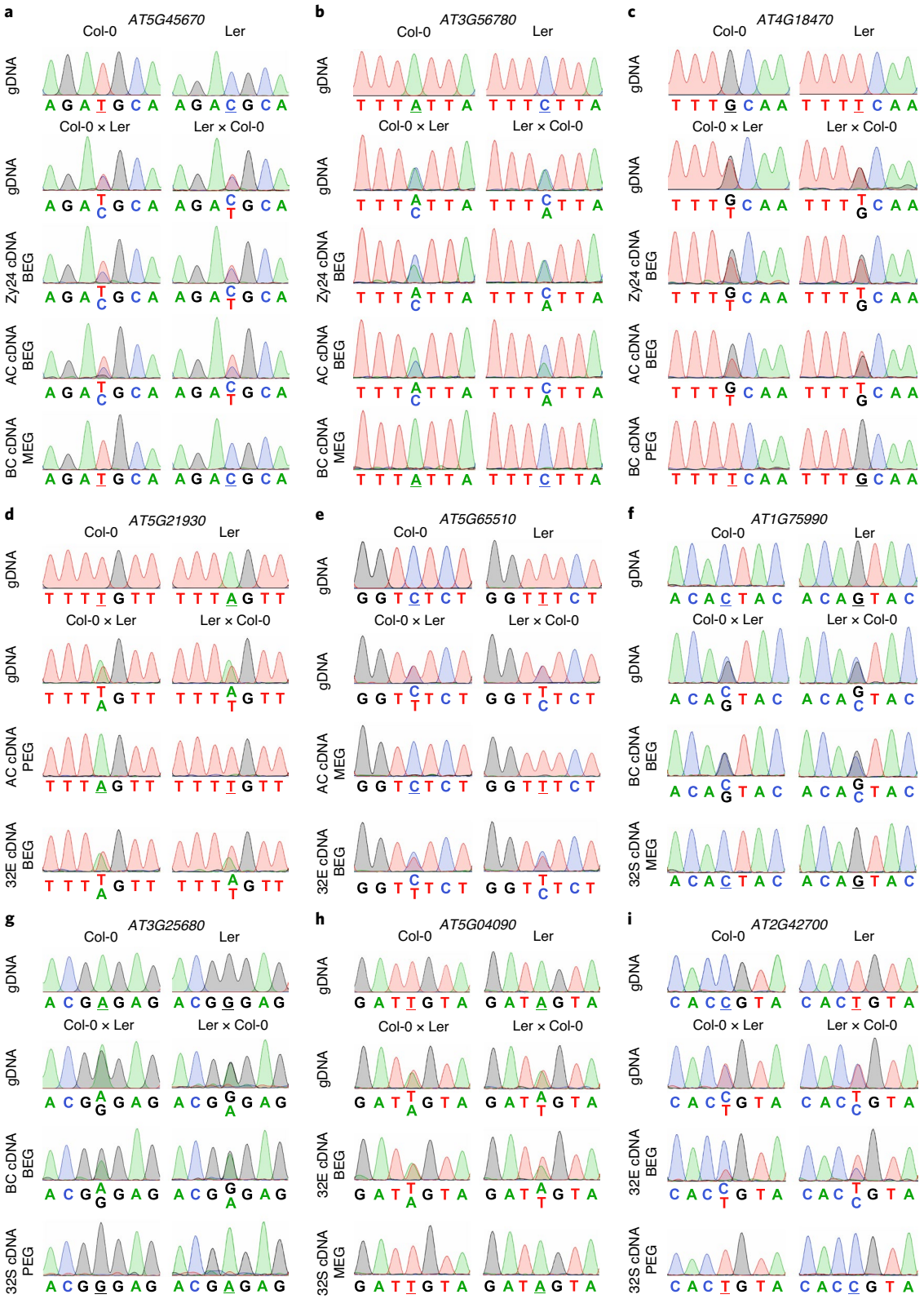

Fig. 3 | Allele-specific expression analysis of parent-of-origin genes in zygotes, ACLs and BCLs of early proembryos. a–i, Sanger sequencing chromatographs of PCR products at the following selected SNP sites, showing the allelic expression patterns of parent of origin genes: *AT5G45670* (**a**), *AT3G56780* (**b**), *AT4G18470* (**c**), *AT5G21930* (**d**), *AT5G65510* (**e**), *AT1G75990* (**f**), *AT3G25680* (**g**), *AT5G04090* (**h**) and *AT2G42700* (**i**). Complementary DNA of hybrid zygotes, ACLs and BCLs of early proembryos from the reciprocal crosses between *Arabidopsis* Col-0 and Ler were prepared for RT–PCR and Sanger sequencing. Genomic DNA (gDNA) of *Arabidopsis* Col-0, Ler and F_1 hybrids from the reciprocal crosses between Col-0 and Ler were used as controls. The polymorphic nucleotide is underlined in each chromatogram.

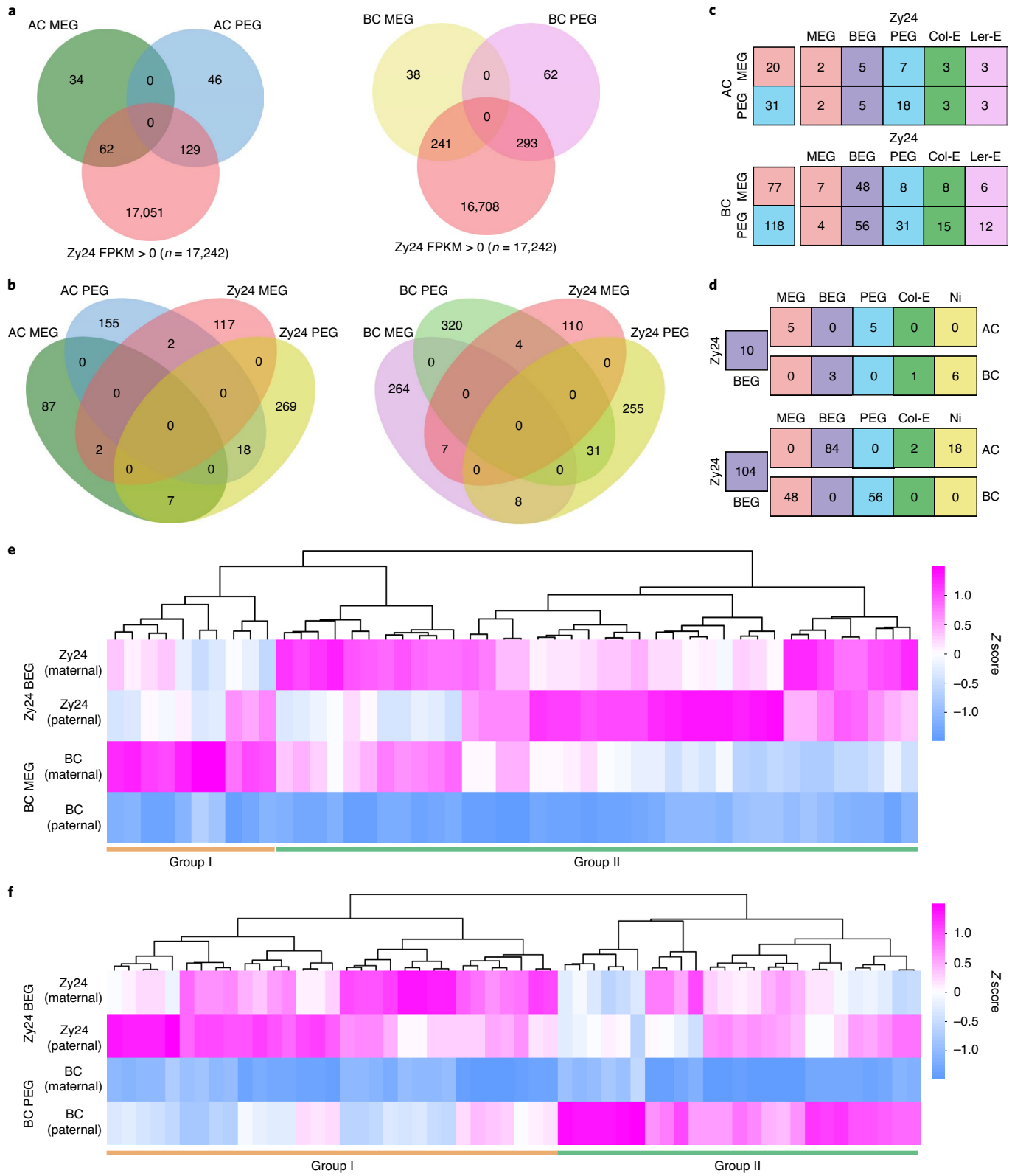

Fig. 4 | ACs and BCs inherit distinct allelic expression patterns of parent-of-origin genes from the zygote upon asymmetric cell division. a, Most PEGs and MEGs identified in ACs and BCs are coexpressed in the zygote. Expressed genes (fragments per kilobase of transcript per million mapped reads (FPKM) > 0) in the zygote were used for the overlap analysis. **b**, Venn diagrams showing MEGs and PEGs identified in zygotes, ACs and BCs. **c**, MEGs and PEGs in ACs and BCs inherit distinct allelic expression patterns from the zygote. Expressed genes with ≥10 SNP-overlapping reads and consistent allelic expression patterns in the reciprocal crosses in zygotes were used for the overlap analysis. **d**, The transitions of BEGs identified in zygotes into MEGs or PEGs in BCs. Ni, no information. **e**,**f**, Heat maps showing the changes in the levels of paternal and maternal transcripts of the BEGs identified in zygotes, which converted into MEGs (**e**) or PEGs (**f**) in BCs after asymmetric division. The paternal and maternal transcript levels of each gene are the mean expression levels of the reciprocal crosses.

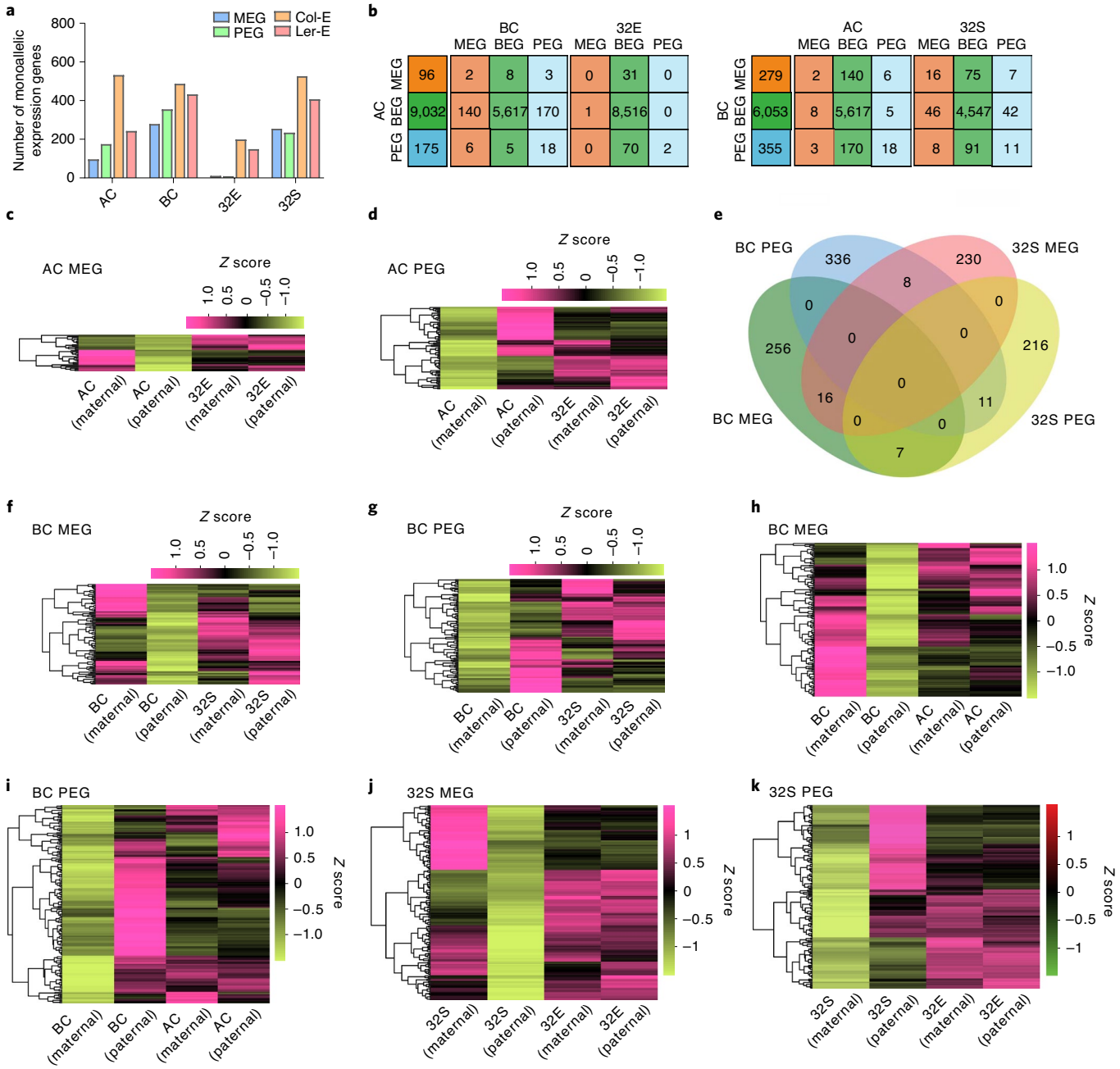

Fig. 5 | Allelic expression patterns of parent-of-origin genes are dependent on developmental stage and cell lineage. a, Numbers of genes with consistent monoallelic expression patterns in reciprocal crosses. **b**, Class transitions of MEGs, BEGs and PEGs during the process of lineage specification and embryo development. **c**,**d**, Heat maps displaying the changes in the expression levels of MEGs (**c**) and PEGs (**d**) identified in the AC in the process of AC development. **e**, Venn diagrams showing MEGs and PEGs identified in the BCL. **f**–**i**, Spatio-temporal allelic expression patterns of MEGs and PEGs identified in BCs. The heat maps show the changes in the expression levels of BC MEGs and PEGs in the 32S (**f**,**g**) and the AC (**h**,**i**). **j**,**k**, Heat maps displaying allelic expression patterns of 32S MEGs (**j**) and PEGs (**k**) in the 32E.

Discussion

Early embryogenesis in animals is controlled by maternal factors, and the transcriptome of early embryos is dominated by maternal transcripts[1]. In contrast to animals, the plant zygotic genome is activated shortly after fertilization[3,16,17]. The maternal and paternal genomes contribute equally to the transcriptome of early embryos[3,7]. However, early proembryos are composed of two distinguishable domains: the ACL and BCL, which show distinct transcriptomes shortly after asymmetric division. The development of the two domains might be differentially regulated. This idea is supported by our recent finding that the morphology and developmental processes of the hybrid tobacco BCL are similar to the maternal line used for the crosses[14]. BCL development thus provides a unique model for transcriptome analysis to confirm whether the development of the two domains is differentially regulated. It also offers an opportunity to correlate the parental contributions to the transcriptome with the role of parental control during proembryonic development.

In this study, we found that the maternal and paternal genomes contributed equally to the transcriptomes of both the ACL and BCL

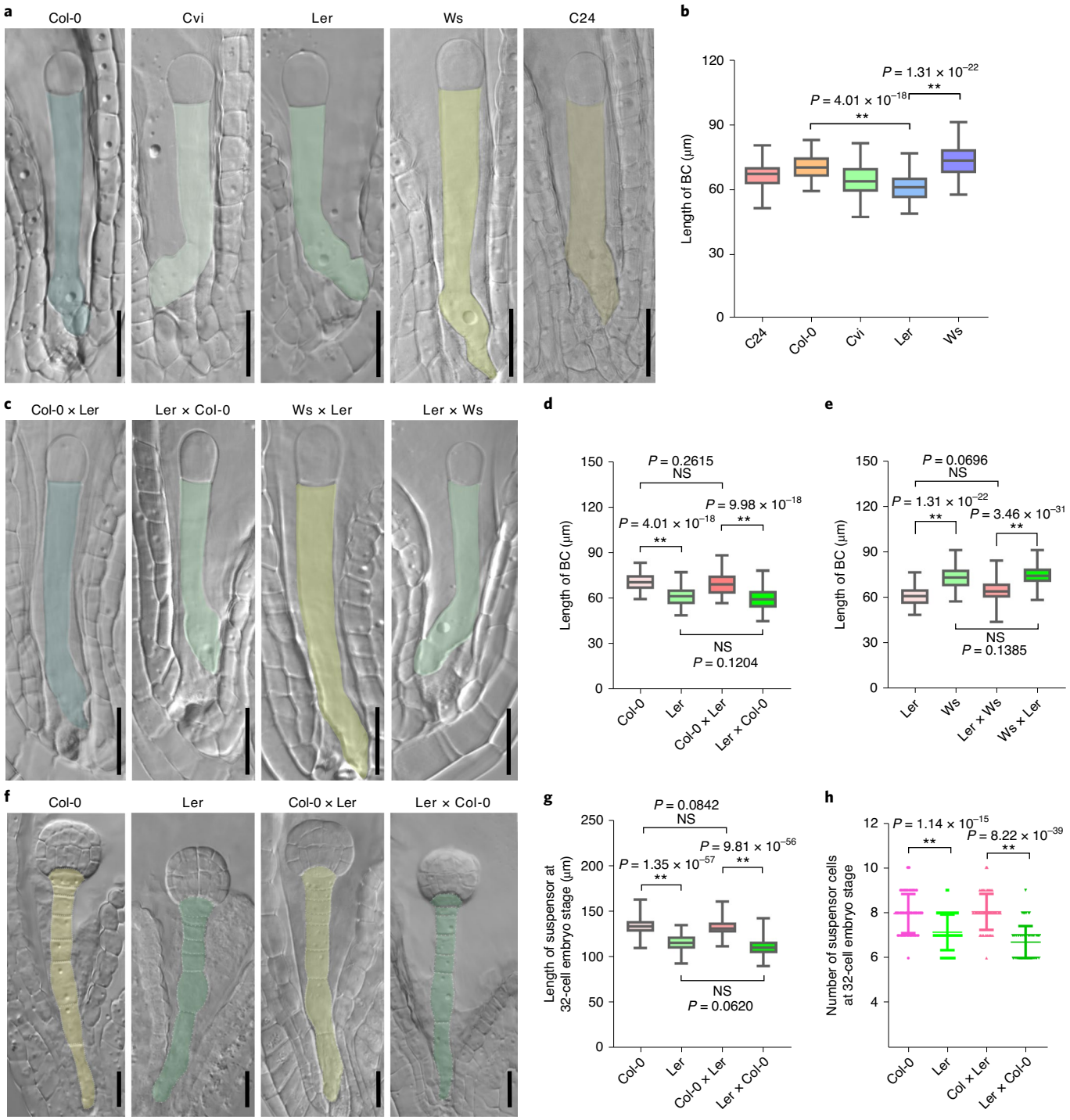

Fig. 6 | Maternal effect on BCL development in early embryogenesis. a,**b**, BC length varies in different *Arabidopsis* ecotypes. Representative pictures of one-cell embryos from different *Arabidopsis* ecotypes are shown in **a**. Statistical data for the length of the BC in different *Arabidopsis* ecotypes are shown in **b** (C24, $n=96$; Col-0, $n=95$; Cape Verde Islands (Cvi), $n=104$; Ler, $n=90$; Ws, $n=97$). **c**–**e**, BC lengths of hybrid embryos are highly consistent with the maternal line. Representative pictures of hybrid one-cell embryos from the reciprocal crosses between Col-0 and Ler or between Ws and Ler are shown in **c**. Statistical data for the lengths of hybrid BCs from the reciprocal crosses are shown in **d**,**e** (Col-0, $n=95$; Ler, $n=90$; Ws, $n=97$; Col-0 × Ler, $n=92$; Ler × Col-0, $n=132$; Ler × Ws, $n=128$; Ws × Ler, $n=134$). **f**–**h**, Cell number and cell length of the BCL of hybrid proembryos are highly consistent with the maternal line. Representative pictures of isogenic and hybrid 32-cell embryos with a fully developed suspensor are shown in **f**. The length (**g**) and number (**h**) of suspensors of isogenic and hybrid 32-cell embryos are also shown. Each datum represents mean ± s.d. (in **g**, Col-0, $n=201$; Ler, $n=136$; Col-0 × Ler, $n=151$; Ler × Col-0, $n=163$; in **h**, Col-0, $n=177$; Ler, $n=125$; Col-0 × Ler, $n=137$; Ler × Col-0, $n=157$). Scale bars, 20 μm. The data for cell length are presented in box-and-whisker plots. The bottom and top of each box represent the 25th and 75th percentiles, respectively. The centre line represents the 50th percentile. The whiskers indicate the minimum and maximum values. The statistical analysis was performed using two-tailed Student's *t*-tests; $^{**}P<0.01$.

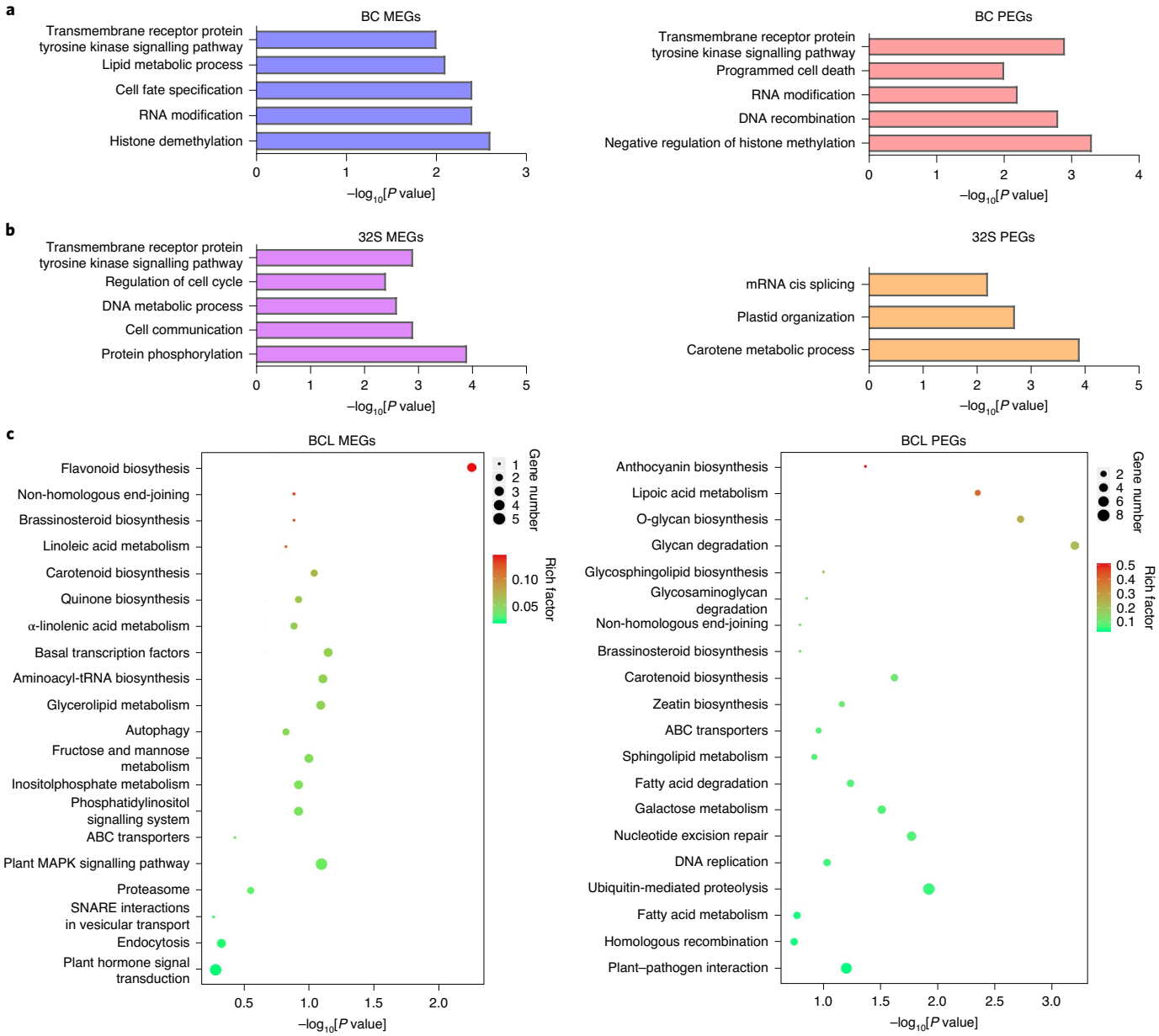

Fig. 7 | GO and KEGG analysis of MEGs and PEGs identified in the BCL. a,**b**, GO analysis of MEGs and PEGs identified in the BCL at the 1-cell embryo (**a**) and 32-cell embryo (**b**) stages. **c**, KEGG analysis of MEGs and PEGs identified in the BCL. The statistical analysis was performed using a hypergeometric test.

in both 1-cell and 32-cell embryos. We also confirmed that BCL development was mainly controlled by maternal factors in hybrid seeds of *Arabidopsis*. Thus, equivalent parental contributions to the transcriptome did not ensure equivalent parental functional contributions to proembryonic development. The parental contributions to early embryonic development may be exerted by a few key parent-of-origin genes. In fact, sperm-delivered *SSP* and paternally expressed *BBM1* have been reported to regulate early embryonic development[18,19].

These findings led us to focus more on the specific roles of MEG and PEG expression during embryogenesis. We identified hundreds of PEGs and MEGs in both the ACL and BCL of early proembryos at different stages. Genes with a parent-of-origin effect showed stage-dependent expression, and the silent allele was not always completely silent during embryogenesis. Notably, the conversion of BEGs in zygotes into MEGs or PEGs in the BCs, but not in the ACs, was found after asymmetric zygote division (Figs. 3a–c and 4d), suggesting that the ACs and BCs may be equipped with different epigenetic mechanisms to regulate the allelic expression of these BEGs soon after asymmetric zygote division. DNA methylation asymmetry and histone modification are two known mechanisms for parent-of-origin gene expression[20]. The genome-wide analysis of epigenetic marks in the zygote and two daughter cells will shed light on the yet unknown mechanism of parent-of-origin expression resetting in the BCL and provide a new mechanism for the cell fate specification of the BCL.

Most PEGs and MEGs identified in the ACL were transformed into BEGs by the 32-cell stage, resulting in a significant decrease in the number of MEGs and PEGs in the ACL of 32-cell embryos. These results suggest that the parent-of-origin effect of ACL development occurs mainly during early embryonic stages. The expression of parent-of-origin genes was also cell type dependent.

We found that the same genes displayed differential allelic expression patterns in the ACL and BCL. Interestingly, there were more parent-of-origin genes in the BCL than in the ACL, suggesting that parent-of-origin effects on BCL development are more significant than those on ACL development.

Our analysis indicates that the parental effects of the two domains of the proembryo are differentially regulated. This finding may be a useful clue for the discovery of the detailed mechanism underlying cell fate determination and selection of distinct developmental pathways of the BCL and ACL. Apart from the effect of MEGs on BCL development, maternal influence from egg-cell-expressed non-parent-of-origin genes and ovule integument are also involved in proembryo development. Recently, egg-cell-expressed transcription factors HDG11/12 and mitochondrial GCD1 were shown to regulate zygote development[21,22]. The fertilization of *hdg11-1*, *hdg12-2* or *gcd1* egg cells with wild-type pollen did not rescue zygote developmental defects in mutants, providing clear evidence that egg-cell-derived maternal factors other than parent-of-origin genes are also involved in early embryogenesis. In addition, ovule-integument-derived auxin and gibberellin as maternal signals were reported to be important for early proembryo development[13,23], suggesting a complex mechanism underlying early embryogenesis. These findings further indicate that maternal and paternal roles in embryogenesis should be discussed in specific developmental events at specific developmental stages. Different developmental events may be regulated by different mechanisms or different combinations of parental and microenvironmental factors at different developmental stages. Therefore, there might be no general mechanism covering the entire process of early embryogenesis in terms of parental control or parental roles. These findings will facilitate further investigation of parental contributions to early embryogenesis and enhance our understanding of the molecular basis of cell fate specification in both the ACL and BCL.

Methods

Plant materials. *Arabidopsis thaliana* polymorphic Col-0, Ler, Ws, Cvi and C24 were used in the present study. *DD45::GFP* reporter lines in the Col-0 or Ler background were provided by G. Drews (University of Utah, Salt Lake City, UT, USA) and Wei-Cai Yang (Institute of Genetics and Developmental Biology, CAS, China), respectively. The plants were cultivated in greenhouses under long-day conditions (16 h light/8 h dark) at 22 ± 1 °C.

Isolation of the ACL and BCL of early proembryos. Early *Arabidopsis* embryo isolation and the separation of the ACL and BCL of early proembryos were performed according to our previous protocol[24]. The isolated ACLs and BCLs were extensively washed four times, transferred to lysis buffer (Life Technologies) and stored in liquid nitrogen for mRNA isolation.

cDNA preparation and library construction for RNA-seq. The RNA extraction and cDNA preparation were performed according to our recent report[3]. Briefly, mRNA was isolated using the Dynabeads mRNA DIRECTTM Micro Kit (Life Technologies). cDNA synthesis and amplification were performed using a SMARTer Ultra Low RNA Kit for Illumina Sequencing (Clontech). The cDNA was then purified using an Agencourt AMPure purification kit (Beckman Coulter). After purification and quantification, the cDNA was used for library construction. The RNA-seq libraries were prepared using a Nextera XT Library Prep Kit (Illumina) according to the reference guide, and sequenced on an Illumina HiSeq 2500 system, generating about six gigabases of raw sequence data for each library.

RNA-seq data analysis. Clean reads of the samples were prepared using Cutadapt v.1.15 (ref. [25]) and an in-house script. The clean reads were then mapped to the *Arabidopsis* Col-0 genome and transcripts using Bowtie 2 v.2.2.2 (ref. [26]). The gene expression levels in each sample were quantified as fragments per kilobase of transcript per million mapped reads using RSEM v.1.3.0 (ref. [27]).

Allele-specific transcriptome analysis. SNPs between the Col-0 and Ler genomes were downloaded from the Arabidopsis Information Resource (TAIR) (https://www.arabidopsis.org/). Allele-specific transcriptome analysis was performed according to our previously published protocol[3]. First, a modified in silicon simulation method was used to overcome the mapping bias of reads generated from hybrid ACLs and BCLs of early proembryos[28]. SNPs between Col-0 and Ler with significant mapping bias were excluded for further analysis. Second, SNP sites with less than 95% Col-0- and Ler-specific bases were filtered out, and then the remaining SNPs covering the annotated genes were used for subsequent analysis. Transcripts with at least ten SNP-overlapping reads in the transcriptomes from reciprocal crosses between Col-0 and Ler were selected for the parental contribution analysis. All SNPs in the same gene were combined to quantify the parental contribution. A binomial distribution test was used to assess the differences in Col-0 and Ler allele expression of the gene. Genes with allelic ratios (Ler/Col-0) of >4 or <0.25 and a false discovery rate of <0.05 were considered to have monoallelic expression.

GO and KEGG analysis. The GO terms for each annotated gene were downloaded from TAIR (https://www.arabidopsis.org/). The pathway information for *Arabidopsis* genes was downloaded from the KEGG database (https://www.genome.jp/). The statistical significance of enriched GO terms and KEGG pathways was calculated using a hypergeometric test in R v.3.6.1.

RT–qPCR. RT–qPCR was used to investigate the expression levels of MEGs and PEGs in sperm cells, egg cells, zygotes, the ACL and the BCL of early proembryos. The RT–qPCR reaction was performed in a 10 µl mixture containing FastStart Essential DNA Green Master (Roche), cDNAs prepared from different cells and 250 nM forward and reverse primers. The PCR protocols were as follows: 95 °C for 10 min, and 40 cycles (95 °C for 15 s, annealing at 60 °C for 20 s; extension at 72 °C for 30 s) on a CFX Connect Real-Time System (Bio-Rad Laboratories). The RT–qPCR data were collected and analysed using the Bio-Rad CFX Manager v.3.1. *AT5G25760*, *AT1G58050* and *AT4G05320* were selected as reference genes according to a previous report[29]. The primers for RT–qPCR are listed in Supplementary Table 3.

RT–PCR and Sanger sequencing. RT–PCR was performed on cDNA prepared from hybrid zygotes, ACLs and BCLs of early proembryos from reciprocal crosses between *Arabidopsis* Col-0 and Ler using Phanta HS Super-Fidelity DNA Polymerase (Vazyme). The PCR products were separated on a 1.2% DNA agarose gel, and the DNA bands were purified using a DNA Cleanup kit (Tiagen). The purified PCR products were then used for Sanger sequencing, and SNP sites for each selected gene on sequencing chromatograms were used to evaluate its parent of origin. PCR amplification of gDNA prepared from the leaves of *Arabidopsis* Col-0, Ler and F_1 hybrids from the reciprocal crosses between Col-0 and Ler with the same primer was used as a control. The primer sequences for the selected genes are listed in Supplementary Table 3.

Image acquisition and cell length measurement. Egg cells, zygotes, and early embryos were observed under a confocal microscope equipped with the LAS X v.3.5.5.19976 (Leica TCS SP8). Cell length was measured using ImageJ v.1.52a (https://imagej.nih.gov/ij/). The box-and-whisker plots and aligned dot plots were prepared using GraphPad Prism v.8.4.2 (https://www.graphpad.com/).

Statistical analyses and plots. R v.3.6.1 (http://www.r-project.org/) was used for the statistical analyses and plotting of results.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All RNA-seq data have been uploaded to the NCBI Gene Expression Omnibus (GEO) under accession no. GSE107700. Previously published RNA-seq data (GEO accession nos GSE120669, GSE121003 and GSE135422) were also used in the present study. The GO terms for each annotated gene and SNPs between the Col-0 and Ler genomes were downloaded from the TAIR (https://www.arabidopsis.org/). The pathway information for *Arabidopsis* genes was downloaded from the KEGG database (https://www.genome.jp/). Source data are provided with this paper.

Code availability

The in-house script for clean reads preparation is available from GitHub at https://github.com/frasergen-rna/FraserQC. Other details of the code for data analysis of the present study are available from the corresponding authors upon reasonable request.

Received: 30 March 2020; Accepted: 16 September 2020;
Published online: 26 October 2020

References

1. Tadros, W. & Lipshitz, H. D. The maternal-to-zygotic transition: a play in two acts. *Development* **136**, 3033–3042 (2009).
2. Zhao, P. & Sun, M. X. The maternal-to-zygotic transition in higher plants: available approaches, critical limitations, and technical requirements. *Curr. Top. Dev. Biol.* **113**, 373–398 (2015).

3. Zhao, P. et al. Two-step maternal-to-zygotic transition with two-phase parental genome contributions. *Dev. Cell* **49**, 882–893 (2019).
4. Schier, A. F. The maternal-zygotic transition: death and birth of RNAs. *Science* **316**, 406–407 (2007).
5. Lee, M. T., Bonneau, A. R. & Giraldez, A. J. Zygotic genome activation during the maternal-to-zygotic transition. *Annu. Rev. Cell Dev. Biol.* **30**, 581–613 (2014).
6. Autran, D. et al. Maternal epigenetic pathways control parental contributions to *Arabidopsis* early embryogenesis. *Cell* **145**, 707–719 (2011).
7. Nodine, M. D. & Bartel, D. P. Maternal and paternal genomes contribute equally to the transcriptome of early plant embryos. *Nature* **482**, 94–97 (2012).
8. Schon, M. A. & Nodine, M. D. Widespread contamination of *Arabidopsis* embryo and endosperm transcriptome data sets. *Plant Cell* **29**, 608–617 (2017).
9. Baroux, C., Autran, D., Raissig, M. T., Grimanelli, D. & Grossniklaus, U. Parental contributions to the transcriptome of early plant embryos. *Curr. Opin. Genet. Dev.* **23**, 72–74 (2013).
10. Del Toro-De Leon, G., Lepe-Soltero, D. & Gillmor, C. S. Zygotic genome activation in isogenic and hybrid plant embryos. *Curr. Opin. Plant Biol.* **29**, 148–153 (2016).
11. Kawashima, T. & Goldberg, R. B. The suspensor: not just suspending the embryo. *Trends Plant Sci.* **15**, 23–30 (2010).
12. Zhao, P. et al. A bipartite molecular module controls cell death activation in the basal cell lineage of plant embryos. *PLoS Biol.* **11**, e1001655 (2013).
13. Shi, C. et al. Maternal control of suspensor programmed cell death via gibberellin signaling. *Nat. Commun.* **10**, 3484 (2019).
14. Luo, A., Zhao, P., Zhang, L. Y. & Sun, M. X. Initiation of programmed cell death in the suspensor is predominantly regulated maternally in a tobacco hybrid. *Sci. Rep.* **6**, 294677 (2016).
15. Borg, M. et al. Targeted reprogramming of H3K27me3 resets epigenetic memory in plant paternal chromatin. *Nat. Cell Biol.* **22**, 621–629 (2020).
16. Anderson, S. N. et al. The zygotic transition is initiated in unicellular plant zygotes with asymmetric activation of parental genomes. *Dev. Cell* **43**, 349–358 (2017).
17. Chen, J. et al. Zygotic genome activation occurs shortly after fertilization in maize. *Plant Cell* **29**, 2106–2125 (2017).
18. Bayer, M. et al. Paternal control of embryonic patterning in *Arabidopsis thaliana*. *Science* **323**, 1485–1488 (2009).
19. Khanday, I., Skinner, D., Yang, B., Mercier, R. & Sundaresan, V. A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. *Nature* **565**, 91–95 (2019).
20. Batista, R. A. & Kohler, C. Genomic imprinting in plants—revisiting existing models. *Genes Dev.* **34**, 24–36 (2020).
21. Ueda, M. et al. Transcriptional integration of paternal and maternal factors in the *Arabidopsis* zygote. *Genes Dev.* **31**, 617–627 (2017).
22. Wu, J. J. et al. Mitochondrial GCD1 dysfunction reveals reciprocal cell-to-cell signaling during the maturation of *Arabidopsis* female gametes. *Dev. Cell* **23**, 1043–1058 (2012).
23. Robert, H. S. et al. Maternal auxin supply contributes to early embryo patterning in *Arabidopsis*. *Nat. Plants* **4**, 548–553 (2018).
24. Zhou, X., Shi, C., Zhao, P. & Sun, M. Isolation of living apical and basal cell lineages of early proembryos for transcriptome analysis. *Plant Reprod.* **32**, 105–111 (2019).
25. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
26. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
27. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinform.* **12**, 323 (2011).
28. Panousis, N. I., Gutierrez-Arcelus, M., Dermitzakis, E. T. & Lappalainen, T. Allelic mapping bias in RNA-sequencing is not a major confounder in eQTL studies. *Genome Biol.* **15**, 467 (2014).
29. Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K. & Scheible, W. R. Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol.* **139**, 5–17 (2005).

Acknowledgements

We thank W.-C. Yang (Institute of Genetics and Developmental Biology, CAS, China) and G. Drews (University of Utah, Salt Lake City, UT, USA) for their kind offer of *DD45::GFP* marker lines in the Ler and Col-0 backgrounds. This work was supported by the National Natural Science Foundation of China (grant nos 31770355 and 31970340) and the Fundamental Research Funds of the Central Universities (grant nos 2042020kf0198 and 2042020kf1074). We thank G. Yuan (Frasergen Bioinformatics Co., Ltd) for a suggestion regarding the data analysis.

Author contributions

P.Z. and M.-x.S. designed the experiments. P.Z., X.Z., Y.Z. and Y.R. performed the experiments. P.Z., X.Z. and M.-x.S. analysed the data. P.Z. and M.-x.S. wrote the manuscript. All authors discussed the results and agreed on the manuscript before submission.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41477-020-00793-x.

Supplementary information is available for this paper at https://doi.org/10.1038/s41477-020-00793-x.

Correspondence and requests for materials should be addressed to P.Z. or M.-x.S.

Peer review information *Nature Plants* thanks Takashi Okamoto and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.