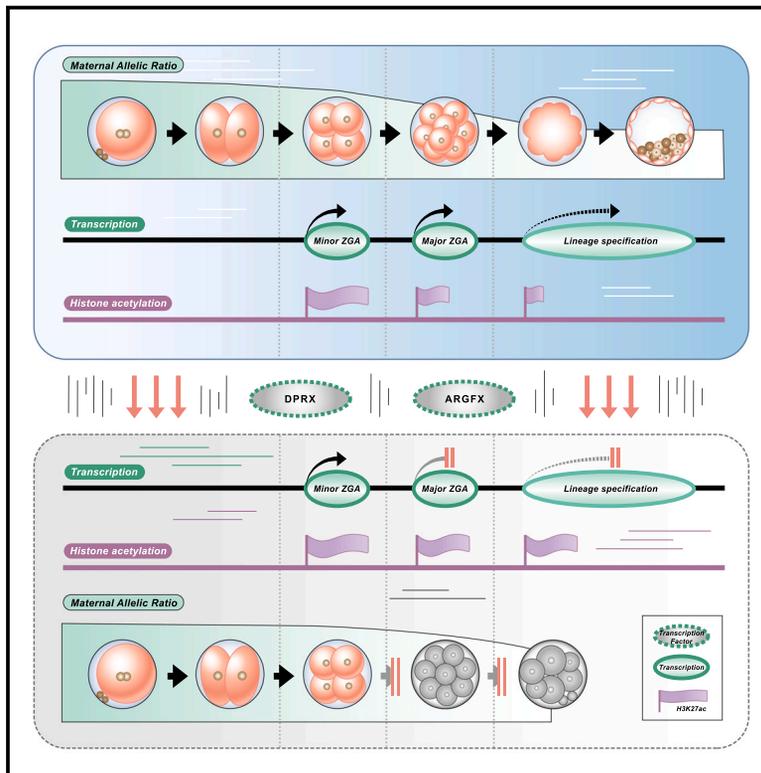


# Developmental Cell

## Allelic transcriptomic profiling identifies the role of PRD-like homeobox genes in human embryonic-cleavage-stage arrest

### Graphical abstract



### Authors

Qianying Guo, Fanqing Xu, Shi Song, ..., Jie Qiao, Liying Yan, Peng Yuan

### Correspondence

zqyan@pku.edu.cn (Z.Y.), jie.qiao@263.net (J.Q.), yanliyingkind@aliyun.com (L.Y.), peng.yuan@bjmu.edu.cn (P.Y.)

### In brief

Cleavage-stage arrest is a leading cause of low blastulation rate in human embryos. Through allelic transcriptome analysis of early human embryos, Guo et al. quantified maternal-to-zygotic transition progression and identified that PRD-like homeobox transcription factors DPRX and ARGFX control human embryonic development and whose deficiency is an etiology underlying human cleavage-stage arrest.

### Highlights

- Maternal allelic ratio can accurately quantify maternal-to-zygotic transition (MZT)
- Some morphologically normal 8-cell embryos are already transcriptionally arrested
- PRD-like homeobox transcription factors DPRX and ARGFX control the MZT process
- DPRX and ARGFX deficiency is an etiology underlying human cleavage-stage arrest

Article

# Allelic transcriptomic profiling identifies the role of PRD-like homeobox genes in human embryonic-cleavage-stage arrest

Qianying Guo,<sup>1,2,3,4,5,8</sup> Fanqing Xu,<sup>1,2,3,4,5,8</sup> Shi Song,<sup>1,2,3,4,5,8</sup> Siming Kong,<sup>1,2,3,4,5,8</sup> Fan Zhai,<sup>1,2,3,4,5,8</sup> Yuwen Xiu,<sup>1,2,3,4,5,6</sup> Dandan Liu,<sup>1,2,3,4,5</sup> Ming Li,<sup>1,2,3,4,5</sup> Ying Lian,<sup>1,2,3,4,5</sup> Ling Ding,<sup>1,2,3,4,5</sup> Qian Liu,<sup>1,2,3,4,5</sup> Ming Yang,<sup>1,2,3,4,5,6</sup> Zhengrong Du,<sup>1,2,3,4,5</sup> Nan Wang,<sup>1,2,3,4,5</sup> Chuan Long,<sup>1,2,3,4,5</sup> Xiaomeng Wang,<sup>1,2,3,4,5</sup> Yuqian Wang,<sup>1,2,3,4,5,6</sup> Zhiqiang Yan,<sup>1,2,3,4,5,\*</sup> Jie Qiao,<sup>1,2,3,4,5,6,7,\*</sup> Liying Yan,<sup>1,2,3,4,5,\*</sup> and Peng Yuan<sup>1,2,3,4,5,9,\*</sup>

<sup>1</sup>Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Peking University Third Hospital, Beijing 100191, China

<sup>2</sup>National Clinical Research Center for Obstetrics and Gynecology (Peking University Third Hospital), Beijing 100191, China

<sup>3</sup>State Key Laboratory of Female Fertility Promotion, Department of Obstetrics and Gynecology Peking University Third Hospital, Beijing 100191, China

<sup>4</sup>Key Laboratory of Assisted Reproduction (Peking University), Ministry of Education, Beijing 100191, China

<sup>5</sup>Beijing Key Laboratory of Reproductive Endocrinology and Assisted Reproductive Technology, Beijing 100191, China

<sup>6</sup>Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China

<sup>7</sup>Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

<sup>8</sup>These authors contributed equally

<sup>9</sup>Lead contact

\*Correspondence: [zqyan@pku.edu.cn](mailto:zqyan@pku.edu.cn) (Z.Y.), [jie.qiao@263.net](mailto:jie.qiao@263.net) (J.Q.), [yanliyingkind@aliyun.com](mailto:yanliyingkind@aliyun.com) (L.Y.), [peng.yuan@bjmu.edu.cn](mailto:peng.yuan@bjmu.edu.cn) (P.Y.)  
<https://doi.org/10.1016/j.devcel.2024.12.031>

## SUMMARY

Cleavage-stage arrest in human embryos substantially limits the success rate of infertility treatment, with maternal-to-zygotic transition (MZT) abnormalities being a potential contributor. However, the underlying mechanisms and regulators remain unclear. Here, by performing allelic transcriptome analysis on human pre-implantation embryos, we accurately quantified MZT progression by allelic ratio and identified a fraction of 8-cell embryos, at the appropriate developmental time point and exhibiting normal morphology, were in transcriptionally arrested status. Furthermore, we identified PAIRED (PRD)-like homeobox transcription factors divergent paired-related homeobox (*DPRX*) and arginine-fifty homeobox (*ARGFX*) as factors involved in MZT, whose deficiency severely impairs MZT and lineage specification and leads to aberrant retention of histone acetylation. By reversing the acetylation retention caused by *DPRX* and *ARGFX* defects, embryonic arrest can be partially rescued. Our study identifies factors involved in human MZT and elucidates the etiology underlying human cleavage-stage arrest.

## INTRODUCTION

Approximately 16.7% of the adult population are affected by infertility in their life time, posing a significant challenge to human health.<sup>1</sup> Although assisted reproductive technology (ART) stands out as the most effective infertility treatment, the live birth rate after ART is lower than 30%.<sup>1</sup> One of the major challenges to the success of ART is the low blastulation rate of human embryos, which hovers around 45%,<sup>2</sup> markedly lower than other species, such as over 80% in murine<sup>3</sup> and swine.<sup>4</sup> Of the embryos that fail to develop into blastocysts, a large proportion suffer developmental arrest at the cleavage stage. In particular, more than 50% of high-quality human 8-cell embryos fail to compact and develop to the blastocyst stage.<sup>5</sup> Until recently, it has remained unclear why human embryos are so prone to arrest at the cleavage stage. Identifying the causes of human cleavage-stage

arrest may help develop clinical treatments and improve the success rate of ART.

Chromosomal abnormality is a common problem in human preimplantation embryo development, yet aneuploidy embryos could develop to blastocysts at a relatively high rate,<sup>6,7</sup> suggesting that other mechanisms are involved in human cleavage-stage arrest. Recent studies of human embryonic arrest have shown that arrested embryos exhibit disruptions in metabolic processes, including oxidative phosphorylation and glycolysis, as well as aberrations in epigenetic modifications, such as DNA methylation.<sup>5,8</sup> However, these studies analyzed embryos that had already arrested, which introduces confounding factors in that these observations may be of phenotypes that occurred after developmental arrest rather than pinpointing the initiating factors that cause the arrest. In addition, beyond the generic characteristics of arrested embryos, the specific molecular

mechanisms driving embryonic arrest, along with the potential interventions, remain largely elusive.

As one of the most pivotal events of the cleavage stage, maternal-to-zygotic transition (MZT) represents the developmental control of embryos shifting from the maternal to the embryonic genome. Initially, the embryonic genome stays in a transcriptionally silent state until zygotic genome activation (ZGA), which is accompanied by the decay of maternally stored mRNA. In humans, it has been shown that maternal decay defects are associated with reduced embryonic developmental potential,<sup>9</sup> suggesting that MZT abnormalities may lead to embryonic developmental arrest. However, consistent with the fact that the cleavage-stage arrest rate is much lower in mouse embryos, the key regulators of MZT appear to be poorly conserved between species,<sup>10,11</sup> making it crucial to study human-specific MZT regulatory mechanisms. Although a recent study revealed the role of tetrapeptide repeat homeobox transcription factors (TPRXs) in ZGA initiation,<sup>12</sup> the intricate regulatory networks governing human MZT remain poorly understood. As MZT is composed of the timely clearance of maternal transcripts and the generation of nascent embryonic transcripts, here, we utilize the allelic ratio to quantify this process, in an effort to identify the factors involved in human MZT and investigate the mechanisms of human cleavage-stage arrest. We identified that PAIRED (PRD)-like homeobox transcription factors (TFs), divergent paired-related homeobox (*DPRX*), and arginine-fifty homeobox (*ARGFX*), which are primate specific and expressed exclusively during preimplantation development,<sup>13</sup> played important roles in the regulation of MZT. Deficiency of these TFs was a contributing mechanism to human embryonic arrest.

## RESULT

### Allelic transcriptional profiling of human preimplantation embryo

We first collected 38 embryos spanning 6 consecutive stages from zygote to blastocyst and performed single-cell full-length transcriptome sequencing. A total number of 467 cells were retained after stringent quality control (Figures 1A, S1A, and S1B; Table S2). The blastomeres were effectively clustered according to developmental stages in principal-component analysis (PCA) (Figure S1D). To uncover the allelic transcription feature of human preimplantation embryos, we distinguished the parent-of-origin of transcripts in the embryos by single nucleotide polymorphism (SNP) allele genotyping analysis based on parental genomic information (Figure 1B). The allelic ratios of the X chromosome (chrX) exhibited different patterns between embryos from the 4-cell stage onward, reflecting the inheritance of chrX in embryos with different sexes (Figure 1C), by which we identified the sex of embryos, including 13 male embryos and 10 female embryos. The assessment of Y chromosome (chrY)-expressed genes confirmed the accuracy of sex identification (Figure S1C). Besides, the maternal allelic ratio (MAR), which represents the proportion of transcripts derived from the maternal genome, was found to be comparable between the chrX and autosomes in female embryos at the blastocyst stage (Figure 1C), indicating that X chromosome inactivation (XCI) has not been initiated yet, consistent with previous studies.<sup>14</sup> These results

demonstrate the feasibility and accuracy of the allelic information in our data, enabling an in-depth exploration of human embryogenesis mechanisms.

We next investigated the dynamics of transcripts derived from the parental genome during early human embryogenesis. A rapid decline of MAR was seen at the 4-cell and 8-cell stages (Figure 1D), indicating the initiation of the paternal genome activation at this time, which was consistent with the timing of minor and major ZGA. At the morula and blastocyst stages, the contribution of parental transcripts reached parity, with a MAR of nearly 50% (Figure 1D), indicating that the decay of maternal mRNA and the activation of the paternal genome was basically completed by the morula stage.

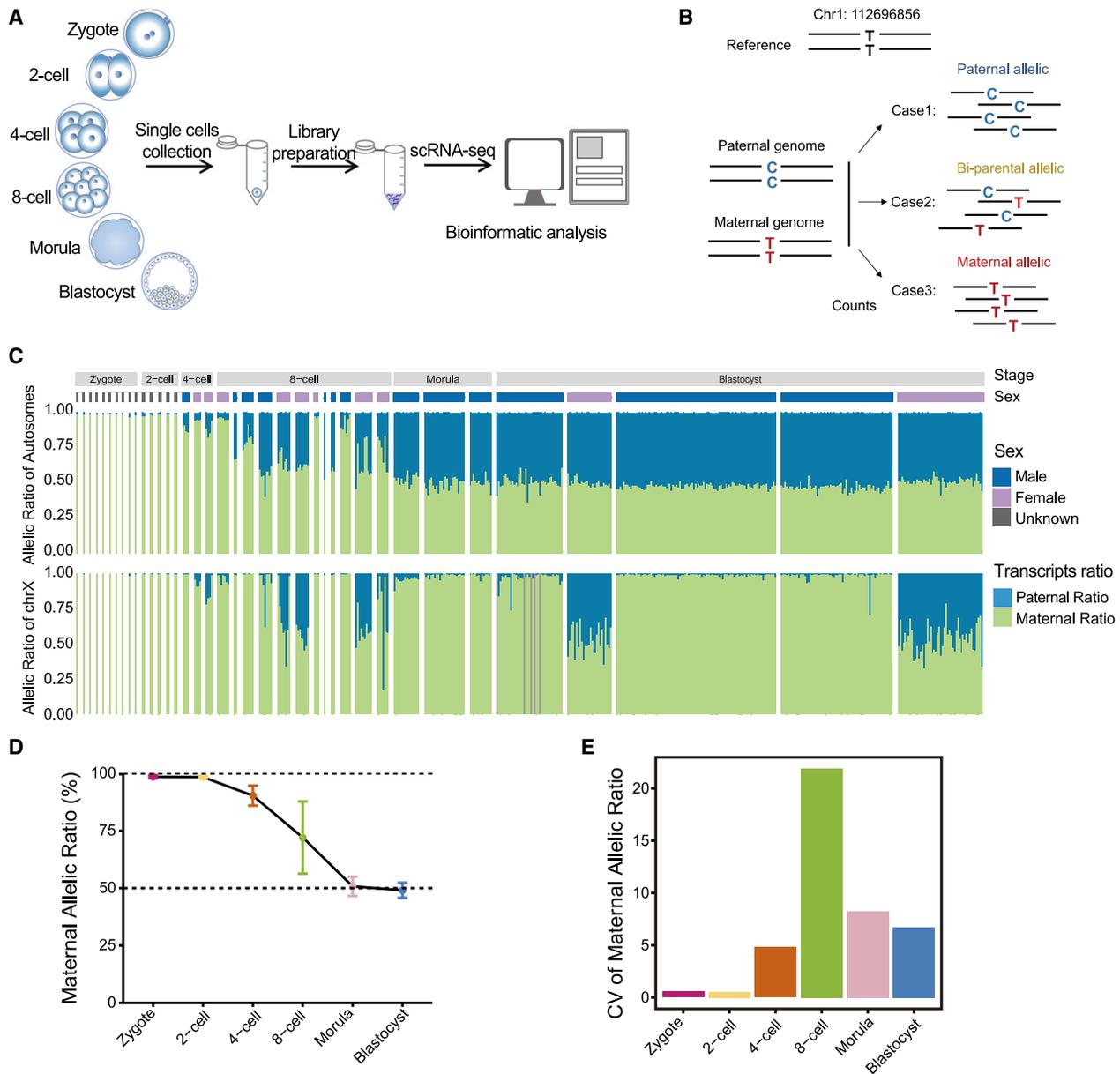
Notably, our results showed the greatest heterogeneity in MAR at the 8-cell stage among all developmental stages (Figure 1E), and the transcriptomes of 8-cell embryos spanned a long distance in the PCA analysis (Figure S1D). The fact that a substantial fraction of human embryos was arrested at the 8-cell stage in clinical ART treatment<sup>5,6,15</sup> prompted us to postulate that the differences in developmental potentials might be largely correlated with the strong heterogeneity of allelic ratio at the 8-cell stage.

### Identification of TAS of 8-cell embryos by allelic ratio

To test this speculation, we examined the allelic ratio at both embryonic and cellular levels. Although all the 8-cell embryos were collected at proper developmental time points and exhibited normal morphology (Figure S1E), the MAR varied from 53.1% to 97.4% among the 12 embryos (Figure 2A) and the MAR of blastomeres ranged from 40% to 100% (Figure S1F). Further exploration of the origins of heterogeneity indicated that inter-embryonic variation, accounting for approximately 80% of the heterogeneity (Figure 2B), rather than intra-embryonic variation, dominated MAR heterogeneity at the 8-cell stage. Therefore, we sought to determine whether the embryonic MAR is related to its transcriptional status.

As the MAR at the blastomere level fits a bimodal distribution<sup>16</sup> (Figure S1F), we determined the cutoff value of MAR with a finite-mixture model,<sup>17</sup> which divided the 8-cell embryos into two groups: the low-MAR group (8C\_E1-E7) and the high-MAR group (8C\_E8-E12) (Figure 2A). Although the two groups could not be distinguished by morphology (Figure S1E), significant differences in gene expression patterns were observed (Figures S1G and S1H; Table S3). Blastomeres with high-MAR were largely delayed along the developmental trajectory at the 8-cell stage (Figure 2C), indicating the developmental arrest status of high-MAR embryos. Furthermore, the transcriptional profile of high-MAR 8-cell embryos was much more similar to 4-cell embryos than to low-MAR 8-cell embryos (Figure 2D). Therefore, we defined these high-MAR 8-cell embryos as “transcriptionally arrested status” (TAS) embryos.

To inquire whether MZT was compromised in the TAS embryos, we comprehensively analyzed the transcriptional differences between TAS and normal 8-cell stage embryos. The expression of maternal factors and major ZGA genes (hereafter referred to as ZGA genes) was significantly altered in TAS embryos (Figure S2A); 47% of downregulated genes in TAS embryos were ZGA genes and up to 64% of ZGA genes were significantly reduced in TAS embryos (Figure 2E), while there was little overlap between the downregulated genes in TAS



**Figure 1. The allelic transcriptional characterization of human preimplantation embryos**

(A) Schematic overview of the experimental design.

(B) Schematic illustration showing the allelic transcripts identification strategy. Only SNPs that are homozygous in parents and heterozygous in embryos are used to trace allelic transcripts.

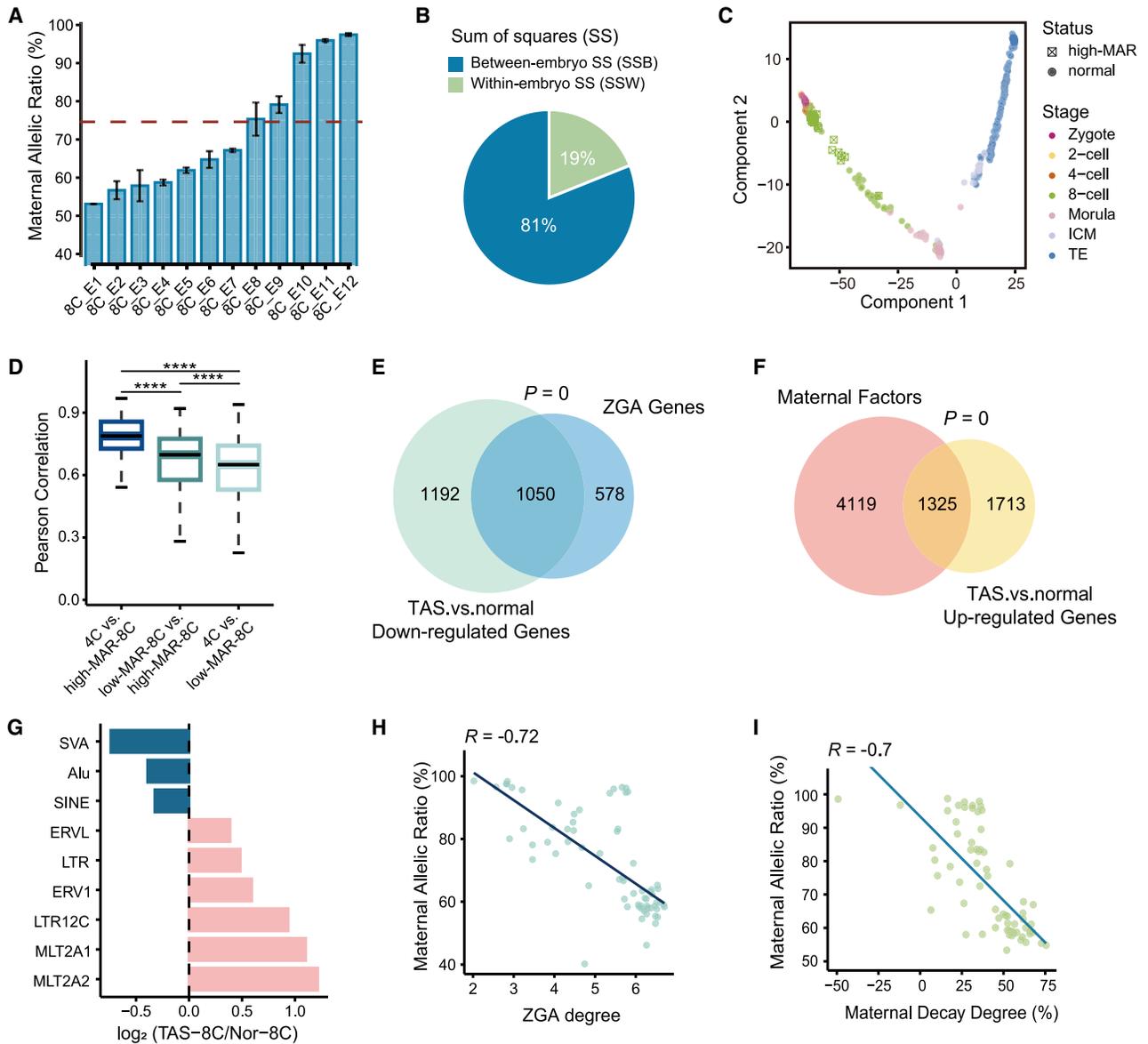
(C) Bar plots showing the parental allelic ratio of autosomes (up) and chrX (down) of preimplantation embryos at single-cell scale. Cells with less than 100 counts on the chrX are marked in gray.

(D) Line chart showing the mean MAR of autosomes at different developmental stages. Error bars represent the standard deviation of the mean.

(E) The coefficient of variation (CV) of MAR of autosomes at different developmental stages.

embryos and maternal factors (Figure S2B). 44% of the upregulated genes were maternal factors (Figure 2F), with none belonging to ZGA genes (Figure S2C). Meanwhile, minor ZGA genes were barely repressed in TAS embryos (Figure S2D), suggesting that the abnormalities of TAS embryos are not due to the defects in minor ZGA. To discern underlying functional differences between normal and TAS 8-cell embryos (TAS-8C), we performed functional enrichment analysis of

differentially expressed genes (DEGs). It was suggested that the energy metabolism, nucleic acid, and protein-synthesis-related categories were severely compromised in TAS embryos (Figure S2E). Meanwhile, actin-filament-based processes, female gamete generation, and the Wnt signaling pathway were enriched for upregulated genes in TAS embryos (Figure S2F), as a result of the failure of maternal genes degradation.<sup>9,18</sup>

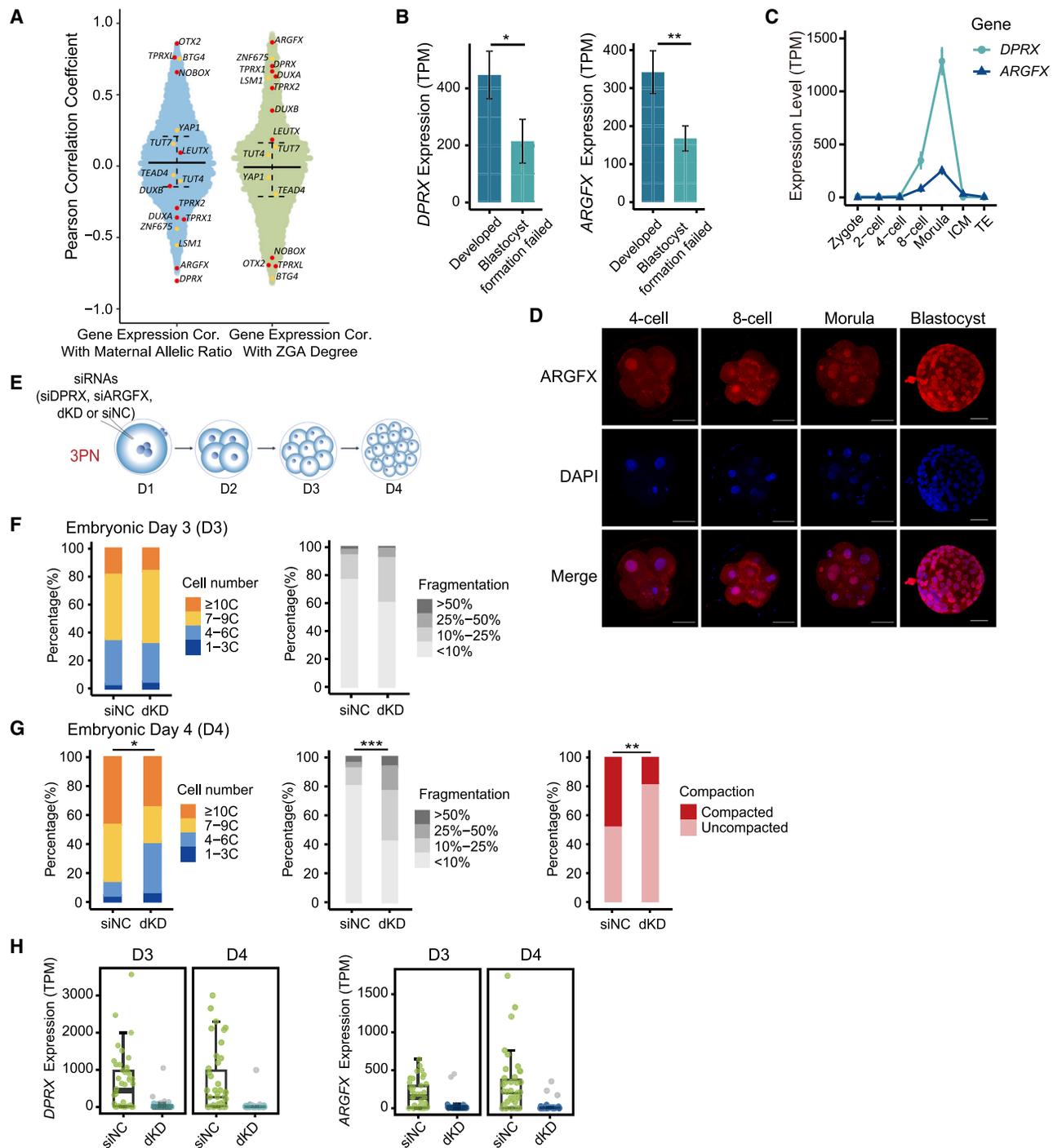


**Figure 2. Characteristics of TAS embryos**

(A) Bar plots showing the median MAR of autosomes of 8-cell stage embryos. Error bars represent the standard error of the median.  
 (B) Pie chart showing the composition of MAR heterogeneity at the 8-cell stage.  
 (C) Monocle pseudotime trajectory showing the development progression of preimplantation embryos. The high-MAR blastomeres are marked in the square box.  
 (D) Boxplot showing the Pearson correlation between 4-cell, high-MAR and low-MAR 8-cell embryos. Statistical analysis was performed using Wilcoxon test. \*\*\*\* $p < 0.0001$ .  
 (E) Venn diagram showing the overlap between ZGA genes and downregulated genes in TAS embryos, the  $p$  value (hypergeometric test) for the overlap is shown.  
 (F) Venn diagram showing the overlap between maternal factors and upregulated genes in TAS embryos, the  $p$  value (hypergeometric test) for the overlap is shown.  
 (G) Bar plot showing the expression changes of representative transposons between normal-8C and TAS-8C embryos.  
 (H) Correlation analysis between the MAR and the expression level of ZGA genes (ZGA degree) in blastomeres of 8-cell stage.  $R = -0.72$ ,  $p = 2.6 \times 10^{-12}$ .  
 (I) Correlation analysis between the MAR and the degree of maternal decay in blastomeres of 8-cell stage.  $R = -0.7$ ,  $p = 1.8 \times 10^{-11}$ .

To further examine the defects of TAS embryos, we applied the Fuzzy C-Means algorithm<sup>19</sup> to cluster up- and downregulated genes, respectively, according to their expression patterns (Figures S2G and S2H). In addition to the failure of maternal mRNA decay (cluster 2 in Figure S2G) and ZGA (clusters 2 and 3 in Figure S2H), the expressions of some lineage-related genes,

such as *TEAD3*, *NANOG*, *KLF4*, *GATA6*, and *GATA2*, were also altered (Figures S2G and S2H), suggesting impairment in subsequent lineage specification. Notably, a subset of genes that are highly expressed at the 4-cell stage, including *LEUTX*, *CCNA1*, *KDM4E*, and *PRAMEF2*, were abnormally prolonged in their activation in TAS embryos (cluster 3 in Figures S2G and S2I).



**Figure 3. Double knockdown of *DPRX* and *ARGFX* resulted in the reduction of embryo developmental potential**

(A) Violin plot showing the Pearson correlation coefficient of gene expression and MAR of all genes (left) and the Pearson correlation coefficient of gene expression and ZGA degree of all genes (right). The PRD-like TFs are marked in red, and other reported MZT factors are marked in yellow.

(B) The expression of *DPRX* and *ARGFX* in blastomeres at 8-cell stage with different developmental potentials. Developed group represents embryos that grew into blastocysts successfully ( $n = 11$ ); blastocyst formation failed group represents embryos that failed to form blastocysts ( $n = 11$ ). Error bars represent the standard error of the mean. Statistical analysis was performed using Wilcoxon test.  $*p < 0.05$ ;  $**p < 0.01$ . The published data were used.<sup>9</sup>

(C) The expression levels of *DPRX* and *ARGFX* during human early embryo development. Error bars represent the standard error of the mean.

(D) Immunostaining images of *ARGFX* at human cleavage and blastocyst stages. Scale bar, 50  $\mu\text{m}$ .

(E) Schematic illustration of gene knockdown experiment. Target gene siRNAs (*DPRX* or/and *ARGFX*) or negative control (NC) siRNA was injected into human 3PN embryos on embryonic day 1, and embryonic developmental status was recorded on embryonic day 2, 3, and 4.

(legend continued on next page)

Transposable elements (TEs), which account for nearly half of the human genome, have been shown to exhibit dynamic changes around MZT, potentially serving as regulatory elements during human embryogenesis and influencing transcriptional networks.<sup>20–22</sup> We further investigated the changes in TE expression in TAS embryos (Figures 2G and S2J). As expected, SINE-VNTR-Alu (SVA) elements, which are hominid-specific and act as enhancers for ZGA genes,<sup>23</sup> were significantly downregulated in TAS embryos. The SINE Alu retrotransposons, which were found to be the main targets of pioneer factors in mouse ZGA,<sup>10</sup> were also markedly repressed in TAS embryos. In addition, *MLT2A1*, *MLT2A2*, and *LTR12C*, which are highly expressed at the 8-cell stage, were abnormally activated in TAS embryos compared with normal 8-cell embryos, indicating the presence of aberrant activation. Taken together, our results identify that a fraction of human 8-cell embryos are morphologically normal but in a state of transcriptional arrest, characterized by extensive impairment of MZT, defects in subsequent lineage differentiation, and aberrant activation of some early embryo-specific genes.

### **DPRX and ARGFX are essential for early embryonic development**

Maternal mRNA decay and ZGA are indispensable to each other and cooperate to complete MZT<sup>24,25</sup>; however, the exact correlation between them was not clear. Leveraging our data, which enables accurate quantification of parental transcripts in biparental embryos, we performed analyses to explore the correlation between allelic ratio, the degree of ZGA, and maternal decay in 8-cell blastomeres. It was found that the degree of maternal decay was positively correlated with ZGA degree (Figure S3A), and MAR was highly negatively correlated with both the degree of ZGA and maternal decay (Figures 2H and 2I), demonstrating that allelic ratio is an accurate indicator to assess the degree of MZT.

Therefore, we postulated that the factors of the MZT process should have strong correlations with the allelic ratio. We analyzed the correlation coefficients between MAR and the expression levels of all genes at the 8-cell stage (Figure 3A, left). As expected, the expression levels of maternal factors that regulate MZT, including *BTG4* and *TPRXL*,<sup>12,25</sup> showed strong positive correlation with MAR. Conversely, ZGA genes that regulate MZT, such as *TPRX1/2*, *LSM1*, and *ZNF675*,<sup>12,26</sup> were highly negatively correlated with MAR.

Among all genes, *DPRX* showed the strongest negative correlation between expression level and MAR (Figures 3A and S3C, left), which was also one of the most significantly downregulated genes in TAS 8-cell embryos (Figure S3B, left). *DPRX* belongs to the PRD-like homeobox TF family.<sup>27</sup> The other gene in this family, *ARGFX*, which was the second strongest TF negatively correlated with the MAR (Figures 3A and S3C, right), also failed

to be normally activated in the TAS embryos (Figure S3B, right). Further evaluation of the correlation between gene expression levels and ZGA degree indicated that both *DPRX* and *ARGFX* exhibited a highly positive correlation with ZGA degree (Figures 3A and S3D). To clarify whether *DPRX* and *ARGFX* affect the developmental potential of the embryos, we used published data of single blastomeres biopsied from 8-cell embryos whose developmental fate was followed after biopsy.<sup>9</sup> Importantly, the expression of *DPRX* and *ARGFX* was significantly higher in the blastocyst formation group than in the blastocyst formation failure group, indicating that the defective activation of *DPRX* and *ARGFX* was associated with impaired embryonic developmental potential (Figure 3B). Based on these findings, we postulated that *DPRX* and *ARGFX* might serve as essential factors of human MZT, and further impact embryonic developmental potential.

Both *DPRX* and *ARGFX* are highly transcribed near the major ZGA, with a peak at the morula stage (Figure 3C). The nuclear localization of *ARGFX* in early embryos was confirmed by immunofluorescence staining (Figure 3D). *DPRX* localization could not be detected due to a lack of commercial antibodies. The roles of *ARGFX* and *DPRX* have only been studied by ectopic overexpression in human embryonic stem cells (hESCs).<sup>13</sup> However, both genes are expressed exclusively in preimplantation embryos in humans.<sup>13</sup> Considering that TFs necessitate a contextually specific genomic environment to execute their functions,<sup>28</sup> it is essential to evaluate their roles directly in preimplantation embryos.

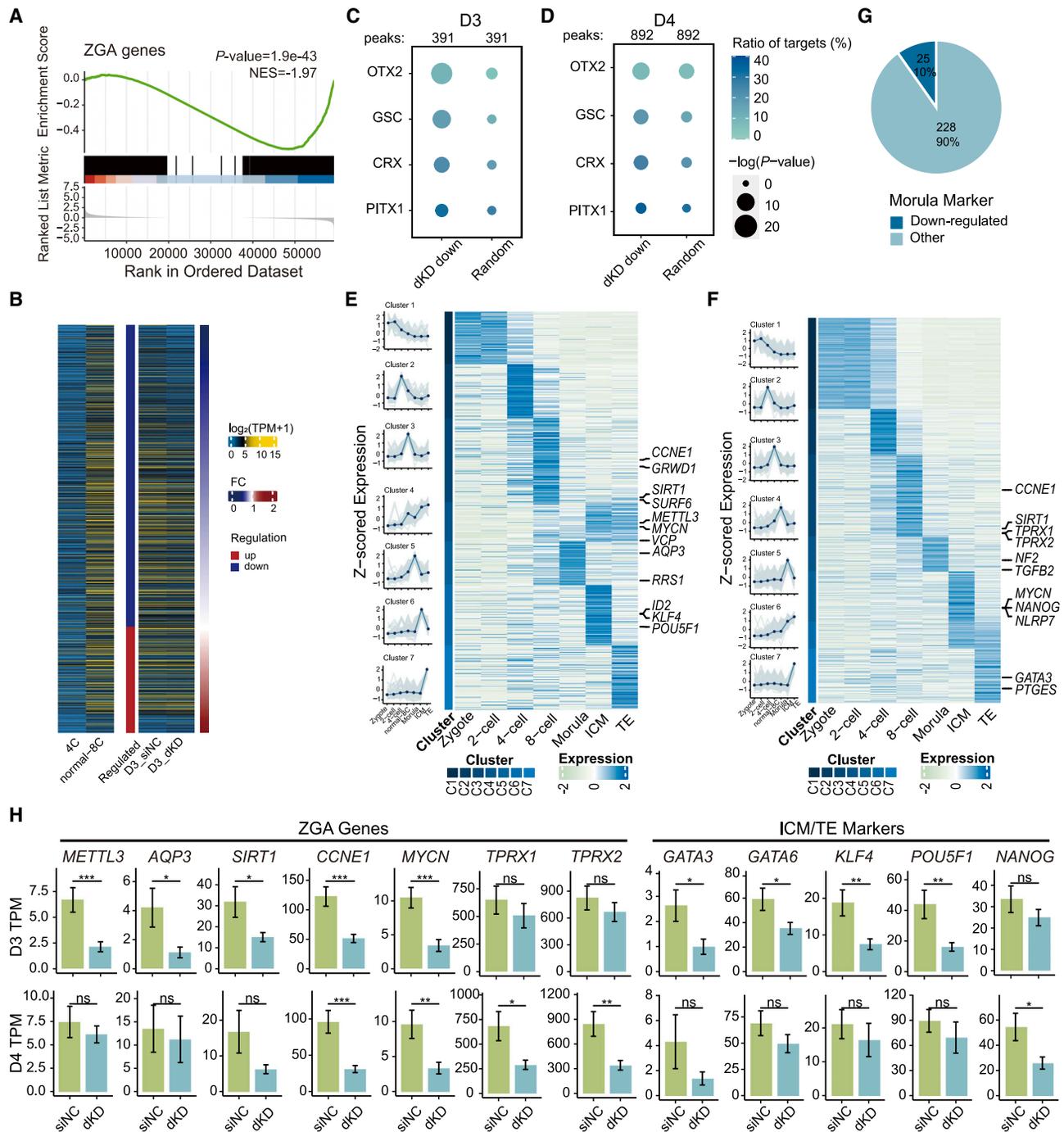
We further knocked down these two genes by injecting small interfering RNAs (siRNAs) in human tri-pronuclear (3PN) zygotes (Figure 3E). The high knockdown efficiency was verified by quantitative real-time PCR (Figures S3E and S3F). Embryo development after *DPRX* or *ARGFX* knockdown was evaluated at embryonic day 3 and 4 based on blastomere numbers, fragmentation degree, and whether they were compacted or not. However, compared with the control embryos (microinjected with negative control siRNA [siNC]), there was only a slight perturbation in embryo developmental potential after *DPRX* or *ARGFX* knockdown (si*DPRX* or si*ARGFX*) separately (Figures S3G and S3H). Because *DPRX* and *ARGFX* are both rapidly evolving genes from the same internal node,<sup>27,29</sup> we asked whether there was a functional redundancy between *DPRX* and *ARGFX*, with developmental deficits being rescued by the other gene after a single knockdown. We then performed single-cell RNA sequencing (scRNA-seq) for si*DPRX* or si*ARGFX* embryos (Figure S3I). As expected, there was a high degree of overlap between DEGs of the two groups (Figures S3J and S3K), suggesting a redundant regulation between *DPRX* and *ARGFX*.

Therefore, we performed the knockdown of *DPRX* and *ARGFX* (double KD [dKD]) simultaneously in 3PN zygotes (Figure S4A; Table S4). Notably, the development of dKD embryos was severely impaired on embryonic day 4, as indicated by a

(F) Embryo phenotypes of *DPRX* and *ARGFX* double knockdown (dKD) on embryonic day 3. Left showing the cell number of embryos in the control and dKD groups. Right showing the degree of fragmentation of the control and dKD groups.  $n$  (dKD) = 81,  $n$  (siNC) = 96.

(G) Embryo phenotypes of dKD embryos on embryonic day 4. Left showing the cell number of embryos in the control and dKD group. Middle showing the degree of fragmentation of the control and dKD group. Right showing the compaction rate of embryos in the control and dKD group.  $n$  (dKD) = 46,  $n$  (siNC) = 49.  $p$  values for cell number (left): Wilcoxon test,  $*p < 0.05$ ; fragmentation (middle): Wilcoxon test,  $***p < 0.001$ ; compaction rate (right): chi-squared test,  $**p < 0.01$ .

(H) Verification of knockdown efficiency by scRNA-seq. Boxplots showing the expression levels of *DPRX* and *ARGFX* in siNC and dKD embryos at embryonic day 3 and 4. Dots representing single blastomeres; samples with low knockdown efficiency were shown in gray dots and discarded in subsequent analyses.



**Figure 4. *DPRX* and *ARGFX* are required for the regulation of major ZGA and subsequent lineage differentiation**

- (A) GSEA for ZGA genes with respect to the global transcriptional changes observed in dKD embryos on embryonic day 3.  
 (B) Heatmap showing the expression of major ZGA genes in dKD embryos and siNC embryos on embryonic day 3. The upregulated genes (26%) are marked in red, and the downregulated genes (74%) are marked in blue.  
 (C) PRD-like TF-binding motifs enrichment in the distal open chromatin regions of downregulated genes upon *DPRX* and *ARGFX* double knockdown on embryonic day 3.  
 (D) PRD-like TF-binding motifs enrichment in the distal open chromatin regions of downregulated genes upon *DPRX* and *ARGFX* double knockdown on embryonic day 4.  
 (E) Clusters of downregulated genes in dKD embryos on embryonic day 3 showing different expression patterns during human preimplantation embryos.  
 (F) Clusters of downregulated genes in dKD embryos on embryonic day 4 showing different expression patterns during human preimplantation embryos.

(legend continued on next page)

significant reduction in the blastomere number and compaction rate, along with a significant increase in the fragmentation rate (Figure 3G), whereas an elevated fragmentation rate was observed on embryonic day 3 (Figure 3F). These results illustrate a synergistic effect of *DPRX* and *ARGFX* in regulating embryonic development during the critical stages of preimplantation.

### ***DPRX* and *ARGFX* regulate ZGA and subsequent lineage specification**

scRNA-seq was further performed on embryonic day 3 and 4 siNC and dKD embryos. Blastomeres of dKD embryos with low knockdown efficiency were excluded from subsequent analyses (Figure 3H). DEGs of dKD embryos on embryonic day 3 showed significant overlap with those between TAS and normal 8-cell embryos (Figures S4B and S4C), and the changes in TEs displayed a resemblance to the TAS 8-cell embryos (Figure S4D), indicating that the knockdown of *DPRX* and *ARGFX* can mimic the phenotype of TAS at the 8-cell stage to some extent.

We then sought to probe how major ZGA was affected upon the loss of *DPRX* and *ARGFX*. Gene set enrichment analysis (GSEA) results indicated that ZGA genes were significantly repressed in dKD embryos on embryonic day 3 (Figure 4A). Of the 1,160 ZGA genes with an expression level (transcript per million [TPM]) greater than 1 in 8-cell embryos, the majority failed to be activated to normal levels and ~8% (85 of 1,160) showed more than 2-fold downregulation (Figure 4B), indicating that dKD embryos exhibit impairment in major ZGA. In addition, by performing the nucleosome occupancy and methylome sequencing (NOME-seq) on siNC and dKD embryos, we identified the putative enhancers (the nearest distal open chromatin regions) based on the detected nucleosome-depleted regions (NDRs). Motif enrichment analysis suggested that, compared with random genes, PRD-like TF motifs (OTX2, GSC, CRX, and PITX1) tended to be more enriched in the putative enhancers of downregulated genes in dKD embryos on both embryonic day 3 and 4 (Figures 4C and 4D), confirming the specificity of knockdown perturbation.

We then classified the downregulated DEGs in dKD embryos on embryonic day 3 into seven clusters, corresponding to the seven key preimplantation developmental stages (Figure 4E). ~22% of the downregulated DEGs were 8-cell specific, including *SIRT1*, *CCNE1*, and *SURF6* (cluster 3 in Figures 4E and 4H), which was consistent with the deficiency of major ZGA. Notably, ~50% of DEGs were highly expressed in morula and blastocysts, including *ID2*, *POU5F1*, *METTL3*, *VCP*, *GATA3*, *GATA6*, and *AQP3* (clusters 4–7 in Figures 4E and 4H), suggesting that a proportion of lineage-specification-related genes were compromised in dKD embryos. To further clarify the impact of *DPRX* and *ARGFX* knockdown on embryonic developmental potential, detailed analysis of downregulated DEGs in dKD embryos on embryonic day 4 was performed (Figure 4F). A subset of 8-cell-specific genes, including *TPRX1* and *TPRX2*, remained in activation failure on embryonic day 4 (cluster 3 in Figures 4F and 4H), indicating the persistence of ZGA dysfunction or a delay in its completion (Figure S4E). ~10%

(25 of 253) of morula-specific genes failed to be properly activated (Figure 4G), and 47% of downregulated DEGs were morula- and blastocyst-stage specific (clusters 4–7 in Figure 4F). Moreover, GSEA results showed defective activation of inner cell mass (ICM) and trophectoderm (TE) markers in dKD embryos on embryonic day 4 (Figure S4F), including key ICM markers *NANOG* and *NLRP7* as well as TE markers *GATA3* and *PTGES* (clusters 6–7 in Figures 4F and 4H). These data demonstrate that *DPRX* and *ARGFX* not only play important roles in the proper activation of major ZGA but also influence the expression of key lineage specification genes.

We next investigated whether maternal mRNA decay was also affected in dKD embryos. As expected, the expression of maternal factors was significantly higher in the dKD group on embryonic day 3 (Figure S4G). However, of all the genes upregulated upon dKD, only ~20% of the upregulated genes were maternal factors on both embryonic day 3 and 4 (cluster 1 in Figures S4H and S4I), indicating that, in addition to the impairment of maternal mRNA decay, *DPRX* and *ARGFX* deficiency also resulted in aberrant activation of other genes, including 4-cell-specific genes *TRIM49* and *TRIM64* as well as lineage-specification-related genes *KRT19* and *FGF4*.

### ***DPRX* and *ARGFX* deficiency resulted in aberrant retention of acetylation in early human embryo**

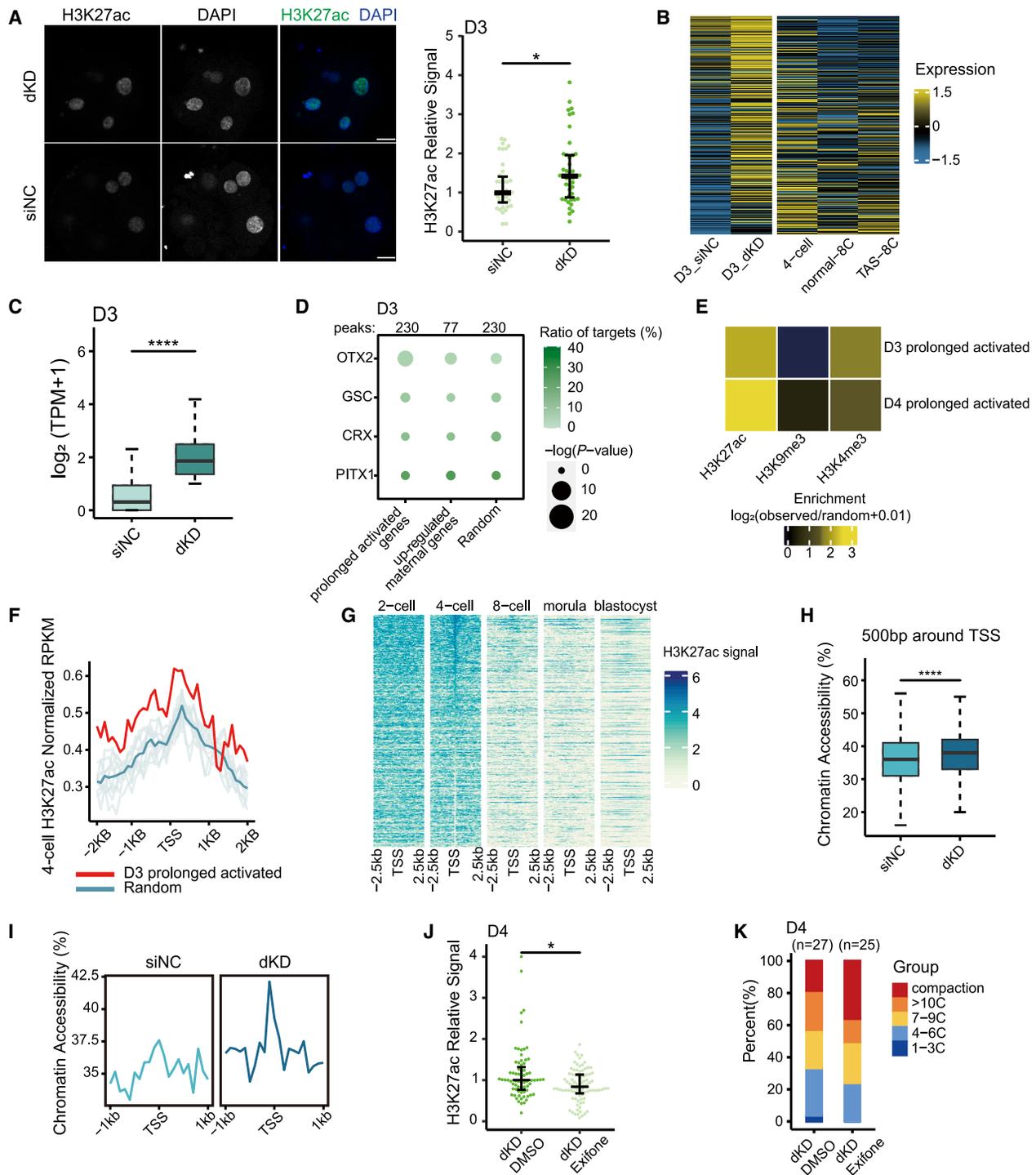
We next investigated whether other defects were present in dKD embryos. Notably, we found that the histone deacetylases (HDACs) pathway was the most significantly enriched category among the downregulated genes in dKD embryos on embryonic day 4 (Figure S5A). Coincidentally, critical deacetylases HDAC1, HDAC2, and *SIRT1*<sup>30,31</sup> failed to be normally activated upon *DPRX* and *ARGFX* deficiency (Figure S5B). Given that HDACs are required for the deacetylation of H3K27ac, which is indispensable for human ZGA,<sup>32</sup> we questioned whether this process was interrupted in dKD embryos. Immunofluorescence staining identified a significantly elevated level of H3K27ac signals in dKD embryos compared with siNC (Figures 5A and S5C), confirming that dKD disrupted the deacetylation of H3K27ac around ZGA, resulting in its aberrant retention.

We further probed whether the aberrant retention of acetylation led to transcriptional overactivation. We focused on the genes upregulated upon dKD in addition to maternal genes. This subset of genes, which should be highly expressed at the 4-cell stage and repressed at the normally developing 8-cell stage, exhibited robust activation upon dKD on embryonic day 3 (Figures 5B and 5C). Similar observations persist through embryonic day 4 (Figures S5D and S5E). We identified this subset of genes as prolonged activated genes. Of note, in dKD embryos, the prolonged activated genes, rather than the upregulated maternal genes, were enriched for PRD-like TF motifs in their putative enhancers (Figures 5D and S5F), demonstrating the specificity of the dKD effect on prolonged activated genes.

By analyzing published histone modification data of human early embryos,<sup>32,33</sup> we found that the putative enhancers of the

(G) Pie chart showing the percentage of morula markers among the genes downregulated after *DPRX* and *ARGFX* double knockdown on embryonic day 4.

(H) The expression levels of representative genes after *DPRX* and *ARGFX* double knockdown on embryonic day 3 (the up row) and embryonic day 4 (the down row). Embryonic day 3: *n* (siNC) = 46, *n* (dKD) = 46. Embryonic day 4: *n* (siNC) = 40, *n* (dKD) = 32. Error bars represent the standard error of the mean. Statistical analysis was performed using unpaired two-sample t test. \**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001; ns, not significant.



**Figure 5. *DPRX* and *ARGFX* double knockdown led to abnormal retention of H3K27ac**

(A) Representative immunostaining images of H3K27ac (left) and quantification of H3K27ac fluorescence intensity (right) in dKD and siNC embryos on embryonic day 3. Each dot represented a nucleus. Scale bar, 20  $\mu$ m.  $n$  (siNC) = 40,  $n$  (dKD) = 43. Statistical analysis was performed using Wilcoxon test. \* $p$  < 0.05.

(B) Heatmap showing the expression pattern of prolonged activated genes after *DPRX* and *ARGFX* double knockdown on embryonic day 3.

(C) The expression level of prolonged activated genes in siNC and dKD embryos on embryonic day 3. Statistical analysis was performed using Wilcoxon test. \*\*\*\* $p$  < 0.0001.

(D) PRD-like TF motifs enrichment in the distal open chromatin regions of prolonged activated genes and upregulated maternal genes on embryonic day 3 after *DPRX* and *ARGFX* double knockdown.

(E) The enrichment of histone modifications in the distal open chromatin regions of prolonged activated genes after *DPRX* and *ARGFX* double knockdown on embryonic day 3 and 4. The published human histone modification ChIP-seq data<sup>23,32,33</sup> was used.

(legend continued on next page)

prolonged activated genes were significantly enriched in H3K27ac peak of wild-type 4-cell embryos rather than H3K9me3 (Figure 5E). There was also some enrichment in H3K4me3, which was in line with the high intersection between the H3K27ac and H3K4me3 domains at the 4-cell stage.<sup>32</sup> Moreover, prolonged activated genes showed stronger H3K27ac signals around their transcription start sites (TSSs) compared with random genes (Figure 5F), which would exhibit deacetylation at the 4-cell to 8-cell stage during normal development (Figure 5G). In addition, promoters of these prolonged activated genes showed higher levels of accessibility in dKD than in siNC embryos (Figures 5H, 5I, and S5G). These results support the notion that the prolonged transcriptional activation is a consequence of the aberrant retention of H3K27ac due to the deficiency of *DPRX* and *ARGFX*.

Notably, aberrant prolonged activation was also observed in TAS embryos (Figures S2I and S5H), which were also deficient in *DPRX* and *ARGFX*. We then investigated whether a parallel mechanism operates in TAS-8C. Consistent with dKD embryos, prolonged activated genes in TAS-8C also showed enrichment for PRD-like TF motifs in their putative enhancers (Figure S5I). Deacetylases of the histone modification H3K27ac were not properly activated in TAS embryos (Figure S5J), and the prolonged activated genes of TAS-8C also showed an increased H3K27ac signal around TSSs, which displayed similar deacetylation from the 4-cell to 8-cell stage (Figures S5K and S5L). Hence, these data suggest the analogous presence of aberrant acetylation retention in TAS-8C.

The above findings prompted us to explore the interventions for cleavage-stage arrest. We further evaluated whether the embryonic developmental arrest could be rescued by compensating for the deficit in deacetylation upon *DPRX* and *ARGFX* defects. Because *HDAC1*, which is indispensable for the erasure of H3K27ac, showed a significant downregulation in both TAS-8C and dKD embryos and, notably, we found that *HDAC1* had a considerably high number of PRD-like TF motifs within its upstream regulatory region among all genes, ranking in the top 93.23% (Figure S5M), suggesting that it may be directly regulated by *DPRX* and *ARGFX*. Therefore, we used exifone, which can bind *HDAC1* and directly induce *HDAC1*-mediated deacetylation,<sup>34</sup> to treat the dKD 3PN embryos from embryonic day 1 to 4 (Figure S6A; Table S5). As expected, the exifone-treated group showed significantly weaker H3K27ac signals compared with the control group (DMSO-treated) (Figures 5J and S6B). Importantly, exifone treatment prominently increased the cell number and compaction rate of embryos (18.5% in DMSO-treated group; 36% in exifone-treated group), rescuing the dKD impairment to some extent (Figure 5K).

Given that the aberrant H3K27ac status also exists in normally developing 8-cell embryos, we further extended our findings to embryonic development under physiological conditions.

Remarkably, by comparing the development of NC 3PN embryos with and without exifone, we found that exifone treatment indeed improved the embryonic developmental potential (Figures S6C and S6D). In sum, our findings identified that the deficiency of *DPRX* and *ARGFX* contributes to human cleavage-stage arrest and that mitigating the acetylation retention resulting from their defects has the potential to alleviate human embryonic developmental arrest (Figure 6).

## DISCUSSION

Cleavage-stage embryonic arrest is a relatively severe phenomenon in humans, in which approximately 50%–60% of 8-cell embryos with good morphology fail to develop to the blastocyst stage.<sup>35,36</sup> It can be seen that the 8-cell stage to the formation of morula is a huge hurdle for embryo development. However, the unique mechanisms responsible for this phenomenon in human embryos remain largely unknown. There have been several studies on the characteristics of arrested embryos<sup>5,8,25</sup>; however, the arrested 8-cell embryos in these studies were collected with additional days of *in vitro* culture, which introduces confounding factors and increases the imprecision of the conclusions. Our study utilized normally developing 8-cell embryos, identifying that up to 42% of 8-cell embryos at proper developmental time points are in a state of transcriptional arrest despite appearing morphologically normal, consistent with the high incidence of cleavage-stage arrest observed in clinically ART treatment.

As one of the most important events in the cleavage stage, the MZT process can be precisely quantified by allelic ratio in our data. Based on this, we identified two essential TFs that regulate MZT, *DPRX* and *ARGFX*. The other members of the PRD-like TFs, *TPRXs*, have been proven as important factors of human ZGA.<sup>12</sup> Our results further identified the impact of *DPRX* and *ARGFX* on MZT and lineage specification in human embryo development. Notably, recent studies showed that PRD-like homeobox TFs *Obox* play an important role in mouse ZGA.<sup>11</sup> There is strong redundancy among the oocyte-specific homeobox (*OBOX*) family in mice, as evidenced by the fact that the existence of either maternal or zygotic *OBOX* can support the normal development of mouse embryos.<sup>11</sup> However, such redundant regulation appears to be much weaker in human embryos, where knockdown of PRD-like homeobox major ZGA TFs *DPRX* and *ARGFX* is sufficient to cause developmental arrest in human embryos. We speculate that such a protective redundancy regulation mechanism is relatively lacking in human embryo development.

Given the specificity of *DPRX* and *ARGFX* expression in early human embryos, their deficiency may be a pivotal factor contributing to human embryonic-cleavage-stage arrest. Of note, we observed the presence of aberrant retention of H3K27ac modification in human embryonic arrest, and erasing H3K27ac retention was able to rescue the cleavage-stage arrest to some

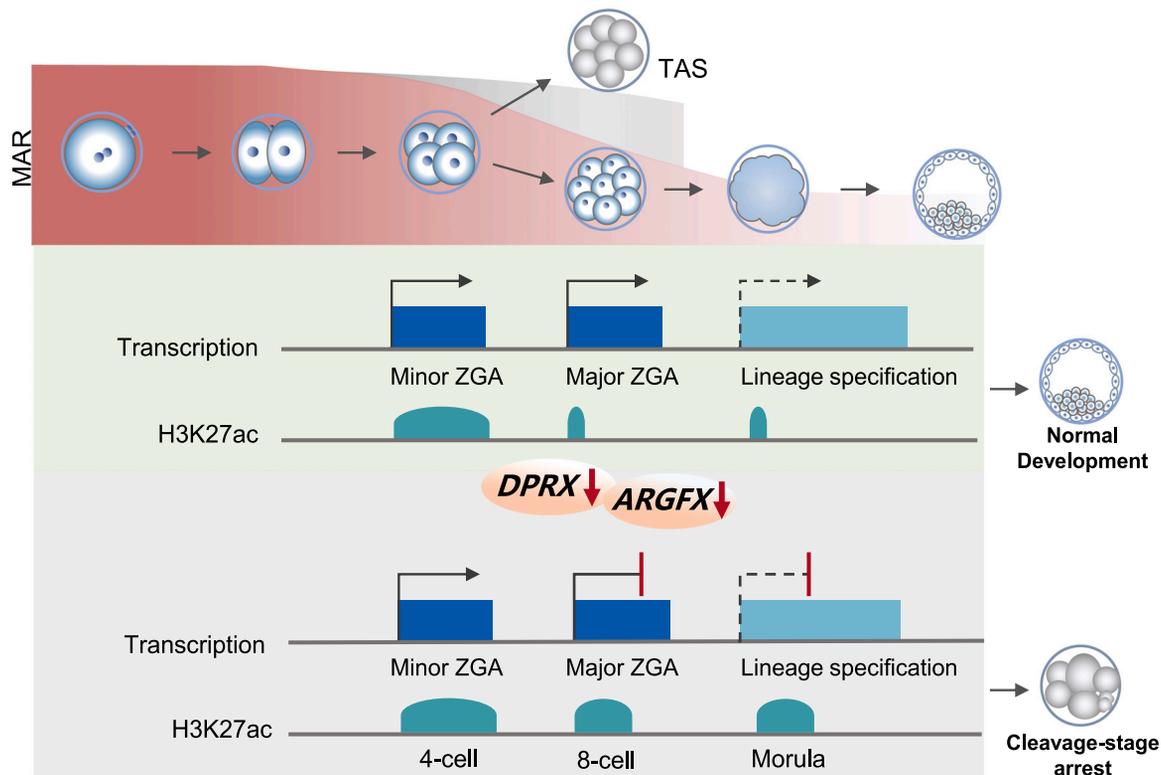
(F) The H3K27ac level around TSSs of prolonged activated genes in siNC and dKD embryos on embryonic day 3.

(G) Heatmap showing H3K27ac signals around TSS regions of prolonged activated genes at different human embryo development stages.

(H and I) Boxplots (H) and line chart (I) showing the GCH methylation level around TSS region of prolonged activated genes in siNC and dKD embryos on embryonic day 3. statistical analysis for (H) was performed using Wilcoxon test. \*\*\*\* $p < 0.0001$ .

(J) Quantification of H3K27ac fluorescence intensity in dKD embryos on embryonic day 4 with DMSO or exifone treatment. Each dot represented a nucleus.  $n$  (DMSO treatment) = 71,  $n$  (exifone treatment) = 86. Statistical analysis was performed using Wilcoxon test. \* $p < 0.05$ .

(K) Embryo phenotypes of dKD embryos on embryonic day 4 with DMSO or exifone treatment.



**Figure 6. A model illustrating the role of *DPRX* and *ARGFX* in human embryonic-cleavage-stage arrest**

The red curve represents the trend of MAR during normal human embryonic development. However, among the 8-cell embryos at the proper developmental time point, there exists a portion of embryos with high-MAR (gray curve), which are in TAS. Through further analysis of the correlation between MAR and gene expression levels, we identified the factors in MZT, *DPRX*, and *ARGFX* whose deficiencies lead to MZT failure, lineage specification impairment, and aberrant retention of H3K27ac, resulting in human embryonic-cleavage-stage arrest.

extent. We envision that our work will provide etiological insights and potential treatments for the high rates of human embryonic-cleavage-stage arrest.

### Limitations of the study

Our study identifies the factors of MZT and lineage specification, *DPRX* and *ARGFX*, in human embryos, and further validates that their deficiency directly contributes to human cleavage-stage arrest. However, due to the scarcity of human embryos, it is an enormous challenge to further explore their regulatory mechanisms directly in human embryos, including their binding target genes. In addition, 3PN embryos exhibit relatively low developmental potential for blastocyst formation,<sup>37,38</sup> and, therefore, this sets a barrier for us to accurately assess the developmental phenotype at the blastocyst stage after *DPRX* and *ARGFX* knockdown. Moreover, although some embryos exhibit transcriptional arrest at the 8-cell stage, it is possible that this phenomenon arises from delayed development and that they may still retain the potential to progress further.

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Peng Yuan ([peng.yuan@bjmu.edu.cn](mailto:peng.yuan@bjmu.edu.cn)).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

All data have been deposited to GSA, with the accession number: GSA: HRA006449. The processed gene expression matrices have been deposited to the OMIX, with accession number: OMIX: OMIX007941. Accession codes of the published data used in this study are as follows: RNA-seq data of single blastomeres biopsied from 8-cell embryos with different developmental potentials were downloaded from HRA001473 in PRJCA006930. The chromatin immunoprecipitation sequencing (ChIP-seq) data for histone H3K27ac modifications of early human embryogenesis were downloaded from HRA002355 in PRJCA009410. The peaks of histone H3K9me3 modifications of human preimplantation embryos were downloaded from GEO: GSE176016. The peaks of histone H3K4me3 of human preimplantation embryos were downloaded from GEO: GSE124718.

This paper does not report original codes. The algorithms used are listed in STAR Methods and are all publicly available. All the other data supporting the findings of this study are available from the corresponding author upon reasonable request.

### ACKNOWLEDGMENTS

We thank the Peking University Third Hospital for supporting this study, all members in Qiao and Yan laboratories for their discussions and comments during manuscript writing, and all the patients who participated in this research. This work was supported by National Natural Science Foundation of China (82125013 and 82288102), China National Key Research and Development Program (2019YFA0801400 and 2022YFC2702200), and the Young

Elite Scientists Sponsorship Program by China Association for Science and Technology (YESS20200398).

### AUTHOR CONTRIBUTIONS

P.Y., L.Y., and J.Q. conceived and designed the project; P.Y., L.Y., J.Q., and Z.Y. supervised the project; Q.G. and F.Z. performed embryo experiments with help from D.L., M.L., and Y.L.; F.X. and S.S. performed bioinformatics analysis with help from P.Y., Z.Y., and S.K.; Q.G., N.W., C.L., and X.W. performed microinjection; Q.G., F.X., and L.D. performed immunofluorescence staining and quantitative real-time PCR; Q.G. and F.X. performed full-length RNA sequencing and NOME-seq sequencing with the help of Y.X., Q.L., M.Y., Y.W., and Z.D.; Q.G., F.X., and S.S. prepared the figures and wrote the manuscript with the help of all authors.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
  - Ethics statement
- [METHOD DETAILS](#)
  - Human embryo culture
  - Target genes knockdown
  - Immunofluorescence
  - Real-time quantitative PCR
  - Isolation of single blastomeres
  - Single-cell full-length RNA sequencing
  - Single-embryo NOME-seq sequencing
  - Whole genome sequencing
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)
  - Single-cell RNA-seq data processing
  - Single-embryo NOME-seq data processing
  - DNA methylation and accessibility analysis
  - ChIP-seq data processing
  - Peak analysis
  - SNP detection of RNA and DNA sequencing data
  - Allele-specific analysis of RNA expression
  - Principal component analysis
  - Inference of embryonic sex
  - Pseudotime analysis
  - Expression estimation of transposable elements
  - Identification of nucleosome-depleted region
  - Transcription factor binding motif analysis
  - Differentially expressed genes analysis
  - ZGA genes and maternal factors identification
  - Maternal decay degree identification
  - Correlation analysis
  - Prolonged activated genes identification
  - Gene ontology analysis
  - Statistical analysis

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.devcel.2024.12.031>.

Received: June 6, 2024

Revised: September 10, 2024

Accepted: December 17, 2024

Published: January 13, 2025

### REFERENCES

1. Qiao, J., Wang, Y., Li, X., Jiang, F., Zhang, Y., Ma, J., Song, Y., Ma, J., Fu, W., Pang, R., et al. (2021). A Lancet Commission on 70 years of women's reproductive, maternal, newborn, child, and adolescent health in China. *Lancet* 397, 2497–2536. [https://doi.org/10.1016/S0140-6736\(20\)32708-2](https://doi.org/10.1016/S0140-6736(20)32708-2).
2. Vanni, V.S., Somigliana, E., Reschini, M., Pagliardini, L., Marotta, E., Faulisi, S., Paffoni, A., Vigano', P., Vegetti, W., Candiani, M., et al. (2017). Top quality blastocyst formation rates in relation to progesterone levels on the day of oocyte maturation in GnRH antagonist IVF/ICSI cycles. *PLoS One* 12, e0176482. <https://doi.org/10.1371/journal.pone.0176482>.
3. Lee, S.H., Liu, X., Jimenez-Morales, D., and Rinaudo, P.F. (2022). Murine blastocysts generated by in vitro fertilization show increased Warburg metabolism and altered lactate production. *eLife* 11, e79153. <https://doi.org/10.7554/eLife.79153>.
4. Gil, M.A., Cuello, C., Parrilla, I., Vazquez, J.M., Roca, J., and Martinez, E.A. (2010). Advances in swine in vitro embryo production technologies. *Reprod. Domest. Anim.* 45 (Suppl 2), 40–48. <https://doi.org/10.1111/j.1439-0531.2010.01623.x>.
5. Yang, Y., Shi, L., Fu, X., Ma, G., Yang, Z., Li, Y., Zhou, Y., Yuan, L., Xia, Y., Zhong, X., et al. (2022). Metabolic and epigenetic dysfunctions underlie the arrest of in vitro fertilized human embryos in a senescent-like state. *PLoS Biol.* 20, e3001682. <https://doi.org/10.1371/journal.pbio.3001682>.
6. Orvieto, R., Jonish-Grossman, A., Maydan, S.A., Noach-Hirsh, M., Dratviman-Storobinsky, O., and Aizer, A. (2022). Cleavage-stage human embryo arrest, is it embryo genetic composition or others? *Reprod. Biol. Endocrinol.* 20, 52. <https://doi.org/10.1186/s12958-022-00925-2>.
7. Shahbazi, M.N., Wang, T., Tao, X., Weatherbee, B.A.T., Sun, L., Zhan, Y., Keller, L., Smith, G.D., Pellicer, A., Scott, R.T., Jr., et al. (2020). Developmental potential of aneuploid human embryos cultured beyond implantation. *Nat. Commun.* 11, 3987. <https://doi.org/10.1038/s41467-020-17764-7>.
8. Hernandez Mora, J.R., Buhigas, C., Clark, S., Del Gallego Bonilla, R., Daskeviciute, D., Monteagudo-Sánchez, A., Poo-Llanillo, M.E., Medrano, J.V., Simón, C., Meseguer, M., et al. (2023). Single-cell multi-omic analysis profiles defective genome activation and epigenetic reprogramming associated with human pre-implantation embryo arrest. *Cell Rep.* 42, 112100. <https://doi.org/10.1016/j.celrep.2023.112100>.
9. He, Q.L., Yuan, P., Yang, L., Yan, Z.Q., Chen, W., Chen, Y.D., Kong, S.M., Tang, F.C., Qiao, J., and Yan, L.Y. (2022). Single-cell RNA sequencing reveals abnormal fluctuations in human eight-cell embryos associated with blastocyst formation failure. *Mol. Hum. Reprod.* 28, gaab069. <https://doi.org/10.1093/molehr/gaab069>.
10. Gassler, J., Kobayashi, W., Gáspár, I., Ruangroengkulrith, S., Mohanan, A., Gómez Hernández, L., Kravchenko, P., Kümmecke, M., Lalic, A., Rifel, N., et al. (2022). Zygotic genome activation by the totipotency pioneer factor Nr5a2. *Science* 378, 1305–1315. <https://doi.org/10.1126/science.abn7478>.
11. Ji, S., Chen, F., Stein, P., Wang, J., Zhou, Z., Wang, L., Zhao, Q., Lin, Z., Liu, B., Xu, K., et al. (2023). OBOX regulates mouse zygotic genome activation and early development. *Nature* 620, 1047–1053. <https://doi.org/10.1038/s41586-023-06428-3>.
12. Zou, Z., Zhang, C., Wang, Q., Hou, Z., Xiong, Z., Kong, F., Wang, Q., Song, J., Liu, B., Liu, B., et al. (2022). Translatome and transcriptome co-profiling reveals a role of TPRXs in human zygotic genome activation. *Science* 378, abo7923. <https://doi.org/10.1126/science.abo7923>.
13. Madisson, E., Jouhilahti, E.M., Vesterlund, L., Töhönen, V., Krjutskov, K., Petropoulos, S., Einarsdottir, E., Linnarsson, S., Lanner, F., Månsson, R., et al. (2016). Characterization and target genes of nine human PRD-like homeobox domain genes expressed exclusively in early embryos. *Sci. Rep.* 6, 28995. <https://doi.org/10.1038/srep28995>.
14. Zhou, F., Wang, R., Yuan, P., Ren, Y., Mao, Y., Li, R., Lian, Y., Li, J., Wen, L., Yan, L., et al. (2019). Reconstituting the transcriptome and DNA methylation landscapes of human implantation. *Nature* 572, 660–664. <https://doi.org/10.1038/s41586-019-1500-0>.

15. Scarica, C., Cimadomo, D., Dovere, L., Giancani, A., Stoppa, M., Capalbo, A., Ubaldi, F.M., Rienzi, L., and Canipari, R. (2019). An integrated investigation of oocyte developmental competence: expression of key genes in human cumulus cells, morphokinetics of early divisions, blastulation, and euploidy. *J. Assist. Reprod. Genet.* 36, 875–887. <https://doi.org/10.1007/s10815-019-01410-3>.
16. Hartigan, J.A., and Hartigan, P.M. (1985). The Dip Test of Unimodality. *Ann. Statist.* 13, 70–84. <https://doi.org/10.1214/aos/1176346577>.
17. Trang, N.V., Choisy, M., Nakagomi, T., Chinh, N.T.M., Doan, Y.H., Yamashiro, T., Bryant, J.E., Nakagomi, O., and Anh, D.D. (2015). Determination of cut-off cycle threshold values in routine RT-PCR assays to assist differential diagnosis of norovirus in children hospitalized for acute gastroenteritis. *Epidemiol. Infect.* 143, 3292–3299. <https://doi.org/10.1017/S095026881500059x>.
18. Coticchio, G., Dal Canto, M., Mignini Renzini, M., Guglielmo, M.C., Brambillasca, F., Turchi, D., Novara, P.V., and Fadini, R. (2015). Oocyte maturation: gamete-somatic cells interactions, meiotic resumption, cytoskeletal dynamics and cytoplasmic reorganization. *Hum. Reprod. Update* 21, 427–454. <https://doi.org/10.1093/humupd/dmv011>.
19. Futschik, M.E., and Carlisle, B. (2005). Noise-robust soft clustering of gene expression time-course data. *J. Bioinform. Comput. Biol.* 3, 965–988. <https://doi.org/10.1142/s0219720005001375>.
20. Göke, J., Lu, X., Chan, Y.S., Ng, H.H., Ly, L.H., Sachs, F., and Szczerbinska, I. (2015). Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells. *Cell Stem Cell* 16, 135–141. <https://doi.org/10.1016/j.stem.2015.01.005>.
21. Liu, L., Leng, L., Liu, C., Lu, C., Yuan, Y., Wu, L., Gong, F., Zhang, S., Wei, X., Wang, M., et al. (2019). An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos. *Nat. Commun.* 10, 364. <https://doi.org/10.1038/s41467-018-08244-0>.
22. Pontis, J., Planet, E., Offner, S., Turelli, P., Duc, J., Coudray, A., Theunissen, T.W., Jaenisch, R., and Trono, D. (2019). Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic Genome Activation and Control Transcription in Naive Human ESCs. *Cell Stem Cell* 24, 724–735.e5. <https://doi.org/10.1016/j.stem.2019.03.012>.
23. Yu, H., Chen, M., Hu, Y., Ou, S., Yu, X., Liang, S., Li, N., Yang, M., Kong, X., Sun, C., et al. (2022). Dynamic reprogramming of H3K9me3 at hominoid-specific retrotransposons during human preimplantation development. *Cell Stem Cell* 29, 1031–1050.e12. <https://doi.org/10.1016/j.stem.2022.06.006>.
24. Sha, Q.Q., Zhang, J., and Fan, H.Y. (2019). A story of birth and death: mRNA translation and clearance at the onset of maternal-to-zygotic transition in mammals<sup>†</sup>. *Biol. Reprod.* 101, 579–590. <https://doi.org/10.1093/biolre/iox012>.
25. Sha, Q.Q., Zheng, W., Wu, Y.W., Li, S., Guo, L., Zhang, S., Lin, G., Ou, X.H., and Fan, H.Y. (2020). Dynamics and clinical relevance of maternal mRNA clearance during the oocyte-to-embryo transition in humans. *Nat. Commun.* 11, 4917. <https://doi.org/10.1038/s41467-020-18680-6>.
26. Yuan, S., Zhan, J., Zhang, J., Liu, Z., Hou, Z., Zhang, C., Yi, L., Gao, L., Zhao, H., Chen, Z.J., et al. (2023). Human zygotic genome activation is initiated from paternal genome. *Cell Discov.* 9, 13. <https://doi.org/10.1038/s41421-022-00494-z>.
27. Booth, H.A.F., and Holland, P.W.H. (2007). Annotation, nomenclature and evolution of four novel homeobox genes expressed in the human germ line. *Gene* 387, 7–14. <https://doi.org/10.1016/j.gene.2006.07.034>.
28. Li, L., Lai, F., Liu, L., Lu, X., Hu, X., Liu, B., Lin, Z., Fan, Q., Kong, F., Xu, Q., et al. (2024). Lineage regulators TFAP2C and NR5A2 function as bipotency activators in totipotent embryos. *Nat. Struct. Mol. Biol.* 31, 950–963. <https://doi.org/10.1038/s41594-023-01199-x>.
29. Holland, P.W.H., Booth, H.A.F., and Bruford, E.A. (2007). Classification and nomenclature of all human homeobox genes. *BMC Biol.* 5, 47. <https://doi.org/10.1186/1741-7007-5-47>.
30. Li, J., Zhang, J., Hou, W., Yang, X., Liu, X., Zhang, Y., Gao, M., Zong, M., Dong, Z., Liu, Z., et al. (2022). Metabolic control of histone acetylation for precise and timely regulation of minor ZGA in early mammalian embryos. *Cell Discov.* 8, 96. <https://doi.org/10.1038/s41421-022-00440-z>.
31. Zhou, J.J., Cho, J.S., Han, H., Blitz, I.L., Wang, W., and Cho, K.W.Y. (2023). Histone deacetylase 1 maintains lineage integrity through histone acetylome refinement during early embryogenesis. *eLife* 12, e79380. <https://doi.org/10.7554/eLife.79380>.
32. Wu, K., Fan, D., Zhao, H., Liu, Z., Hou, Z., Tao, W., Yu, G., Yuan, S., Zhu, X., Kang, M., et al. (2023). Dynamics of histone acetylation during human early embryogenesis. *Cell Discov.* 9, 29. <https://doi.org/10.1038/s41421-022-00514-y>.
33. Xia, W., Xu, J., Yu, G., Yao, G., Xu, K., Ma, X., Zhang, N., Liu, B., Li, T., Lin, Z., et al. (2019). Resetting histone modifications during human parental-to-zygotic transition. *Science* 365, 353–360. <https://doi.org/10.1126/science.aaw5118>.
34. Patnaik, D., Pao, P.C., Zhao, W.N., Silva, M.C., Hylton, N.K., Chindavong, P.S., Pan, L., Tsai, L.H., and Haggarty, S.J. (2021). Exifone Is a Potent HDAC1 Activator with Neuroprotective Activity in Human Neuronal Models of Neurodegeneration. *ACS Chem. Neurosci.* 12, 271–284. <https://doi.org/10.1021/acscchemneuro.0c00308>.
35. Speyer, B., O'Neill, H., Saab, W., Seshadri, S., Cawood, S., Heath, C., Gaunt, M., and Serhal, P. (2019). In assisted reproduction by IVF or ICSI, the rate at which embryos develop to the blastocyst stage is influenced by the fertilization method used: a split IVF/ICSI study. *J. Assist. Reprod. Genet.* 36, 647–654. <https://doi.org/10.1007/s10815-018-1358-3>.
36. Van Landuyt, L., De Vos, A., Joris, H., Verheyen, G., Devroey, P., and Van Steirteghem, A. (2005). Blastocyst formation in in vitro fertilization versus intracytoplasmic sperm injection cycles: Influence of the fertilization procedure. *Fertil. Steril.* 83, 1397–1403. <https://doi.org/10.1016/j.fertnstert.2004.10.054>.
37. Reichman, D.E., Jackson, K.V., and Racowsky, C. (2010). Incidence and development of zygotes exhibiting abnormal pronuclear disposition after identification of two pronuclei at the fertilization check. *Fertil. Steril.* 94, 965–970. <https://doi.org/10.1016/j.fertnstert.2009.04.018>.
38. Ezoë, K., Takahashi, T., Shimazaki, K., Miki, T., Tanimura, Y., Amagai, A., Sawado, A., Akaike, H., Mogi, M., Kaneko, S., et al. (2022). Human 1PN and 3PN zygotes recapitulate all morphokinetic events of normal fertilization but reveal novel developmental errors. *Hum. Reprod.* 37, 2307–2319. <https://doi.org/10.1093/humrep/deac177>.
39. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
40. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. <https://doi.org/10.1093/bioinformatics/btt656>.
41. Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27, 1571–1572. <https://doi.org/10.1093/bioinformatics/btr167>.
42. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. <https://doi.org/10.1038/nmeth.1923>.
43. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
44. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.

45. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoerckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell* *177*, 1888–1902.e21. <https://doi.org/10.1016/j.cell.2019.05.031>.
46. Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* *32*, 381–386. <https://doi.org/10.1038/nbt.2859>.
47. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* *25*, 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
48. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* *38*, 576–589. <https://doi.org/10.1016/j.molcel.2010.05.004>.
49. Kharchenko, P.V., Silberstein, L., and Scadden, D.T. (2014). Bayesian approach to single-cell differential expression analysis. *Nat. Methods* *11*, 740–742. <https://doi.org/10.1038/nmeth.2967>.
50. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* *10*, 1523. <https://doi.org/10.1038/s41467-019-09234-6>.
51. Li, R., Qiao, J., Wang, L., Zhen, X., and Lu, Y. (2008). Serum progesterone concentration on day of HCG administration and IVF outcome. *Reprod. Biomed. Online* *16*, 627–631. [https://doi.org/10.1016/s1472-6483\(10\)60475-0](https://doi.org/10.1016/s1472-6483(10)60475-0).
52. Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* *9*, 171–181. <https://doi.org/10.1038/nprot.2014.006>.
53. Wang, Y., Yuan, P., Yan, Z., Yang, M., Huo, Y., Nie, Y., Zhu, X., Qiao, J., and Yan, L. (2021). Single-cell multiomics sequencing reveals the functional regulatory landscape of early embryos. *Nat. Commun.* *12*, 1247. <https://doi.org/10.1038/s41467-021-21409-8>.
54. Dahl, J.A., Jung, I., Aanes, H., Greggains, G.D., Manaf, A., Lerdrup, M., Li, G., Kuan, S., Li, B., Lee, A.Y., et al. (2016). Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* *537*, 548–552. <https://doi.org/10.1038/nature19360>.
55. Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* *33*, 495–502. <https://doi.org/10.1038/nbt.3192>.

**STAR★METHODS**

**KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Anti-ARGFX	Thermo Scientific	Cat# PA535190; RRID: AB_2552500
Anti-H3K27ac	Cell Signaling Technology	Cat# 8173S; RRID: AB_10949503
Goat anti-rabbit secondary antibody AlexaFluor 594	Invitrogen	Cat# A-11005; RRID:AB_2534073
Goat anti-rabbit secondary antibody AlexaFluor 488	Invitrogen	Cat# A-11008; RRID:AB_143165
<b>Biological samples</b>		
Human embryos	This study	N/A
<b>Chemicals, peptides, and recombinant proteins</b>		
G-1 plus	Vitrolife	Cat# 10128
G-2 plus	Vitrolife	Cat# 10132
HYASE-10X	Vitrolife	Cat# 10017
Exifone	TCI America	Cat# H17830
DMSO	Sigma-Aldrich	Cat# D2650
DAPI	Solarbio	Cat# C0065
Accutase	Invitrogen	Cat# A1110501
Trypsin-EDTA	Thermo Fisher Scientific	Cat# 25200056
SuperScript™ II Reverse Transcriptase	Invitrogen	Cat# 18080044
Ambion RNase Inhibitor	Thermo Fisher Scientific	Cat# AM2682
KAPA HiFi HotStart ReadyMix	KAPA Biosystems	Cat# KK2602
Agencourt AMPure XP beads	Beckman	Cat# A63881
M.CviPI GpC Methyltransferase	New England Biolabs	Cat# M0227S
S-adenosylmethionine	New England Biolabs	Cat# B9003S
EDTA	Sigma-Aldrich	Cat# T9285
phenylmethylsulfonyl fluoride	Roche	Cat# 10837091001
lambda DNA	Thermo Fisher Scientific	Cat# SD0011
protease	QIAGEN	Cat# 19155
carrier RNA	QIAGEN	Cat# 1017647
SYBR Master Mix	Applied Biosystems	Cat# A25742
<b>Critical commercial assays</b>		
NEBNext Ultra™ DNA Library Prep kit	New England Biolabs	Cat# E7370L
NEBNext Ultra™ II DNA Library Prep kit	New England Biolabs	Cat# E7645L
MethylCode™ Bisulfite Conversion Kit	Thermo Scientific	Cat# MECOV-50
QIAamp DNA Blood Mini kit	QIAGEN	Cat# 51104
Kapa Hyper Prep Kit	Kapa Biosystems	Cat# KK8504
DNA Clean & Concentrator-5	Vistech	Cat# DC2005
<b>Deposited data</b>		
RNA-seq data (human embryo)	This paper	GSA: HRA006449
NOMe-seq data (human embryo)	This paper	GSA: HRA006449
RNA-seq data (single blastomeres biopsied from human 8-cell embryos)	He et al. <sup>9</sup>	GSA: HRA001473
H3K27ac ChIP-seq data (human embryo)	Wu et al. <sup>32</sup>	GSA: HRA002355
H3K4me3 ChIP-seq data (human embryo)	Xia et al. <sup>33</sup>	GEO: GSE124718
H3K9me3 ChIP-seq data (human embryo)	Yu et al. <sup>23</sup>	GEO: GSE176016
<b>Oligonucleotides</b>		
siRNA targeting sequences (Table S1)	This paper	N/A

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Primers for quantitative real-time PCR (Table S1)	This paper	N/A
<b>Software and algorithms</b>		
Trim Galore v0.4.1	<a href="http://www.bioinformatics.babraham.ac.uk/projects">http://www.bioinformatics.babraham.ac.uk/projects</a>	<a href="http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/">http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/</a>
STAR v2.7.0e	Dobin et al. <sup>39</sup>	<a href="https://github.com/alexdobin/STAR">https://github.com/alexdobin/STAR</a>
featureCounts v1.6.4	Liao et al. <sup>40</sup>	<a href="http://subread.sourceforge.net/">http://subread.sourceforge.net/</a>
Bismark v0.22.3	Krueger and Andrews <sup>41</sup>	<a href="https://www.bioinformatics.babraham.ac.uk/projects/bismark/">https://www.bioinformatics.babraham.ac.uk/projects/bismark/</a>
PICARD v2.19.0	N/A	<a href="https://broadinstitute.github.io/picard/">https://broadinstitute.github.io/picard/</a>
Bowtie2 v2.4.5	Langmead and Salzberg <sup>42</sup>	<a href="http://bowtie-bio.sourceforge.net/bowtie2/index.shtml">http://bowtie-bio.sourceforge.net/bowtie2/index.shtml</a>
Samtools v0.1.18	Li et al. <sup>43</sup>	<a href="http://samtools.sourceforge.net/">http://samtools.sourceforge.net/</a>
MACS2 v2.2.6	Zhang et al. <sup>44</sup>	<a href="https://pypi.org/project/MACS2/">https://pypi.org/project/MACS2/</a>
R v4.1.0	<a href="https://www.R-project.org/">https://www.R-project.org/</a>	<a href="https://www.R-project.org/">https://www.R-project.org/</a>
Seurat v4.4.0	Stuart et al. <sup>45</sup>	<a href="https://satijalab.org/seurat/">https://satijalab.org/seurat/</a>
Monocle2	Trapnell et al. <sup>46</sup>	<a href="https://cole-trapnell-lab.github.io/monocle-release/">https://cole-trapnell-lab.github.io/monocle-release/</a>
BWA	Li and Durbin <sup>47</sup>	<a href="https://github.com/lh3/bwa">https://github.com/lh3/bwa</a>
HOMER v4.11.1	Heinz et al. <sup>48</sup>	<a href="http://homer.ucsd.edu/homer/">http://homer.ucsd.edu/homer/</a>
SCDE	Kharchenko et al. <sup>49</sup>	<a href="https://hms-dbmi.github.io/scde/">https://hms-dbmi.github.io/scde/</a>
Metascape	Zhou et al. <sup>50</sup>	<a href="http://metascape.org/gp/index.html#/main/step1">http://metascape.org/gp/index.html#/main/step1</a>
ImageJ	ImageJ	<a href="https://imagej.nih.gov/ij/">https://imagej.nih.gov/ij/</a>

**EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**

**Ethics statement**

This study was conducted under the guidance of the Ethics of Human Biomedical Research (issued by the National Health Commission of the People's Republic of China, 2016) and was approved by the Reproductive Medicine Ethics Committee of Peking University Third Hospital (Research License M2021628), which was reviewed through human Embryo Research Oversight (EMRO) process. The committee evaluated this study's scientific merit and ethical justification and conducted a comprehensive review of donations, subsequent uses of samples, and potential risks to donors. The ethical and regulatory framework established by the Reproductive Medicine Ethics Committee of Peking University Third Hospital clearly states that informed consent can only be obtained if eligible participants are provided with all necessary information about the study and opportunities for appropriate counseling. The informed consent form clearly described this study's objectives and related clinical procedures.

The embryos used for allelic transcriptome analysis were generated through intracytoplasmic sperm injection (ICSI) from donated oocytes and semen not used for fertility treatment, which were voluntarily donated by donors at the Center for Reproductive Medicine, Peking University Third Hospital. The oocytes and semen donors were provided with the relevant information regarding this research, informed that their gametes would be used to form embryos only for research purposes. They were fully aware of the potential risks and benefits that may arise from this study. No financial inducements were offered for donation. 3PN embryos were clinically discarded and donated by patients voluntarily for scientific research purposes. All human embryos involved in this study were not used for fertility treatment, and the culture of all embryos was terminated before 14 days post-fertilization. All procedures in this study were in accordance with the International Society for Stem Cell Research (ISSCR) guidelines.

**METHOD DETAILS**

**Human embryo culture**

Human oocytes were obtained from five fertile women, the superovulation procedure was performed according to a published method.<sup>51</sup> Hyaluronidase (Vitrolife, 10017) was used to remove the surrounding granulosa cells of metaphase II (MII) oocytes. Human semen was obtained from two healthy men with normal semen parameters. Swim-up sperms were used, and the intracytoplasmic sperm injection (ICSI) was performed to obtain embryos. Normal fertilization was evaluated 16-18 hours after ICSI based on the presence of 2PN (pronuclei), and only 2PN zygotes were used for subsequent culture. The embryos were first cultured in G-1 plus medium (Vitrolife, 10128) at 37°C, 6% CO<sub>2</sub>, and then transferred into G-2 plus medium (Vitrolife, 10132) on Day 3 for further culture. Embryos with normal morphology and developmental speed were collected at different stages, including zygote, 2-cell, 4-cell, 8-cell, morula, and blastocyst stages. The zygotes were collected 16-24 hours after ICSI, the 2-cell, 4-cell and 8-cell embryos were collected at 26

hours, 44 hours, and 66–68 hours after fertilization respectively. The morulae and blastocysts were harvested on Day 4 and Day 5–Day 6. All 2PN embryos used in this study were described in [Table S2](#).

For 3PN embryos that were clinically discarded and donated by patients after fertilization evaluation, the oocytes were inseminated by conventional in-vitro fertilization (IVF) or ICSI, and the fertilization status was checked 16–18 hours later.

In the exifone treatment experiment, exifone (TCI America, H17830) was first dissolved in DMSO (Sigma-Aldrich, D2650) at a concentration of 1M and stored in  $-20^{\circ}\text{C}$ , G-1 plus or G-2 plus medium with  $0.02\ \mu\text{M}$  exifone were used for embryo culture to activate HDAC1.

### Target genes knockdown

To study the function of *DPRX* and *ARGFX* in human embryonic development, 3PN zygotes were microinjected with siRNA (Ribobio) to knockdown target genes, and non-silencing siRNA was used as negative control. Three siRNAs were designed for each target genes, and diluted to  $50\ \mu\text{M}$  in nuclease-free water. The three siRNAs were mixed equally, and about 5–10 pL mixture was injected into embryos by electronic microinjectors FemtoJet 4i (Eppendorf). Injected zygotes were cultured in G-1 plus medium and transferred into G-2 plus medium on Day 3 in an incubator at  $37^{\circ}\text{C}$  and 6%  $\text{CO}_2$ , as described before. The phenotype of injected embryos was recorded with microscopy. The sequences of siRNAs against human *DPRX* or *ARGFX* are listed in [Table S1](#).

### Immunofluorescence

Embryos with normal morphology were fixed in 4% paraformaldehyde (PFA) at room temperature for 30 min, and then transferred into washing buffer (0.1% Tween and 0.01% Triton X-100 in PBST) at  $4^{\circ}\text{C}$  for preservation. After permeabilized with 0.3% Triton X-100 for 40 min, the samples were transferred into 1% bovine serum albumin (BSA) for 1 h for blocking, and then incubated with the primary antibody (*ARGFX* antibody, Thermo Scientific, PA535190, 1:10; H3K27ac antibody, Cell Signaling Technology, 8173S, 1:200) overnight at  $4^{\circ}\text{C}$ . The samples were washed five times in washing buffer and then transferred into DAPI solution (10  $\mu\text{g}/\text{ml}$ , Solarbio, C0065) with secondary antibody (Alexa Fluor 594 goat anti-rabbit, Invitrogen, A11005, 1:200; Alexa Fluor 488 goat anti-rabbit, Invitrogen, A11008, 1:200) for 2 h at room temperature. After another five washes, embryos were transferred to glass bottom dishes (NEST, 801001) or mounted on slides of glass. The images were captured by a laser scanning confocal microscope platform (Carl Zeiss LSM 880, ZEISS, Germany). Multichannel images captured by confocal z-stacks were processed in ImageJ. The quantification of H3K27ac nucleus fluorescence intensity was conducted by normalizing the ratio between H3K27ac signal intensity and DNA signal intensity using the built-in functions of ImageJ.

### Real-time quantitative PCR

RNA from an entire embryo was extracted and reverse-transcribed into cDNA by SuperScript<sup>TM</sup> II Reverse Transcriptase, the cDNA was pre-amplified and purified before performing qPCR. SYBR Master Mix (Applied Biosystems, A25742) and QuantStudio 3 Real-Time PCR System were used for qPCR analyses, and relative gene expression was analyzed based on the  $\Delta\Delta\text{CT}$  method. Each group contained 8–9 samples. The qPCR primers used in this study are listed in [Table S1](#).

### Isolation of single blastomeres

Embryos were moved into the acidic solution containing 0.3% HCl by mouth pipette for several seconds to remove zona pellucida. After being washed in DPBS containing 0.5% HSA (Vitrolife, 10064) several times, the embryos were moved into a mixture of Accutase (Invitrogen, A1110501) and 0.25% trypsin-EDTA (Thermo Scientific 25200056) at a ratio of 1:1 for 5–60 minutes in an incubator to digest. Single blastomeres were separated by repeat gentle pipetting, and transferred into Smart-seq lysis after three washes in DPBS containing 0.5% HAS. The zona-free zygotes were transferred into lysis directly without digestion.

### Single-cell full-length RNA sequencing

The RNA-seq library construction strategy was improved based on Smart-seq2 protocol.<sup>52</sup> In brief, reverse transcription reaction was performed with SuperScript<sup>TM</sup> II Reverse Transcriptase (Invitrogen, 18080044), KAPA HiFi HotStart ReadyMix (KAPA Biosystems, KK2602) was used for pre-amplification. After being purified with Ampure XP beads (Beckman, A63881), the amplification products were sheared into fragments by the Covaris S2 system ( $\sim 300$ -bp long). DNA Clean & Concentrator-5 (Vistech, DC2005) and Ampure XP beads were used to purify the fragmented cDNAs. Approximately 25 ng of the cDNA fragments were used to construct library, and it was performed by NEBNext Ultra<sup>TM</sup> or NEBNext Ultra<sup>TM</sup> II DNA Library Prep kit (New England Biolabs, E7370L or E7645L). Sequencing was performed on an Illumina X-ten platform or Nova platform.

### Single-embryo NOMe-seq sequencing

Libraries were generated following the NOMe-seq protocol as previously described.<sup>53</sup> Briefly, single embryos were moved into lysis containing  $1 \times$  M.CviPI reaction buffer (New England Biolabs, M0227S), 5U M.CviPI GpC Methyltransferase (New England Biolabs, M0227S),  $160\ \mu\text{M}$  S-adenosylmethionine (New England Biolabs, B9003S), 0.25mM EDTA (Sigma-Aldrich, T9285), 0.25mM phenylmethylsulfonyl fluoride (Roche, 10837091001), and 1 pg of lambda DNA (Thermo Scientific, SD0011) after removal of zona pellucida. The mixture was incubated at  $37^{\circ}\text{C}$  for 60min to methylate GpC, followed by heat inactivation at  $65^{\circ}\text{C}$  for 25min. Then, 0.5 $\mu\text{l}$  of protease (QIAGEN, 19155) and 10 ng of carrier RNA (QIAGEN, 1017647) were introduced and incubated at  $50^{\circ}\text{C}$  for 3h to release genomic DNA. Heat inactivation was then performed at  $75^{\circ}\text{C}$  for 30min. The genomic DNA underwent bisulfite conversion (Thermo

Scientific, MECOV-50), and the DNA libraries were generated according to scBS-seq protocol. Sequencing was performed on an Illumina X-ten platform.

### Whole genome sequencing

Peripheral blood was donated by male and female donors and collected in EDTA anticoagulant tubes. The whole genome DNA was extracted from 500  $\mu$ L anticoagulant blood with the QIAamp DNA Blood Mini kit (QIAGEN, 51104). DNA was fragmented as approximately 300 bp by Covaris S2 system before library construction with a Kapa Hyper Prep Kit (Kapa Biosystems, KK8504). Sequencing was performed on an Illumina X-ten platform.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Single-cell RNA-seq data processing

Raw reads were first preprocessed using Trim Galore (v0.4.1) to trim low-quality bases and primer sequences (with the parameters: trim\_galore –paired –quality 20 –phred33 –stringency 3 –gzip –length 36). The trimmed reads were then mapped to the human reference genome of GRCh38 using STAR software (v2.7.0e)<sup>39</sup> with default settings, and only uniquely mapped reads were kept. featureCounts (v1.6.4)<sup>40</sup> was used to quantify the reads mapped to each annotated gene (parameters: featureCounts -p -t exon -g gene\_name). Samples were excluded from future analysis based on the following cutoffs: (1) less than 4,000 genes detected, (2) the library size was less than 0.1 million total counts, (3) reads for mitochondrial genes accounted for were over 30%, (4) outlier cells identified by PCA dimensionality reduction. Genes expressed in less than 3 cells were filtered out.

### Single-embryo NOME-seq data processing

Raw reads were first preprocessed using Trim Galore (v0.4.1) to remove adaptors and low-quality bases (with the parameters: –clip\_R1 9 –clip\_R2 9 –quality 20 –phred 33 –stringency 3 –length 50 –paired). Next, the trimmed reads were mapped to the lambda DNA and the human reference genome of GRCh38 using Bismark (v0.22.3)<sup>41</sup> (with the parameters: bismark –bowtie2 –non\_directional). The mapped reads were merged and PCR duplicates were removed using PICARD (v2.19.0).

### DNA methylation and accessibility analysis

For each cytosine site, the methylation level was computed as the ratio of the sum of methylated reads to the total number of covered reads (sum of both methylated and unmethylated reads). The DNA methylation level was referred to as the methylation level of WCG (ACG, TCG) sites, and the chromosome accessibility level was referred to as the methylation level of GCH (GCA, GCC, and GCT) sites.

### ChIP-seq data processing

The ChIP-seq data for H3K27ac modifications of early human embryogenesis were downloaded from HRA002355 in PRJCA009410<sup>26</sup> from the Genome Sequence Archive (GSA) database. Low-quality reads were first removed using Trim Galore (with the parameters: trim\_galore –paired –quality 20 –phred33 –stringency 3 –gzip –length 36). Paired reads were then mapped to the human reference genome of GRCh38 using Bowtie2 (v2.4.5) (with the parameters: -X 800 –no-mixed –no-discordant). Reads with Mapping Qualities (MAPQ) < 10 were removed by Samtools (v0.1.18), and PCR duplicates were removed by PICARD (v 2.19.0).

### Peak analysis

H3K27ac peaks were called by MACS2 (v2.2.6)<sup>44</sup> (with the parameters: -p 1e-3 –broad –broad-cutoff 0.01 –max-gap 500). The peak signal of a domain was calculated as the sum of RPKM values of consecutive 100bp bins within that domain. Signal normalization across typical peaks and broad domains was performed according to the published method.<sup>26,54</sup>

### SNP detection of RNA and DNA sequencing data

We followed a previously published protocol<sup>14</sup> (code available at [https://github.com/WRui/Post\\_Implantation](https://github.com/WRui/Post_Implantation)) to generate VCF files from single-cell RNA sequencing and parental genome DNA sequencing data. After performing adapter removal and quality filtering on the sequencing reads using the Trim Galore software, we employed BWA software to map the processed reads to the hg38 reference genome. Next, the “MarkDuplicates” function of PICARD software were utilized to remove the redundant alignments. Subsequently, the GATK pipeline were employed to identify SNPs from high-quality read alignments. The acquired VCF files were further used for allelic expression analysis.

### Allele-specific analysis of RNA expression

The gametes from two male and five female donors with different genetic backgrounds were used to generate embryos. Seven different parental combinations were utilized, resulting in a total of 383,125 average traceable hybrid SNP, which were used to separate paternal and maternal traceable reads. The detailed protocol can be found in the published pipeline.<sup>53</sup> In general, each mapped read covering an informative SNP site was analyzed according to the specific base at that SNP position. The reads were then assigned as either paternal or maternal based on whether the base matched the paternal or maternal allele, respectively. The generated paternal and maternal separated bam files were counted via featureCounts to quantify the allelic gene expression profile.

Maternal allelic ratio was calculated as the ratio of maternal transcript counts to the sum of maternal and paternal transcript counts. When calculating the allelic ratio of samples, only samples with a sum of maternal and paternal transcript counts greater than 10,000 were retained.

### Principal component analysis

Principal component analysis (PCA) was performed with the top 2000 high variable genes, which were identified using the “FindVariableFeatures” function in the Seurat (v4.4.0) package<sup>55</sup>

### Inference of embryonic sex

Two methods were used in this study to determine the sex of each embryo. (1) The expression of Y-linked genes. Considering female embryos do not express Y-linked genes, we calculated the sum expression (TPM) of Y chromosome-specific genes of each cell in each embryo, and embryos that did not express Y-linked genes (sum TPM = 0) were defined as female. (2) The difference in allelic ratios between autosomes (Auto) and X chromosome (chrX). Considering male embryos inherit chrX from the maternal genome, whereas female embryos inherit chrX from both parental genomes, the sex of embryos could be inferred by comparing the paternal allelic ratios (PAR) of chrX with that of autosomes. Male embryos: (chrX PAR/Auto PAR) < 0.1; Female embryos: (chrX PAR/Auto PAR) >= 0.1. Because the human embryos initiate minor ZGA at the 4-cell stage, we cannot identify the sex of zygotes and 2-cell embryos with transcriptome information.

### Pseudotime analysis

Monocle2<sup>19</sup> was used to reconstruct single-cell trajectories. Embryonic stage specific genes were first identified by Seurat “FindAllMarkers” (logfc.threshold = log(1.5), min.pct = 0.25, only.pos = T and test.use = “roc”) and then used as input for Monocle2. Dimensionalities were reduced by performing “reduceDimension” function with the method of “DDRTree” and the parameter “max\_components = 2”. After reducing the expression data into a lower-dimensional space, the cells were ordered along the trajectory by the “orderCells” function. Cells were then colored by embryonic stage, confirming that pseudotime analysis can well match the physiological development of embryos, thus reflecting developmental arrest in some embryos.

### Expression estimation of transposable elements

The RepeatMasker output files for the human genome (hg38) were downloaded from the UCSC Genome Browser. In brief, the clean data was mapped to hg38 reference genome using the BWA tools (with the parameters: aln -o 1 -e 60 -i 15 -q 10 -t 8). The featureCounts was used to calculate aligned read counts for transposable elements and annotated genes. The read counts mapped to transposons was further normalized with the read counts of annotated genes to estimate the expression level of transposable elements.

### Identification of nucleosome-depleted region

First, GCH data from individual embryos within the same group were aggregated. Next, the genome was segmented into 100bp windows with 20bp spacing, and the number of C and T reads within each window was quantified. The significance of methylation differences from the genomic baseline was assessed using the chi-square test. Finally, the nucleosome-depleted regions (NDRs) were identified as regions exhibiting significantly increased GCH methylation, indicated by  $P \leq 1 \times 10^{-15}$ , spanning a minimum length of 140bp, and encompassing at least 3 GCH sites.

### Transcription factor binding motif analysis

To find the enriched TF motifs at putative enhancers for different gene sets, we identified the nearest distal open chromatin regions as the closest NDR peak more than 2kb away from the annotated promoter. The identification of known motifs was then performed using the function “findMotifsGenome.pl” in HOMER (v4.11.1)<sup>48</sup> with the “-size given” parameter. To determine the counts of PRD-like TF motifs in the distal open chromatin regions of each gene, “scanMotifGenomeWide.pl” in HOMER was used to map the genome-wide motif locations based on the “TAATCC” sequence.

### Differentially expressed genes analysis

After filtering out cells and genes, differentially expressed genes of normal and TAS 8-cells were determined with SCDE,<sup>49</sup> a Bayesian-based approach that enabled the detection of differential expression signatures in a way that was more tolerant to noise. DEGs were defined as genes  $|\log_2(\text{fold change})| \geq 1$  and  $P < 0.01$ .

In target gene knockdown experiments, DEGs in 3PN embryos were defined as genes with at least a 2-fold change between siNC and knockdown groups.

### ZGA genes and maternal factors identification

Minor ZGA genes were defined as the genes that were significantly upregulated from the 2-cell stage to the 4-cell stage, which was also calculated by SCDE (4-cell vs 2-cell,  $\log_2(\text{fold change}) \geq 2$ ,  $P < 0.01$ ). Major ZGA genes were defined as the genes that were significantly upregulated from 4-cell stage to normal 8-cell stage (normal 8-cell vs 4-cell,  $\log_2(\text{fold change}) \geq 2$ ,  $P < 0.01$ ). Maternal factors were defined as genes highly expressed in zygote stage (average TPM  $\geq 1$ ) without up-regulated expression ( $P < 0.05$  and  $\log_2(\text{fold change}) > 0$ ) at the later stages.

### Maternal decay degree identification

The ratio of the reduction in expression of maternal factors between zygote and 8-cell stage to their expression in zygotes was defined as the maternal decay degree.

### Correlation analysis

To profile the relationships between gene expression level and maternal allelic ratio, ZGA degree, the Pearson correlation was calculated between gene expression level ( $\log_2(\text{TPM}+1)$ ) and the maternal allelic ratio of blastomeres at the 8-cell stage. The mean expression level of ZGA genes at the 8-cell stage ( $\log_2(\text{Mean ZGA gene TPM}+1)$ ) was used as an indicator of ZGA degree. Pearson correlation was calculated between gene expression level and ZGA degree.

### Prolonged activated genes identification

The prolonged activated genes in dKD embryos or TAS 8-cell embryos were defined as up-regulated genes in dKD embryos or TAS 8-cell embryos, while maternal factors and genes with expression levels (TPM) > 10 in zygotes were excluded.

### Gene ontology analysis

GO terms enrichment analysis was performed using Metascape.<sup>50</sup> (<http://metascape.org/gp/index.html#/main/step1>)

### Statistical analysis

Wilcoxon test, unpaired two-sample t-test, Chi-squared test, dip test, hypergeometric test or Pearson correlation were used as appropriate. Statistical assays, sample sizes, definition of center and precision measures are reported in figures and figure legends. All statistical analyses were performed in R.