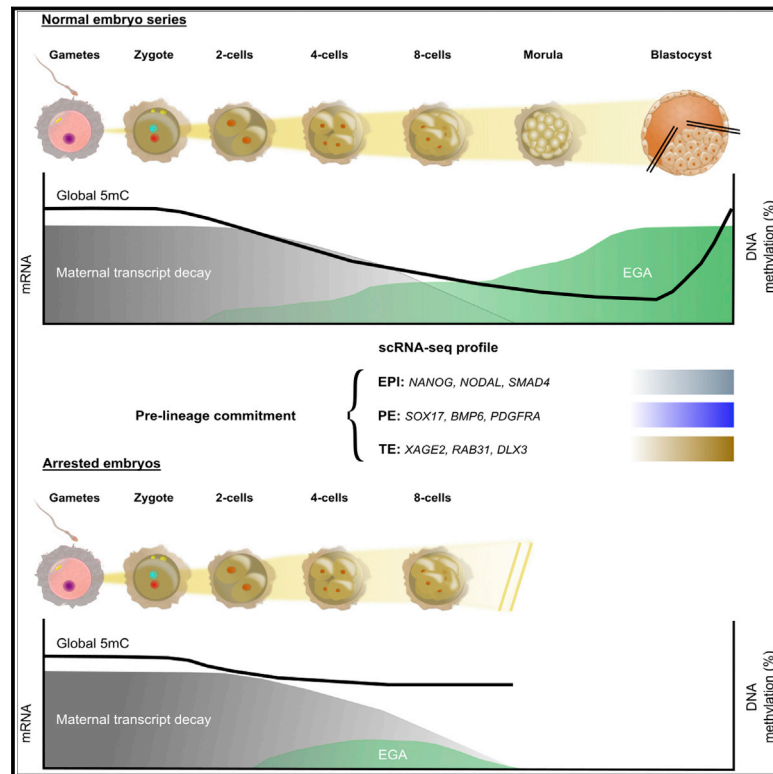


# Cell Reports

## Single-cell multi-omic analysis profiles defective genome activation and epigenetic reprogramming associated with human pre-implantation embryo arrest

### Graphical abstract



### Authors

Jose Ramon Hernandez Mora, Claudia Buhigas, Stephen Clark, ..., Marcos Meseguer, Gavin Kelsey, David Monk

### Correspondence

d.monk@uea.ac.uk

### In brief

Hernandez Mora et al. use single-cell M&T-seq to profile genome activation and DNA methylation dynamics in a developmental series of high-quality pre-implantation embryos. The resulting profiles are compared with developmentally arrested embryos with good morphology. Their findings clarify that aberrant molecular mechanisms involved in maternal-to-zygotic transition may result in embryo demise.

### Highlights

- Combined single-cell transcriptional and DNA methylation profiling of human embryos
- Distinct genes/molecular pathways identified for early developmental stages
- Trophectoderm differentiation occurs largely independent of DNA methylation
- Aberrant EGA and DNA methylation dynamics observed in arrested embryos



## Resource

# Single-cell multi-omic analysis profiles defective genome activation and epigenetic reprogramming associated with human pre-implantation embryo arrest

Jose Ramon Hernandez Mora,<sup>1,10</sup> Claudia Buhigas,<sup>2,10</sup> Stephen Clark,<sup>3</sup> Raquel Del Gallego Bonilla,<sup>4</sup> Dagne Daskeviciute,<sup>2</sup> Ana Monteagudo-Sánchez,<sup>1</sup> Maria Eugenia Poo-Llanillo,<sup>5</sup> Jose Vicente Medrano,<sup>4,6</sup> Carlos Simón,<sup>6,7</sup> Marcos Meseguer,<sup>4</sup> Gavin Kelsey,<sup>3,8,9</sup> and David Monk<sup>1,2,11,\*</sup>

<sup>1</sup>Bellvitge Institute for Biomedical Research, 08908 Barcelona, Spain

<sup>2</sup>Biomedical Research Centre, School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7JT, UK

<sup>3</sup>The Babraham Institute, Babraham, Cambridge CB22 3AT, UK

<sup>4</sup>IVI-RMA Global and Health Research Institute la Fe, 46026 Valencia, Spain

<sup>5</sup>Igenomix SL, 46980 Valencia, Spain

<sup>6</sup>Department of Obstetrics and Gynecology, Valencia University and INCLIVA, 46010 Valencia, Spain

<sup>7</sup>Department of Obstetrics and Gynecology, BIDMC, Harvard University, Boston, MA 02215, USA

<sup>8</sup>Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EL, UK

<sup>9</sup>Wellcome-MRC Institute of Metabolic Science-Metabolic Research Laboratories, Cambridge CB2 0QQ, UK

<sup>10</sup>These authors contributed equally

<sup>11</sup>Lead contact

\*Correspondence: [d.monk@uea.ac.uk](mailto:d.monk@uea.ac.uk)

<https://doi.org/10.1016/j.celrep.2023.112100>

## SUMMARY

During pre-implantation stages of mammalian development, maternally stored material promotes both the erasure of the sperm and oocyte epigenetic profiles and is responsible for concomitant genome activation. Here, we have utilized single-cell methylome and transcriptome sequencing (scM&T-seq) to quantify both mRNA expression and DNA methylation in oocytes and a developmental series of human embryos at single-cell resolution. We fully characterize embryonic genome activation and maternal transcript degradation and map key epigenetic reprogramming events in developmentally high-quality embryos. By comparing these signatures with early embryos that have undergone spontaneous cleavage-stage arrest, as determined by time-lapse imaging, we identify embryos that fail to appropriately activate their genomes or undergo epigenetic reprogramming. Our results indicate that a failure to successfully accomplish these essential milestones impedes the developmental potential of pre-implantation embryos and is likely to have important implications, similar to aneuploidy, for the success of assisted reproductive cycles.

## INTRODUCTION

In mammals, development starts with fertilization, the process of which triggers events essential for the development of the totipotent embryo, collectively referred to as the maternal-to-zygotic transition (MZT).<sup>1</sup> Initial events include paternal protamine-histone exchange and the erasure of the gametic DNA methylation and post-translational histone tail modifications, in a process referred to as global epigenetic reprogramming.<sup>2,3</sup> Upon reaching cleavage day 2–3 (CL2–CL3, respectively) containing 2–8 cells, respectively, transcription products from the embryonic genome have replaced most oocyte-derived material in a process known as embryonic genome activation (EGA) that takes control of further development during MZT.<sup>4</sup> By day 5–6 of development, dynamic events correlate with a progressive loss in totipotency, until the approximate time of blastocyst formation. At this time the

embryo has undergone committed differentiation resulting in the inner cell mass (ICM) and trophectoderm (TE), presumably as the result of differential expression of a few lineage-specific transcripts.<sup>5,6</sup>

To better understand the relationship between these early processes and early human development, we have utilized single-cell parallel methylation and transcriptome sequencing (scM&T-seq), a technique that simultaneously allows for quantitative expression and underlying DNA methylation profiles to be obtained from the same cell.<sup>7</sup> Here, we characterize the main transcriptomic and epigenetic events in metaphase II (MII) oocytes and high-quality CL2–CL3 embryos and day 5 blastocysts. These datasets act as references that subsequently allow us to describe the extent of heterogeneity during this critical developmental window and compare the profiles with embryos that have undergone spontaneous developmental arrest.



## RESULTS

### Methylation and transcriptome profiling in good quality embryos

To obtain integrated epigenetic and transcriptional maps of early human development, we performed scM&T-seq on 11 MII oocytes and human cleavage embryos (day 2, 2–4 cells,  $n = 8$ ; day 3, 5–8 cells,  $n = 7$ ) (referred to as CL2 and CL3, respectively) and blastocysts (day 5–6,  $n = 4$ ) suitable for transfer (i.e., cleavage embryos scoring G1 or G2; expanded or hatching blastocysts scoring ICM and TE A or B), with the latter subject to laser dissection to physically separate the mural TE (mTE) from the polar TE (pTE)/ICM (Table S1). A total of 128 cells were manually isolated and processed from 19 embryos across all stages (Figure 1A). After quality control, we retained the scRNA-seq datasets for 10 oocytes and 117 embryonic cells with an average of 14,295 (SD 2,693) expressed transcripts with at least one read in three different cells, which represented 19,582 (SD 1,066) transcripts per oocyte, 16,521 (SD 2,342) per blastomere of a CL2 embryo, 14,782 (SD 1,297) per cell of CL3 embryos, and 12,889 (SD 2,250) per cell from blastocysts. These expression levels are of comparable quality to those generated using stand-alone scRNA-seq in human embryos.<sup>8</sup>

The scM&T-seq protocol allows for parallel methylome and transcriptome sequencing within single cells since RNA is physically separated from DNA. We subjected the genomic DNA to bisulfite conversion following the post-bisulfite adaptor tagging (PBAT) strategy, which enables us to detail the complex relationship between DNA methylation and transcription in individual heterogeneous cells of early human embryos in unprecedented detail. Sequencing of the single-cell bisulfite sequence (scBS-seq) libraries was performed (after trimming average of 7.9 M reads/cell, SD = 4 M) with an average of 3 M (SD = 2.1 M) uniquely mapped reads per cell. The mean mapping efficiency for oocytes was 51.6% (SD = 2.5%), while CL2 and CL3 embryos had 39.7% (SD = 8.9%) and 33.5% (SD = 11%), respectively. Cells from blastocysts had a mapping efficiency of 35.1% (SD = 10.7%). Combined, we have 10 oocytes and 106 embryonic cells that passed quality control for both transcriptomes and methylomes, allowing direct inference between datasets.

To ascertain the sex of each embryo, we utilized both the expression of Y-linked genes and chromosome XY mapped bisulfite sequence reads. Furthermore, we assigned the cell-cycle stage for each individual cell, as this is known to be linked with fundamental biological processes relevant to early development and a source of potential transcriptional heterogeneity. Reassuringly, this revealed that oocytes were mostly in G2/M, while cells from cleavage-stage embryos and blastocysts were asynchronous, being assigned to different stages of the cell cycle (Table S1).

To determine whether gene expression was correlated with developmental stage, transcriptomes from the isolated cells were subject to t-distribution stochastic neighbor embedding (t-SNE) analysis that showed that the dataset separated by developmental order but failed to efficiently separate the ICM and TE lineages (Figure 1B). In fact, increasing the number of potential clusters or clustering only blastocyst cells resulted in separation by embryo rather than cell type, suggesting origin

was more influential than ICM and TE classification. The accuracy of the initial cell-type classification was verified when we analyzed the expression of known ICM and TE marker genes,<sup>4</sup> along with expression of *CCR7* allowing for the identification of pTE cells<sup>5</sup> (Table S1).

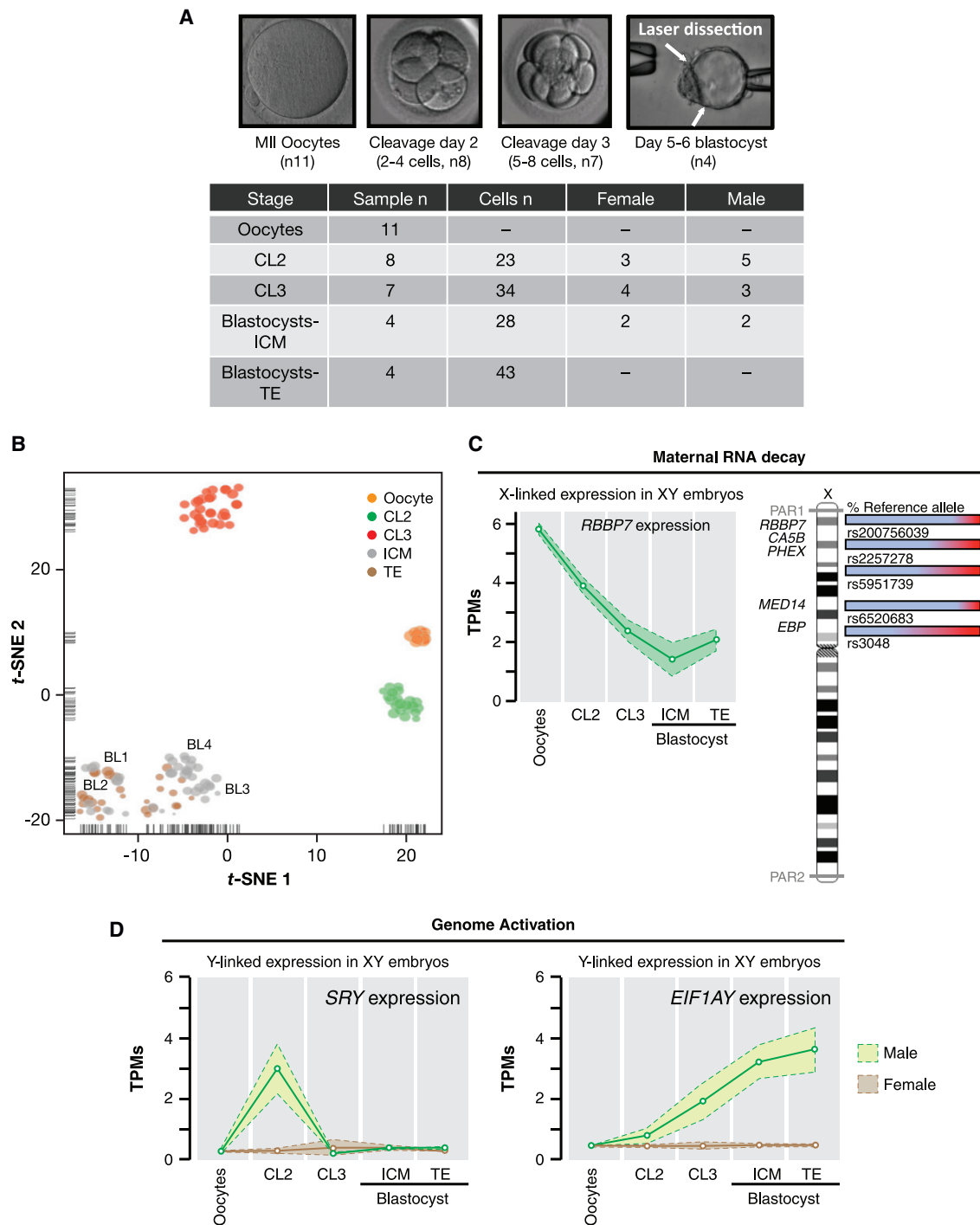
### Assessing the maternal-to-zygotic transition and genome activation

One often overlooked process during pre-implantation development is the exclusion of maternal inherited RNA from the oocyte. To confirm the timing of this event, we performed allelic expression analysis using single nucleotide polymorphisms (SNPs) to discriminate oocyte-derived X chromosome alleles in male embryos. We observed biallelic expression of X-linked genes in CL2 embryos that represent detection of maternally derived mRNAs inherited from the oocyte, which gradually disappeared by the blastocyst stage, indicative of maternal RNA clearance (Figure 1C). Consistent with recent studies in mice, *BTG4* and *CNOT6L*, as well as *YTHDF2*, a m<sup>6</sup>A RNA reader required for post-transcriptional transcript dosage during MZT, are all expressed in oocytes and CL2–CL3 embryos, therefore potentially licensing maternally derived mRNA clearance programs in humans also<sup>9–11</sup> (Table S2). Concomitant with these events is EGA. To study the transcriptional activation of the embryonic genome, we first quantified the expression of Y-linked genes in male embryos. We observed a peak of *SRY* expression in CL2 embryos, confirming that EGA occurs at the 2–4 cells stage, as well as gradual activation of *EIF1AY* from CL2 to blastocyst stage (Figure 1D).

### Initial EGA and DUX-family transcription factor expression

To determine whether global expression profiles were associated with specific developmental stages, we performed differential gene expression analysis combining two different approaches, the Wilcoxon test and MAST analysis.<sup>12</sup> To identify mutually exclusive upregulated markers, we focused on only those intersecting transcripts identified by both approaches with a differentially doubling of expression following Bonferroni correction for multiple comparisons ( $p$  value < 0.05) (Figure 2A and Table S2). A total of 1,265 genes were differentially expressed in CL2 embryos (including the *ELOA2/3* cluster and *PRR20* family of genes), 1,159 in CL3 (including the *PRAMEF* and *EIF1A* families), and 2,261 in blastocysts. Consistent with the poor performance separating ICM and mTE cells, we only observed 93 upregulated genes in ICM. In contrast, we found 1,185 specifically upregulated genes in mTE cells, suggesting that mTE cells represent a pure and differentiated population. For genes subject to this initial EGA, gene ontology (GO) analysis revealed enrichment for genes involved in mRNA processing and RNA processing, nucleosome assembly, and organization, as well as mitochondrial ATP synthesis, indicating that CL2 embryos have independently initiated these essential biological processes. To further confirm the developmental timing of each individual cell, we utilized the stage-specific markers identified to assign pseudo-timing, which corroborated the linear progression of development (Figure 2B).

To classify the developmental transitions in greater detail, we performed pairwise comparisons of successive time points. For



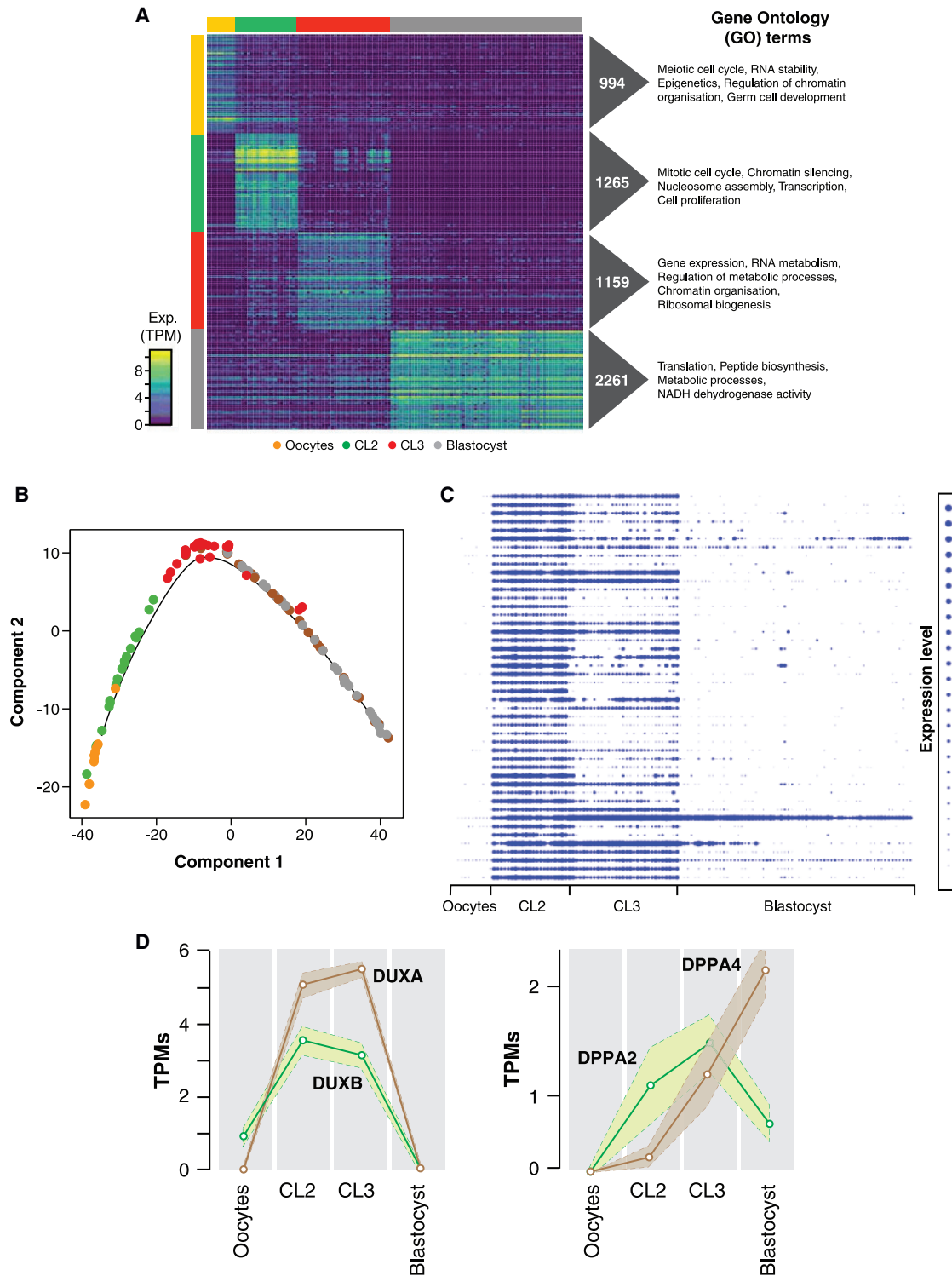
**Figure 1. scM&T-seq of human reference embryos**

(A) Table listing the number of high-quality reference embryos and cells.

(B) t-SNE analysis of expression data for cells that passed quality control and filtering criteria.

(C) Expression of X-linked RB binding protein 7 of *RBBP7* showing diminishing abundance across our developmental series consistent with maternal-derived mRNA clearance. Allele-specific expression of X-linked SNPs showing the percent reads assigned to the reference and alternative alleles in cell isolated from CL2 male embryos.

(D) Genome activation as determined by expression of Y-link genes in male and female embryos at different developmental time points.



**Figure 2. Characterization of gene activation at different stages of pre-implantation development**

(A) Heatmap showing the transcriptional profile for the top 50 stage-specifically activated genes. The total number of activated genes for each developmental stage is shown, as are the most significant gene ontology terms.

(legend continued on next page)

MZT that happens between the oocyte and CL2 embryos, we detected 281 differentially expressed transcripts. 20 of the top 25 most significant genes were activated in CL2 embryos from quiescent state in oocytes, although only 35% ( $n = 98$ ) showed this profile, with the majority representing diminished abundance of oocyte-derived transcripts (Table S2).

Recently *DUX4*, which is encoded within *D4Z4* telomeric microsatellite repeats on chromosomes 4q, has been shown to be a pioneer factor expressed in two-cell embryos responsible, in part, for minor EGA.<sup>13,14</sup> Unfortunately, we could not determine the transcription profile of *DUX4* in our datasets due to the fact its RNA is not polyadenylated, and mRNA capture is central to the scM&T-seq technique; however, 5% of the genes that initiate transcription in our CL2 embryos overlap previously reported *DUX4* target genes<sup>13,14</sup> (Figure 2C). Interestingly, we observe that *DUXA/B*, two single-copy *DUX* homologs on chromosome 19 and 16 respectively,<sup>15</sup> which are polyadenylated and therefore captured during the scM&T-seq protocol, share similar expression profiles, being readily detectable at the CL2 stage. This suggests that additional *DUX*-family members may influence transcription initiation in cleavage embryos. Two reports have described a role of *Dppa2* and *Dppa4* in regulating *DUX*-driven EGA in mice,<sup>16,17</sup> but such a relationship was not apparent in our human dataset. *DPPA2* was weakly induced in CL2 embryos, and *DPPA4* expression progressively increases from CL3, peaking in blastocysts, illustrating that these two genes are not maternally derived or expressed prior to minor EGA (Figure 2D).

### Major EGA activates pluripotency genes

A subsequent major wave of EGA happens at the CL3 stage. We observed significant upregulation of 1,519 genes and significant downregulation of 986 genes between CL2 and CL3. Analysis of gene terms associated with upregulated transcripts revealed enrichment for an involvement in gene expression, RNA processing, covalent chromatin modifications and gene silencing, translation, and ER protein localization, indicating embryonic competence in transcription and translation processes (Figure 2A).

However, the most pronounced change in expression was observed between CL3 and blastocyst with 2,920 genes differentially expressed, with equal numbers up- and downregulated (1,542 and 1,378, respectively). Key pluripotency-associated factors, including *NANOG* and *ZFP42* (also known as *REX1*), were upregulated from CL3 through to blastocysts. Interestingly, the expression of *GATA6*, which also initiated at the CL3 stage, was present in 74% of ICM cells (cutoff >1 transcripts per million [TPM]), endorsing previously described salt-and-pepper staining patterns at this stage<sup>5</sup> (Table S2). Examination of genes associated with early, mid, and late blastocyst development defined by pseudo-time analysis<sup>6</sup> revealed that blastocyst 4 expressed late markers, suggesting there was developmental asynchrony in our

day 5 embryos with this blastocyst being more advanced (Figure 3A and Table S2). Furthermore, concomitant with the overall increase in transcription up to the blastocyst stage, we observed increased global variation that progresses with developmental stage as revealed by kernel density estimations of single-cell transcriptomes. This observation is consistent with increased transcriptional noise in uncommitted cells of the embryo<sup>18,19</sup> (Figure S1).

### Dynamic markers of pre-lineage specification and potency

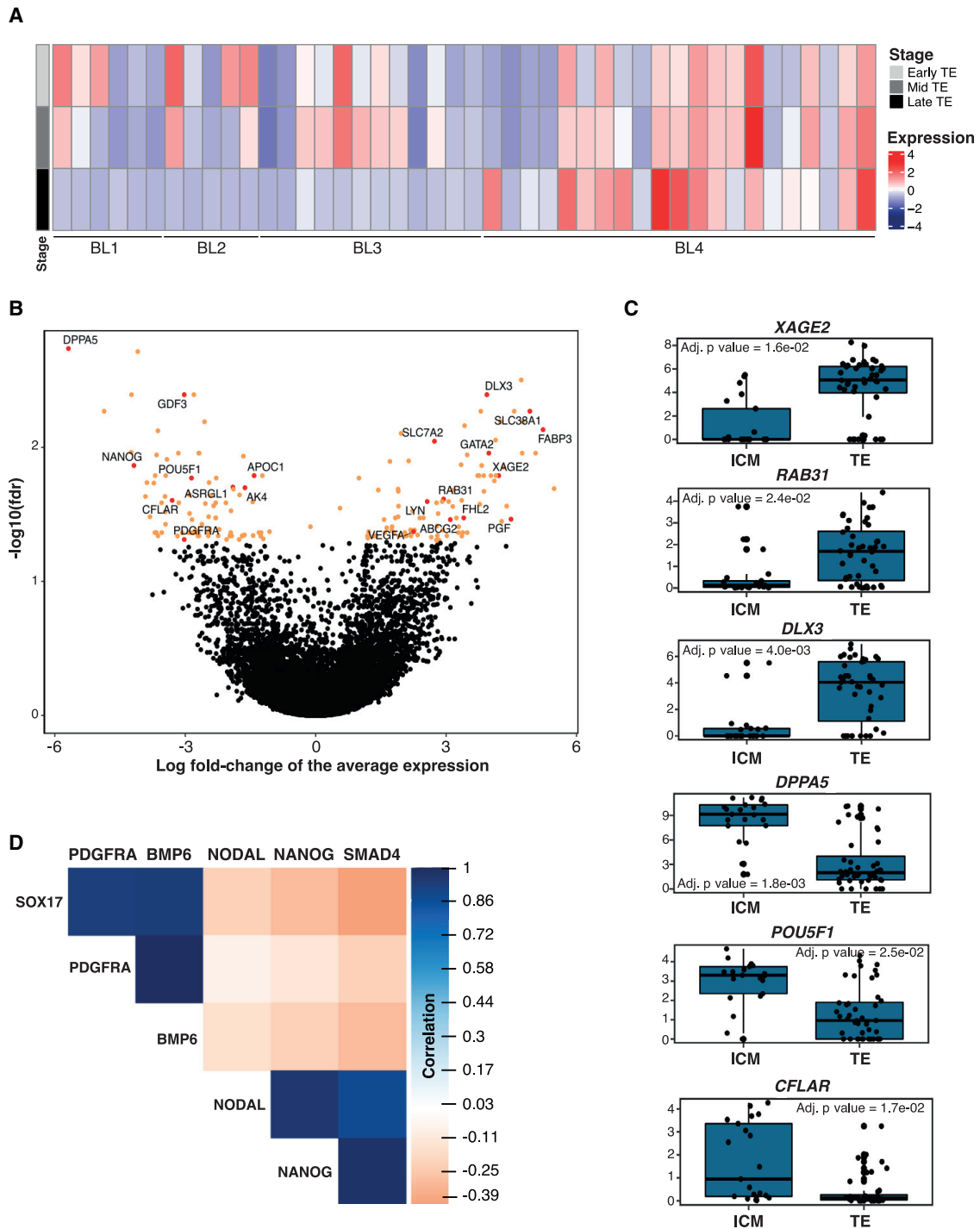
Comparatively little is known about the molecular mechanisms driving early cell fate decision, which results in the segregation of the ICM and TE. This is immediately followed by the ICM differentiation into the pluripotent epiblast (EPI) that forms the embryo proper and the primitive endoderm (PE), which along with the TE becomes the extra-embryonic layer and placenta. Our earlier attempts to classify ICM and TE in blastocysts using unsupervised clustering were not successful (Figure 1B), but following assignment of pTE cells in the ICM fraction using *CCR7* as a marker, differential expression analysis between ICM and mTE + pTE revealed a number of prominently upregulated genes, allowing the identity of TE cells to be confirmed. As highlighted by *XAGE2*, *RAB31*, and *DLX3*, many TE genes had minimal expression in cleavage embryos and were expressed in ICM cells, albeit at lower levels compared with TE (Figures 3B and 3C and Table S2) when all blastocysts were analyzed together. Furthermore, we observed upregulation of traditional TE markers,<sup>4</sup> including *FHL2* and *GATA2* in three TE datasets from the four blastocysts as well as factors important to placenta development, including *SLC38A1*, *PGF*, and *VEGFA* (Figure 3B). Interestingly, *CDX2* expression, a transcription factor essential for mouse TE development,<sup>20,21</sup> was variable in the TE cells, in agreement with the previous staining of human blastocysts<sup>4,22</sup> (Table S2). Furthermore, several pluripotency-associated genes were more abundant in ICM cells than TE, highlighted by *DPPA5*, *POU5F*, *GDF3*, and *NANOG* (Figures 3B and 3C and Table S2).

It has previously been reported that cells undergoing pre-lineage commitment can be identified in the ICM cells with presumptive EPI and PE cells identifiable in late-stage blastocysts. Using proven EPI and PE markers,<sup>4,5</sup> we looked for cells within each embryo that expressed these lineage markers. We observe mutually exclusive expression of these markers in only one embryo, blastocyst 4. Highly correlated expression of the EPI markers (*NANOG*, *NODAL*, and *SMAD4*) was detected in a subset of cells, while another population expressed the PE markers (*SOX17*, *BMP6*, and *PDGFRA*), suggesting that this initial specification had already begun (Figure 3D). However, lineage commitment was restricted to TE, EPI, and PE, as markers of advanced downstream lineages including mesoderm, ectoderm, and endoderm were not detected in any

(B) A pseudo-time was assigned to each cell by filtering a principle curve of the most differentially expressed transcripts to cells in the two-dimensional *t-SNE* subspace.

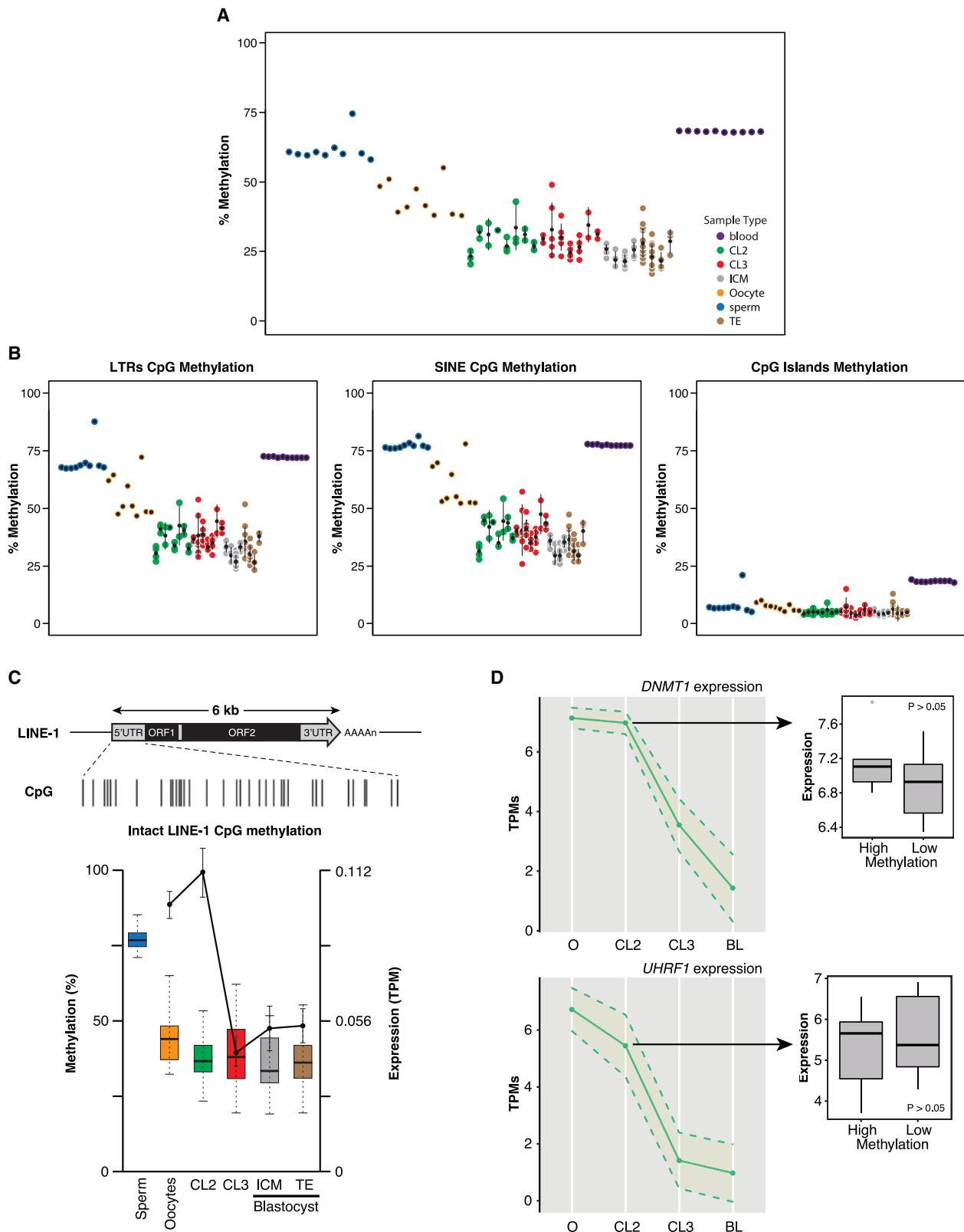
(C) Gene expression profiling of known *DUX4* target genes.

(D) Gene expression dynamics of the *DUX4* single copy homologs, *DUXA* and *DUXB*, mapping to chromosomes 19 and 16, respectively, as well as the pluripotent embryonic transcription factors *DPPA2* and *DPPA4*.



**Figure 3. Assessment of expression variability and lineage commitment in human pre-implantation embryos**

(A) Expression profiling of developmental marker genes associated with early and late blastocyst development in cells isolated from our day 5 blastocysts. (B) Volcano plot representation of up- and downregulated genes comparing TE-derived cells with ICM. Colors represent on a log<sub>2</sub>-scale differential expression with upregulated known TE-associated genes and downregulated pluripotent marks highlighted in red. (C) Boxplot representation of differentially expressed individual marker genes for TE (*XAGE2*, *RAB31*, and *DLX3*) and pluripotency (*DPPA5*, *POU5F1*, and *CFLAR*). The black bar indicates the median of distribution. Whiskers represent the 95% confidence intervals. (D) Heatmaps showing correlation coefficients in blastocyst 4 and the mutually exclusive correlated expression for known epiblast (*NANOG*, *NODAL*, and *SMAD4*) and primitive endoderm (*SOX17*, *BMP6*, and *PDGFRA*) markers in individual cells.



(legend on next page)



ICM cells, confirming that additional differentiation programs had not been initiated.

### Global methylation dynamics at single-cell resolution

Global epigenetic reprogramming occurs during human pre-implantation development. Our group and others have analyzed the fate of germline-derived DNA methylation in human embryos,<sup>23,24</sup> but the dynamic nature and cell-to-cell variability are not fully appreciated since most descriptive studies have been performed on whole embryos. The average methylation for MII oocytes was 32.5%–54.7% (individual oocyte mean 39.5%, SD = 4.8%), which is considerably less than that reported for individual spermatozoa from a previously published dataset.<sup>23</sup> Consistent with literature, oocytes were preferentially methylated in highly transcribed gene bodies (mean expressed 78.6%, SD = 0.5%; mean silent 50.6%, SD = 11.7%) compared with non-expressed genes<sup>23,25,26</sup> (Figures 4A and S2). Tracking the methylation dynamics during development, we confirmed pronounced global demethylation had already occurred in CL2 embryos (individual blastomere cell mean 25.4%, SD = 5.4%), which progressed during pre-implantation development (individual blastomeres from CL3 cell mean 25.2%, SD 5.1%) to blastocysts (individual cell mean 20.3%, SD = 3.3%) (Figure 4A). It must be noted, however, that analyzing methylation using bisulfite-based technologies does not allow us to discriminate 5mC from 5hmC. This is especially poignant for CL2 embryos since our 5mC measurements will be inflated by paternally derived 5hmC. Despite this, embryonic global methylation is considerably lower than in somatic tissues. Previous studies have reported non-CpG methylation to be a prominent feature in oocytes.<sup>23,26</sup> We found non-CpG methylation to be most abundant in oocytes in line with previous reports, with a mean of 2.9% (SD = 0.9%) and 4.3% (SD = 1.3%) for CHG and CHH sites, respectively (where H is adenine, cytosine, or thymine), decreasing progressively in blastocysts (CHG individual cell mean 1.5%, SD = 1.6% and CHH individual cell mean 2.1%, SD = 2.5%).

The pattern of demethylation observed globally is not restricted to individual genomic features but was observed at most sequences investigated. This includes gene bodies, promoters, and enhancer regions, along with all major families of repeat elements, including short interspersed nuclear elements (SINEs), long interspersed nuclear elements (LINEs), and long terminal repeats (LTRs) (Figure 4B). Promoter regions with low CpG density tend to follow the global demethylation pattern, while high CpG density promoters and CpG islands tend to be constitutively hypomethylated (Figures 4B and S3). Recently the methylation dynamics of highly methylated domains have been suggested to differ from partially methylated domains (PMDs), where the latter are characterized by solo-CpG WCGW tetranucleotides (where W is adenine or thymine), late

replication timing, and localization to nuclear lamina.<sup>27</sup> The reported presence of specific patterns of solo-WCGW methylation in gametes led us to further investigate their dynamics during pre-implantation reprogramming. We observed demethylation patterns resembling those obtained for retrotransposons, suggesting that methylation at PMDs/solo-WCGW sequences are also erased (Figure S3) but re-established in a tissue-specific pattern in somatic tissues.

### Concomitant demethylation and expression of LINE-1 elements

Experiments in mice have shown that there is a burst of expression from LINEs, concomitant with initial EGA, that decreases globally by the eight-cell stage, being essential for global chromatin accessibility and gene expression in early development.<sup>28</sup> Since a subset of full-length LINE-1 transcripts possess a poly-A tail,<sup>29</sup> thus allowing for capture during the scM&T-seq protocol, we profiled the expression of these retrotransposons in our sample set. In parallel with the strong demethylation, we detected an increase of LINE-1 expression in CL2 embryos, suggesting that this is an integral program for mammalian development (Figure 4C).

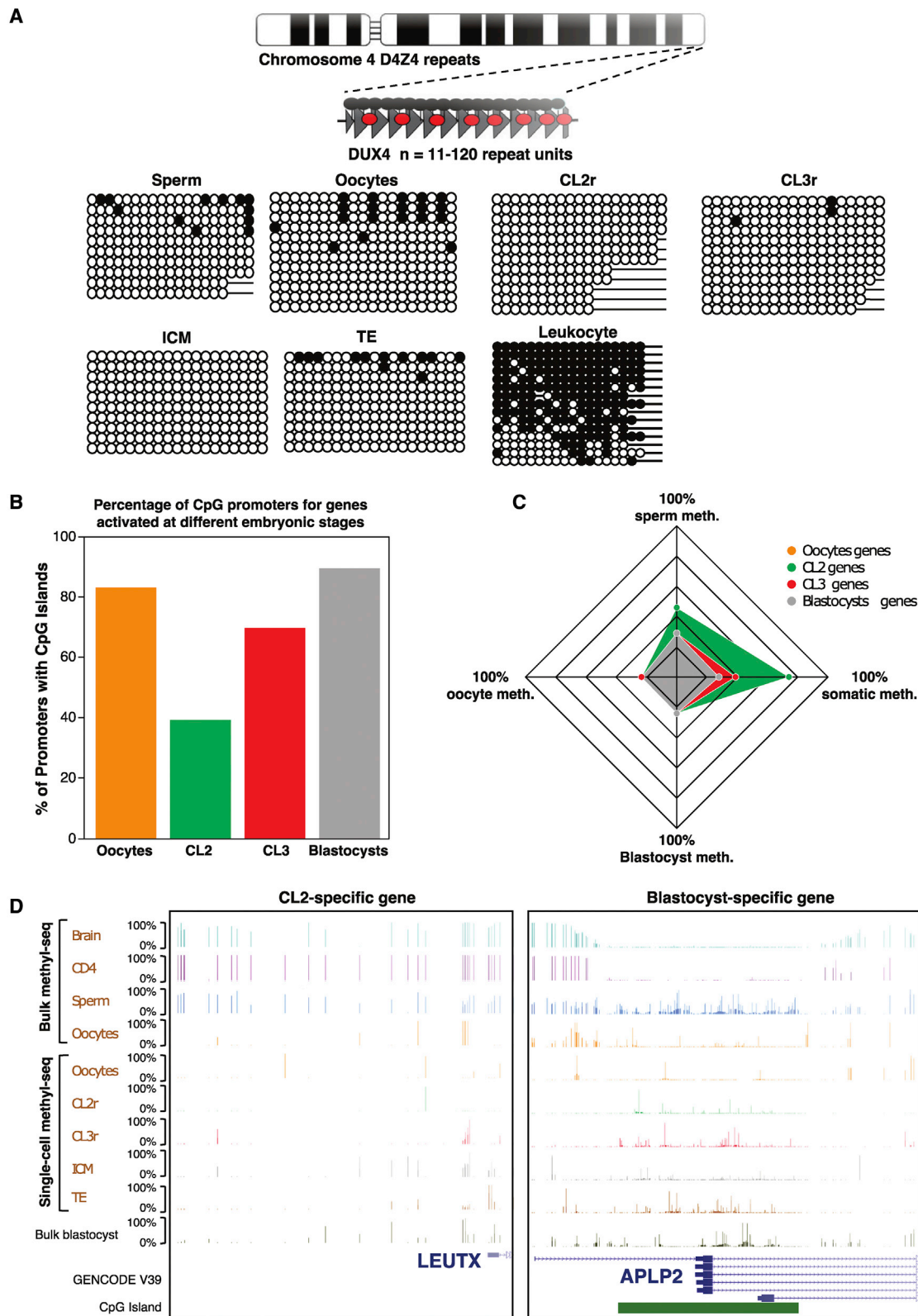
### Specific interplay between DNA methylation and expression during early development

The global changes in methylation that we observed coincide with a developmental time window in which changes in cellular identity and potency occur as a consequence of transcriptional event(s) amalgamated with EGA. We next explored how expression of known epigenetic factors influences global methylation. We confirmed that demethylation during cleavage stages is concomitant with ablated DNMT1 and UHRF1 levels in the presence of TET3, factors responsible for passive replication-dependent and active demethylation, respectively. Furthermore, we observed high expression of *KDM3A/JMJD1A* in oocytes and CL2-specific expression of *KDM4E/JMJD2E*, both of which are demethylases that participate in the active removal of di- and trimethylation of H3K9, repressive histone modifications often found associated with DNA methylation (Table S2). Within our embryo dataset, there is a positive correlation between global methylation and *DNMT1* ( $R = 0.57$ ,  $p = 2.09e-11$ ) and *UHRF1* ( $R = 0.52$ ,  $p = 1.5e-09$ ) expression levels, suggesting that cells with higher expression retain more residual methylation (Figure S3). Furthermore, when analyzing single cells with extreme methylation from each embryo, it was evident that those with more methylation had higher mean levels of *DNMT1* and *UHRF1* at the CL2 stage, although this did not reach significance (Figure 4D).

Subsequently, we explored additional links between methylation and gene expression. Surprisingly, we observe that the majority of CL2-specific genes, including DUX targets, do not

#### Figure 4. Methylation patterns in gametes and pre-implantation embryos

- (A) Average global DNA methylation for individual cells across different developmental stages.  
 (B) Single-cell methylation dynamics of different annotated genomic elements including LTRs, non-LTR SINE retrotransposons, and CpG islands.  
 (C) Methylation and expression dynamics of full-length LINE1 retrotransposons during the pre-implantation window. Whiskers represent SD.  
 (D) Expression dynamics of *DNMT1* and its co-factor *UHRF1* during pre-implantation development. Boxplots show the expression levels for individual cells isolated from CL2 embryos with highest and lowest quartile methylation. Whiskers represent the 95% confidence intervals.



(legend on next page)

possess CpG island promoters, but an intriguing methylation profile. This pattern was not observed for DUX4 itself, as the *D4Z4* repeats harboring this transcription factor were unmethylated in gametes and embryos (Figure 5A). The interval surrounding the transcription start sites of the CL2-specific genes are frequently methylated in sperm, unmethylated in oocytes, and progressively demethylated in cleavage-stage embryos (Figures 5B–5D). Transcriptional silencing following day 3 is methylation independent, as blastocysts remain unmethylated, with DNA methylation only increasing following implantation consistent with their silencing in somatic tissues. Genes that commence expression in day 5 blastocysts largely initiate from transcription start site (TSS) embedded in ubiquitously unmethylated CpG islands that do not show any methylation dynamics during development (Figures 5B–5D).

### DNA methylation signatures associated with TE differentiation

It is well known that differentiation of somatic cells is closely associated with cell-specific epigenetic silencing of lineage-inappropriate genes. We compared the DNA methylation profiles for the top 100 most differentially expressed genes between ICM and TE and found no evidence for differential methylation at intervals overlapping promoter CpG islands, +5 kb upstream of the selected gene promoters, or at the most prominent ATAC-seq peaks<sup>30</sup> (Figure S4) in our CL3, ICM, and TE scBS-seq datasets. This suggests that TE and ICM identity is largely independent of DNA methylation, an observation that has been previously reported for *Elf5*, *Pou5f1*, and *Nanog* in mouse TE and ICM.<sup>31</sup>

### Compromised EGA is widespread in cleavage-stage arrested embryos

Naturally about half of all human pre-implantation embryos fail to implant. To address if embryos subject to cleavage-stage arrest fail to complete either epigenetic reprogramming and EGA, we performed scRNA-seq on 22 embryos (21 arresting between 4 and 10 cells and one at the 16-cell stage) that were morphologically high quality with limited signs of fragmentation (Table S1). These samples had an average of 12,890 (SD = 5.52) expressed transcripts per cell. Unsupervised clustering was carried out by using 10% of the most variable transcripts between all cells, allowing us to resolve the stage of development arrest by determining their clustering with CL2 or CL3 reference embryos described above (CL2r and CL3r, respectively). Interestingly, cells from individual arrested embryos ceased at different points of the cell cycle (Table S1). Within the arrested embryo group, 20% of cells possessed a transcriptome profile indistinguishable from cells derived from the CL2 reference embryos, suggesting

they had completed EGA on day 2 but not day 3 of development (termed CL2a). A further 40% of cells clustered autonomously with the reference CL3 embryos, indicating they had activated their genomes appropriately (CL3a). Interestingly 24% (CL3b) clustered away from CL3r, implying initial CL2 activation proceeded appropriately but faltered at subsequent stages, whereas 16% clustered independently of both reference embryo sets, signifying they had suffered catastrophic EGA failure (CL2b) (Figures 6A and 6B). Supporting this, CL2b cells cluster with oocytes and separate from CL2r/CL2a cells when using oocyte- and CL2-specific gene expression for t-SNE analysis (Figure 6C). Blastomeres from the majority of embryos (14/21) were uniform in their EGA status, with five embryos showing asynchrony between developmental stages (both D2 and D3 calls), while two embryos were mosaic for developmental stage and EGA status (i.e., embryo Q possesses 2 x CL2a, 2 x CL2b, and 4 x CL3a cells) (Table S1).

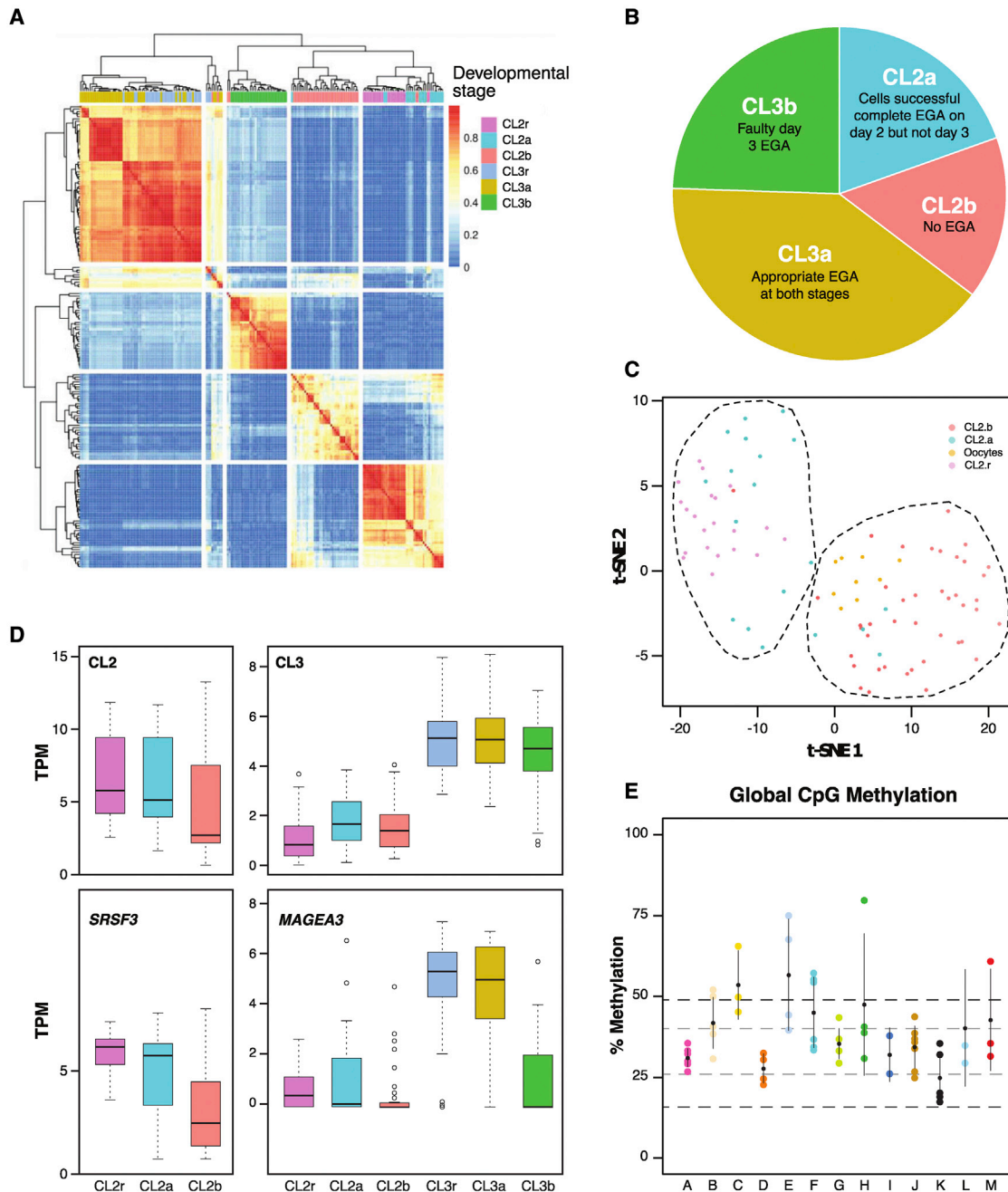
Next, we focused on determining the differential expression profile of transcripts that we had previously identified as either specifically activated or expressed in the reference embryo datasets. This revealed that the overall expression levels of CL2a arrested cells were similar to CL2r, although there was often greater variance, while CL2b cells were significantly less abundant (Figure 6D). Importantly, CL2b embryos did not activate a different set of transcripts compared with reference embryos with GO terms including G2/M phase transition, oocyte differentiation, and meiotic progression, signifying their similarity to oocytes. Interrogation of the most differential expressed transcripts revealed that genes involved in transcriptional initiation and alternative splicing (*TAF11*, *SSU72*, *EIF4A1*, *SRSF3*, and *ELOF1*) as well as H3K9 demethylation (*KDM4E* and *KDM4F*) were overrepresented, suggesting these events are critical during early cleavage-stage development. Similar comparisons revealed significant reductions in expression levels for CL3-specific activated or expressed genes in the CL3b embryos, including zinc-finger (*ZNF280A*, *ZNF669*, *ZNF735*, and *ZNF878*) and *MAGEA* genes (Figure 6D and Table S2), with GO term analysis revealing RNA processing and translation were still features of CL3 arrested cells.

### Aberrant DNA demethylation in cleavage-stage arrested embryos

To determine whether aberrant epigenetic reprogramming occurs in the arrested embryos, we quantified DNA methylation in a subset of arrested embryos, primarily those with transcription profiles independent from the reference dataset (Table S1). The scBS-seq libraries were sequenced to an average of 2.8 M (SD = 2.2 M) uniquely mapped reads per cell. Of the 13 embryos assessed, eight presented with significantly

#### Figure 5. Locus-specific methylation profiling

- (A) Example of strand-specific methylation as resolved by bisulfite PCR, sub-cloning, and direct Sanger sequencing at *D4Z4* repeats for sperm, oocytes, reference embryos (CL2r and CL3r) ICM, TE, and blood leukocytes. Each circle represents a single CpG dinucleotide on the strand, (●) a methylated cytosine, (○) unmethylated cytosines.
- (B) Bar graphs show the percentage of CpG island promoters for genes activated at different developmental stages.
- (C) Radar chart highlighting unique sperm and somatic tissue methylation signatures for genes associated with CL2 activation.
- (D) Methylation profiling for *LEUTX*, a CL2 non-CpG island promoter gene, and *APLP2*, a blastocyst expressed gene originating expression from a CpG island promoter.



**Figure 6. Expression and methylation profiling of cleavage-stage embryos following spontaneous developmental arrest**

(A) Heatmap showing clustering of individual cells from reference and arrest embryos using the 10% most variable genes. (B) Distribution of arrested cells according to aberrant EGA classification. (C) *t*-SNE analysis of oocyte and CL2-specific markers showing clustering of oocytes with CL2b and CL2r with CL2a cells. (D) Differential expression of genes exhibiting CL2r-specific or CL3r-specific activation in cells isolated from arrested embryos. The top panel represents all genes, whereas the lower panel highlights individual gene examples. Whiskers represent SD. (E) Average global DNA methylation for individual cells isolated from embryos designated as CL3b. The dots represent the mean of all cells from an individual embryo, while the black dashed lines depict the methylation range and gray dashed line the IQR for the equivalently aged reference embryos (CL3r).

increased global methylation levels when compared with the CL3r dataset. The five embryos with largely normal methylomes possessed mixed CL2b/CL3b expression signatures. This was

present at all genomic features investigated (Figures 6E and S5). Quantitative pyrosequencing for LINE-1 elements confirmed the sequencing observations that arrested embryos retained

higher levels of DNA methylation compared with appropriately aged controls (Figure S5) and that levels were highly correlated between techniques (Pearson's  $r = 0.73$ ). Interestingly, both *DNMT1* and *UHRF1* show significantly decreased expression levels (*DNMT1*  $p = 8.9e-13$ ; *UHRF1*  $p = 1.1e-08$ ) in CL2b when compared with CL2r. Furthermore, it was still evident that those cells with more methylation had higher levels of *DNMT1* and *UHRF1*, although expression levels were lower than CL2r (Figure S5), although the difference was not significant.

### Detecting aneuploidy cells

Chromosomal abnormalities are a major contributor to human reproductive failure, with embryonic aneuploidy being a major contributor to implantation failure, especially in mothers of advanced maternal age.<sup>32</sup> Aneuploidy is observed in a mosaic state, even in embryos that contribute to successful pregnancies.<sup>33</sup> Pre-implantation genetic testing for aneuploidy (PGT-A) has been scrutinized recently<sup>34</sup> since embryo selection is generally reliant on biopsies comprising just a few cells and, as such, may underestimate the frequency of mitotic aneuploidies.<sup>35</sup> Therefore, mitotically derived, mosaic aneuploidies are only truly detectable with single-cell technologies. We determined the chromosomal landscape of each embryo in our cohort by utilizing the read counts of our bisulfite sequencing datasets. When coverage was sufficient, this was also performed for individual cells, allowing us to detect aneuploidies for each embryo as well as aberrations restricted to individual cells. Reassuringly, our reference embryos had significantly fewer aneuploidy events compared with the arrested cohort, with the highest irregular chromosome complements identified in the arrested embryos (Figure S6 and Table S1). When the analysis was performed at single-cell rather than at embryo resolution, cell-specific aberrations were detected consistent with mosaicism (Figure S6).

## DISCUSSION

Early embryonic development represents a unique time during which the parental genomes of arguably the most terminally differentiated cells, the sperm and oocyte, are reprogrammed following fertilization to give rise to a totipotent embryo. Concomitant with epigenetic reprogramming that encompasses global DNA demethylation,<sup>3</sup> paternal protamine exchange,<sup>2</sup> redistribution of histone modifications,<sup>36</sup> and gradual formation of higher-order architecture<sup>37</sup> is the initiation of transcription from the embryonic genome. It is, therefore, not surprising that a failure in any of these co-dependent processes could lead to early embryonic failure, which occurs for over 50% of cleavage-stage embryos.<sup>38</sup> Previous studies have suggested that high rates of aneuploidy are a major cause of this arrest, but recent studies have questioned this as expression from trisomic chromosomes seems to be subject to dosage compensation in blastocysts,<sup>39</sup> which may in part be linked to chromosome-specific methylation profiles.<sup>40</sup> Here, we present data supporting the hypothesis that early developmental arrest is frequently associated with genome-wide failure of EGA. Previous observations are controversial, with initial investigations suggesting that EGA occurred normally, regard-

less of the morphological quality of cell number,<sup>41</sup> and developmental demise is not linked to chromosomal abnormalities or EGA failure, but misregulation of individual genes. However, it must be noted that the characterization of embryos in the aforementioned study was performed using qRT-PCR targeting only two genes subject to EGA. More recently, whole embryo RNA-seq profiling suggest that EGA failure may account for embryo demise,<sup>42</sup> so our study adds valuable knowledge to the intricate expression profiles required for development and has allowed us to catalog events associated with cleavage-stage embryo arrest. Endorsing our observations, Asami and colleagues<sup>43</sup> recently reported disrupted genome activation at the pronuclear stage of human development, suggesting that abnormal genome activation possibly occurs in early-stage arrested embryos. It would be fascinating to determine if cell-specific variation in EGA is linked to advanced maternal age, since a previous study from Kawai and colleagues<sup>44</sup> reported that maternal age had a strong impact on expression, with more than 800 genes being less abundant in embryos from older women. Since many processes involved in the MZT are reliant on stored maternally derived factors, oocyte quality is unsurprisingly critical in determining the outcome of early embryonic development, but further work on larger embryo cohorts is warranted to confirm these observations.

By employing the multi-omic scM&T-seq technology, we were able to analyze the dynamic remodeling of DNA methylation, as well as to determine expression status, to understand if epigenetic reprogramming influences developmental progression. We observe that the majority of demethylation has occurred by day 2 in agreement with immunofluorescence<sup>45</sup> and previous single-cell whole-genome bisulfite sequencing in human embryos.<sup>23,46</sup> The unique methylation profile immediately following fertilization is permissive to transcription, as highlighted by the brief burst of LINE-1 expression and the induction of the *DUX4* transcription factor, which is partly responsible for activating a subset of CL2-specific transcripts. However, as the *DUX4-D4Z4* repeats are unmethylated in gametes, additional, yet to be identified, maternally derived factors are ultimately responsible for the burst of *DUX4* expression in humans. Given the dynamic and variable nature of global DNA demethylation during pre-implantation development and the low success rates of early human embryos, we interrogated the single-cell methylation profiles in a subset of arrested embryos as identified by non-invasive time-lapse imaging. Over half of the embryos assessed presented with global hypermethylation, a pattern comparable with recent studies.<sup>23,45,46</sup> Importantly, these embryos had aberrant expression profiles suggesting that unsuccessful epigenetic reprogramming may subsequently influence EGA. It will be important to determine if other epigenetic events, including chromatin reorganization and remodeling of histone modifications, also influence the success of an embryo to transition the cleavage stage.

In addition to better characterizing the methylation and expression dynamics related to normal pre-implantation human embryo development, our study has added to our understanding of why embryos undergo spontaneous cleavage-stage arrest. Whether such developmental failures are due to aberrant EGA

or can be attributed to just a few crucial genes remains to be determined. However, our datasets catalog the timing of EGA and demethylation events, which will be useful in developing biomarkers to improve our capacity to select the most competent embryos for transfer or cryopreservation.

### Limitations of the study

Obtaining high-quality human embryos for research is fraught with challenges, and their characterization at single-cell resolution is still in its infancy. Here, we describe an additional reference dataset that not only includes scRNA-seq but has accompanying DNA methylation, which revealed variability within individual cells of an embryo. scBS-seq profiling is an emerging field within epigenetic research and is currently hampered by the random nature of genome coverage between samples. Despite this, our analysis of defined genetic features revealed that key developmental programs, such as TE differentiation, are largely independent of *de novo* methylation. To fully understand the intricacies of TE differentiation, early development should be monitored using time-lapse techniques. Arrested human cleavage-stage embryos with good morphology are extremely rare, and to obtain the numbers used in our study required the time-lapse imaging of several thousand IVF embryos. Due to local legislation, the earliest that arrested embryos can be collected is D5, and it is therefore not possible to address if the profiles we observe are due to extended cultivation. Future studies will rely on wider collaborative efforts to obtain additional arrested embryos associated with different infertility parameters, including advanced maternal age and different paternal infertility etiologies. In addition, not all arrested embryos had accompanying DNA methylation profiles, and future work should include both omic datasets as well as immunofluorescence to ensure the observed transcriptional profiles translate to quantitative protein measurements.

### STAR★METHODS

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

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Human oocytes and pre-implantation embryos
- **METHOD DETAILS**
  - Human oocytes and pre-implantation embryos
  - Embryo culture and time-lapse imaging
  - Single-cell isolation
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Single-cell sequencing
  - Sequence data processing and raw data analysis
  - Sequence features
  - Data analysis
  - Locus-specific methylation analysis

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2023.112100>.

### ACKNOWLEDGMENTS

We would like to thank all the donating couples that participated in this study and the clinical staff at IVI Valenica. The Monk lab received funding from the Spanish Ministry of Economy and Competitiveness (MINECO; BFU2014-53093-R and BFU2017-85571-R), co-funded with the European Union Regional Development Fund (FEDER), the H2020 HUTER project from the European Union's Horizon 2020 research and innovation program under grant agreement No 874867, and the UK Biotechnology and Biological Sciences Research Council (BB/V016156/1). A.M.S. was a recipient of a FPI PhD studentship from MINECO (BES-2015-072547), and D.D. is a BBSRC DTP PhD fellow. The Meseguer lab received funding from the Spanish Ministry of Economy and Competitiveness (CDTI IDI-20180248), co-funded with the European Union Regional Development Fund (FEDER). R.D.G.B. was supported by Agència Valenciana de la Innovació and European Social Fund (ACIF/2018/034). Work in the Kelsey lab was funded by the BBSRC (BBS/E/B/000C0421) and the Medical Research Council (MR/K011332/1, MR/S000437/1).

### AUTHOR CONTRIBUTIONS

J.R.H.M. and C.B. performed all bioinformatic analyses. S.C. and G.K. generated the scM&T-seq libraries. C.S., M.E.P.L., and J.V.M. were responsible for reference embryo selection and isolation of single cells. R.D.G.B. and M.M. assessed time-lapse imaging data and isolated single cells from arrested embryos. D.D. and A.M.S. performed verification-related experiments. D.M. conceived the project, helped with data interpretation, and prepared the manuscript with help from all authors.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

### INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: June 14, 2022

Revised: October 14, 2022

Accepted: January 26, 2023

Published: February 9, 2023

### REFERENCES

1. Daškevičiūtė, D., Sanchez-Delgado, M., and Monk, D. (2022). Epigenetics from oocytes to embryos Genetic Diagnostic technologies in Reproductive Medicine: improving patient success rates and Infant health. In Carlos Simon & Carmen Rubio (CRC Press), pp109–120. <https://doi.org/10.1201/9781003024941>.
2. Jenkins, T.G., and Carrell, D.T. (2012). Dynamic alterations in the paternal epigenetic landscape following fertilization. *Front. Genet.* 3, 143. <https://doi.org/10.3389/fgene.2012.00143>.
3. Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* 484, 339–344. <https://doi.org/10.1038/nature10960>.
4. Blakeley, P., Fogarty, N.M.E., del Valle, I., Wamaitha, S.E., Hu, T.X., Elder, K., Snell, P., Christie, L., Robson, P., and Niakan, K.K. (2015). Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* 142, 3151–3165. <https://doi.org/10.1242/dev.123547>.
5. Petropoulos, S., Edsgård, D., Reinius, B., Deng, Q., Panula, S.P., Code-luppi, S., Plaza Reyes, A., Linnarsson, S., Sandberg, R., and Lanner, F.

- (2016). Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell* 165, 1012–1026. <https://doi.org/10.1016/j.cell.2016.03.023>.
6. Meistemann, D., Bruneau, A., Loubersac, S., Reignier, A., Firmin, J., François-Campion, V., Kilens, S., Lelièvre, Y., Lammers, J., Feyeux, M., et al. (2021). Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification. *Cell Stem Cell* 28, 1625–1640.e6. <https://doi.org/10.1016/j.stem.2021.04.027>.
  7. Angermueller, C., Clark, S.J., Lee, H.J., Macaulay, I.C., Teng, M.J., Hu, T.X., Krueger, F., Smallwood, S., Ponting, C.P., Voet, T., et al. (2016). Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods* 13, 229–232. <https://doi.org/10.1038/nmeth.3728>.
  8. Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* 20, 1131–1139. <https://doi.org/10.1038/nsmb.2660>.
  9. Zheng, W., Zhou, Z., Sha, Q., Niu, X., Sun, X., Shi, J., Zhao, L., Zhang, S., Dai, J., Cai, S., et al. (2020). Homozygous mutations in BTG4 cause zygotic cleavage failure and female infertility. *Am. J. Hum. Genet.* 107, 24–33. <https://doi.org/10.1016/j.ajhg.2020.05.010>.
  10. Sha, Q.Q., Yu, J.L., Guo, J.X., Dai, X.X., Jiang, J.C., Zhang, Y.L., Yu, C., Ji, S.Y., Jiang, Y., Zhang, S.Y., et al. (2018). CNOT6L couples the selective degradation of maternal transcripts to meiotic cell cycle progression in mouse oocyte. *EMBO J.* 37, e99333. <https://doi.org/10.15252/embj.201899333>.
  11. Ivanova, I., Much, C., Di Giacomo, M., Azzi, C., Morgan, M., Moreira, P.N., Monahan, J., Carrieri, C., Enright, A.J., and O’Carroll, D. (2017). The RNA m(6)A reader YTHDF2 is essential for the post-transcriptional regulation of the maternal transcriptome and oocyte competence. *Mol. Cell* 67, 1059–1067.e4. <https://doi.org/10.1016/j.molcel.2017.08.003>.
  12. Sonesson, C., and Robinson, M.D. (2018). Bias, robustness and scalability in single-cell differential expression analysis. *Nat. Methods* 15, 255–261. <https://doi.org/10.1038/nmeth.4612>.
  13. De Iaco, A., Planet, E., Coluccio, A., Verp, S., Duc, J., and Trono, D. (2017). DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.* 49, 941–945. <https://doi.org/10.1038/ng.3858>.
  14. Hendrickson, P.G., Doráis, J.A., Grow, E.J., Whiddon, J.L., Lim, J.W., Wike, C.L., Weaver, B.D., Pflueger, C., Emery, B.R., Wilcox, A.L., et al. (2017). Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* 49, 925–934. <https://doi.org/10.1038/ng.3844>.
  15. Leidenroth, A., and Hewitt, J.E. (2010). A family history of DUX4: phylogenetic analysis of DUXA, B, C and Duxbl reveals the ancestral DUX gene. *BMC Evol. Biol.* 10, 364. <https://doi.org/10.1186/1471-2148-10-364>.
  16. Eckersley-Maslin, M., Alda-Catalinas, C., Blotenburg, M., Kreibich, E., Krueger, C., and Reik, W. (2019). Dppa2 and Dppa4 directly regulate the Dux-driven zygotic transcriptional program. *Genes Dev.* 33, 194–208. <https://doi.org/10.1101/gad.321174.118>.
  17. De Iaco, A., Coudray, A., Duc, J., and Trono, D. (2019). DPPA2 and DPPA4 are necessary to establish a 2C-like state in mouse embryonic stem cells. *EMBO Rep.* 20, e47382. <https://doi.org/10.15252/embr.201847382>.
  18. Piras, V., Tomita, M., and Selvarajoo, K. (2014). Transcriptome-wide variability in single embryonic development cells. *Sci. Rep.* 4, 7137. <https://doi.org/10.1038/srep07137>.
  19. Mohammed, H., Hernando-Herraez, I., Savino, A., Scialdone, A., Macaulay, I., Mulas, C., Chandra, T., Voet, T., Dean, W., Nichols, J., et al. (2017). Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. *Cell Rep.* 20, 1215–1228. <https://doi.org/10.1016/j.celrep.2017.07.009>.
  20. Wu, G., Gentile, L., Fuchikami, T., Sutter, J., Psathaki, K., Esteves, T.C., Araúzo-Bravo, M.J., Ortmeier, C., Verberk, G., et al. (2010). Initiation of trophectoderm lineage specification in mouse embryos is independent of Cdx2. *Development* 137, 4159–4169. <https://doi.org/10.1242/dev.056630>.
  21. Deglincerti, A., Croft, G.F., Pietila, L.N., Zernicka-Goetz, M., Siggia, E.D., and Brivanlou, A.H. (2016). Self-organization of the in vitro attached human embryo. *Nature* 533, 251–254. <https://doi.org/10.1038/nature17948>.
  22. Chen, Y., Wang, K., Gong, Y.G., Khoo, S.K., and Leach, R. (2013). Roles of CDX2 and EOMES in human induced trophoblast progenitor cells. *Biochem. Biophys. Res. Commun.* 437, 197–202. <https://doi.org/10.1016/j.bbrc.2012.12.135>.
  23. Guo, H., Zhu, P., Yan, L., Li, R., Hu, B., Lian, Y., Yan, J., Ren, X., Lin, S., Li, J., et al. (2014). The DNA methylation landscape of human early embryos. *Nature* 511, 606–610. <https://doi.org/10.1038/nature13544>.
  24. Sanchez-Delgado, M., Court, F., Vidal, E., Medrano, J., Monteagudo-Sánchez, A., Martín-Trujillo, A., Tayama, C., Iglesias-Platas, I., Kondova, I., Bontrop, R., et al. (2016). Human oocyte-derived methylation differences persist in the placenta revealing widespread transient imprinting. *PLoS Genet.* 12, e1006427. <https://doi.org/10.1371/journal.pgen.1006427>.
  25. Smallwood, S.A., Tomizawa, S.I., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., Sato, S., Hata, K., Andrews, S.R., and Kelsey, G. (2011). Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat. Genet.* 43, 811–814. <https://doi.org/10.1038/ng.864>.
  26. Yu, B., Dong, X., Gravina, S., Kartal, Ö., Schimmel, T., Cohen, J., Tortorello, D., Zody, R., Hawkins, R.D., and Vijg, J. (2017). Genome-wide, single-cell DNA methylomics reveals increased non-CpG methylation during human oocyte maturation. *Stem Cell Rep.* 9, 397–407. <https://doi.org/10.1016/j.stemcr.2017.05.026>.
  27. Zhou, W., Dinh, H.Q., Ramjan, Z., Weisenberger, D.J., Nicolet, C.M., Shen, H., Laird, P.W., and Berman, B.P. (2018). DNA methylation loss in late-replicating domains is linked to mitotic cell division. *Nat. Genet.* 50, 591–602. <https://doi.org/10.1038/s41588-018-0073-4>.
  28. Jachowicz, J.W., Bing, X., Pontabry, J., Bošković, A., Rando, O.J., and Torres-Padilla, M.E. (2017). LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat. Genet.* 49, 1502–1510. <https://doi.org/10.1038/ng.3945>.
  29. Roy-Engel, A.M. (2012). LINEs, SINEs and other retroelements: do birds of a feather flock together? *Front. Biosci.* 17, 1345–1361. <https://doi.org/10.2741/3991>.
  30. 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>.
  31. Nakanishi, M.O., Hayakawa, K., Nakabayashi, K., Hata, K., Shiota, K., and Tanaka, S. (2012). Trophectoderm-specific DNA methylation occurs after the segregation of the trophectoderm and inner cell mass in the mouse preimplantation embryo. *Epigenetics* 7, 173–182. <https://doi.org/10.4161/epi.7.2.18962>.
  32. Rienzi, L., Cimadomo, D., Delgado, A., Minasi, M.G., Fabozzi, G., Gallego, R.D., Stoppa, M., Bellver, J., Giancani, A., Esbert, M., et al. (2019). Time of morulation and trophectoderm quality are predictors of a live birth after euploid blastocyst transfer: a multicentre study. *Fertil. Steril.* 112, 1080–1093.e1. <https://doi.org/10.1016/j.fertnstert.2019.07.1322>.
  33. Bielanska, M., Jin, S., Bernier, M., Tan, S.L., and Ao, A. (2005). Diploid-an euploid mosaicism in human embryos cultured to the blastocyst stage. *Fertil. Steril.* 84, 336–342. <https://doi.org/10.1016/j.fertnstert.2005.03.031>.
  34. Gleicher, N., Albertini, D.F., Patrizio, P., Orvieto, R., and Adashi, E.Y. (2022). The uncertain science of preimplantation and prenatal genetic testing. *Nat. Med.* 28, 442–444. <https://doi.org/10.1038/s41591-022-01712-7>.
  35. Starostik, M.R., Sosina, O.A., and McCoy, R.C. (2020). Single-cell analysis of human embryos reveals diverse patterns of aneuploidy and mosaicism. *Genome Res.* 30, 814–825. <https://doi.org/10.1101/gr.262774.120>.

36. 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>.
37. Du, Z., Zheng, H., Huang, B., Ma, R., Wu, J., Zhang, X., He, J., Xiang, Y., Wang, Q., Li, Y., et al. (2017). Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* 547, 232–235. <https://doi.org/10.1038/nature23263>.
38. French, D.B., Sabanegh, E.S., Jr., Goldfarb, J., and Desai, N. (2010). Does severe teratozoospermia affect blastocyst formation, live birth rate, and other clinical outcome parameters in ICSI cycles? *Fertil. Steril.* 93, 1097–1103. <https://doi.org/10.1016/j.fertnstert.2008.10.051>.
39. Licciardi, F., Lhakhang, T., Kramer, Y.G., Zhang, Y., Heguy, A., and Tsirigos, A. (2018). Human blastocysts of normal and abnormal karyotypes display distinct transcriptome profiles. *Sci. Rep.* 8, 14906. <https://doi.org/10.1038/s41598-018-33279-0>.
40. Yang, M., Tao, X., Scott, K., Zhan, Y., Scott, R.T., and Seli, E. (2021). Evaluation of genome-wide DNA methylation profile of human embryos with different developmental competences. *Hum. Reprod.* 36, 1682–1690. <https://doi.org/10.1093/humrep/deab074>.
41. Dobson, A.T., Raja, R., Abeyta, M.J., Taylor, T., Shen, S., Haqq, C., and Pera, R.A.R. (2004). The unique transcriptome through day 3 of human preimplantation development. *Hum. Mol. Genet.* 13, 1461–1470. <https://doi.org/10.1093/hmg/ddh157>.
42. 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 dysfunction underlies 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>.
43. Asami, M., Lam, B.Y.H., Ma, M.K., Rainbow, K., Braun, S., Ver Milyea, M.D., Yeo, G.S.H., and Perry, A.C.F. (2022). Human embryonic genome activation initiates at the one-cell stage. *Cell Stem Cell* 29, 209–216.e4. <https://doi.org/10.1016/j.stem.2021.11.012>.
44. Kawai, K., Harada, T., Ishikawa, T., Sugiyama, R., Kawamura, T., Yoshida, A., Tsutsumi, O., Ishino, F., Kubota, T., and Kohda, T. (2018). Parental age and gene expression profiles in individual human blastocysts. *Sci. Rep.* 8, 2380. <https://doi.org/10.1038/s41598-018-20614-8>.
45. Arand, J., Reijo Pera, R.A., and Wossidlo, M. (2021). Reprogramming of DNA methylation is linked successful human preimplantation development. *Histochem. Cell Biol.* 156, 197–207. <https://doi.org/10.1007/s00418-021-02008-6>.
46. Zhu, P., Guo, H., Ren, Y., Hou, Y., Dong, J., Li, R., Lian, Y., Fan, X., Hu, B., Gao, Y., et al. (2018). Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat. Genet.* 50, 12–19. <https://doi.org/10.1038/s41588-017-0007-6>.
47. Larsen, M., Rost, S., El Hajji, N., Ferbert, A., Deschauer, M., Walter, M.C., Schoser, B., Tacik, P., Kress, W., and Müller, C.R. (2015). Diagnostic approach for FSHD revisited: SMCHD1 mutations cause FSHD2 and act as modifiers of disease severity in FSHD1. *Eur. J. Hum. Genet.* 23, 808–816. <https://doi.org/10.1038/ejhg.2014.191>.
48. Martin-Trujillo, A., Vidal, E., Monteagudo-Sánchez, A., Sanchez-Delgado, M., Moran, S., Hernandez Mora, J.R., Heyn, H., Guitart, M., Esteller, M., and Monk, D. (2017). Copy number rather than epigenetic alterations are the major dictator of imprinted methylation in tumors. *Nat. Commun.* 8, 467. <https://doi.org/10.1038/s41467-017-00639-9>.
49. Coello, A., Campos, P., Remohí, J., Meseguer, M., and Cobo, A. (2016). A combination of hydroxypropyl cellulose and trehalose as supplement for vitrification of human oocytes: a retrospective cohort study. *J. Assist. Reprod. Genet.* 33, 413–421. <https://doi.org/10.1007/s10815-015-0633-9>.
50. Macaulay, I.C., Haerty, W., Kumar, P., Li, Y.L., Hu, T.X., Teng, M.J., Goolam, M., Saurat, N., Coupland, P., Shirley, L.M., et al. (2015). G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat. Methods* 12, 519–522. <https://doi.org/10.1038/nmeth.3370>.
51. Clark, S.J., Smallwood, S.A., Lee, H.J., Krueger, F., Reik, W., and Kelsey, G. (2017). Genome-wide base-resolution mapping of DNA methylation in single cells using single-cell bisulfite sequencing (scBS-seq). *Nat. Protoc.* 12, 534–547. <https://doi.org/10.1038/nprot.2016.187>.
52. Linker, S.M., Urban, L., Clark, S.J., Chhatrivala, M., Amatya, S., McCarthy, D.J., Ebersberger, I., Vallier, L., Reik, W., Stegle, O., and Bonder, M.J. (2019). Combined single-cell profiling of expression and DNA methylation reveals splicing regulation and heterogeneity. *Genome Biol.* 20, 30. <https://doi.org/10.1186/s13059-019-1644-0>.
53. McCarthy, D.J., Campbell, K.R., Lun, A.T.L., and Wills, Q.F. (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* 33, 1179–1186. <https://doi.org/10.1093/bioinformatics/btw777>.
54. Lun, A.T.L., McCarthy, D.J., and Marioni, J.C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res.* 5, 2122. <https://doi.org/10.12688/f1000research.9501.2>.
55. Kiselev, V.Y., Kirschner, K., Schaub, M.T., Andrews, T., Yiu, A., Chandra, T., Natarajan, K.N., Reik, W., Barahona, M., Green, A.R., and Hemberg, M. (2017). SC3: consensus clustering of single-cell RNA-seq data. *Nat. Methods* 14, 483–486. <https://doi.org/10.1038/nmeth.4236>.
56. 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>.
57. 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>.
58. Xie, W., Schultz, M.D., Lister, R., Hou, Z., Rajagopal, N., Ray, P., Whitaker, J.W., Tian, S., Hawkins, R.D., Leung, D., et al. (2013). Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* 153, 1134–1148. <https://doi.org/10.1016/j.cell.2013.04.022>.
59. Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470, 279–283. <https://doi.org/10.1038/nature09692>.
60. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. <https://doi.org/10.1038/75556>.
61. Scialdone, A., Natarajan, K.N., Saraiva, L.R., Proserpio, V., Teichmann, S.A., Stegle, O., Marioni, J.C., and Buettner, F. (2015). Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods* 85, 54–61. <https://doi.org/10.1016/j.ymeth.2015.06.021>.
62. Garvin, T., Aboukhalil, R., Kendall, J., Baslan, T., Atwal, G.S., Hicks, J., Wiggler, M., and Schatz, M.C. (2015). Interactive analysis and assessment of single-cell copy-number variations. *Nat. Methods* 12, 1058–1060. <https://doi.org/10.1038/nmeth.3578>.



## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Bacterial and virus strains</b>		
pGEM-T Easy Vector System and JM109 Competent bacteria	Promega	A1380
<b>Biological samples</b>		
Human pre-implantation embryos and oocytes	The fertility clinic Instituto Valenciano de Infertilidad (IVI), Valencia, Spain	Authorized projects: 4/2014 & 10/2017
Leukocyte DNA samples for methylation profiling	IDIBELL	Authorized project: PR292/14
<b>Chemicals, peptides, and recombinant proteins</b>		
Fertilization medium	Origio, Cooper Surgical®	MEO163/83010060A
Pre-equilibrated EmbryoSlides	EmbryoSlide®, Vitrolife	FL0392/FT-S-ES-D
EmbryoScope single-step medium	Gems, Genea Biomedx®	MEO134/UEE1GE33
1% polyvinylpyrrolidone	Sigma-Aldrich	9003-39-8
Accutase medium	Chemicon - Sigma	SF006
AMPure XP beads	Beckman Coulter	A63882
EZ Methylation Direct bisulphite reagent	Zymo	D5021
<b>Deposited data</b>		
Oocyte and reference embryo scRNA-seq	This study; PMID: 33053156	NCBI-BioProject PRJNA630371
Oocyte and reference embryo scBS-seq	This study	NCBI-BioProject PRJNA631209
Arrested embryo scRNA-seq	This study	NCBI-BioProject PRJNA813525
Arrested embryo scBS-seq	This study	NCBI-BioProject PRJMA81367
Human embryo ATAC-seq	●PMID: 30664750	NCBI Sequence Read Archive SRP163205
Bulk BS-seq for brain		NCBI GEO GSM916050
Bulk BS-seq for blood	●PMID: 22689993	NCBI GEO GSM848927
Bulk BS-seq for preimplantation embryos	●PMID: 25501653	Japanese Genotype-phenotype Archive JGAS00000000006
Bulk BS-seq for sperm	PMID: 28899353	NCBI GEO GSE30340
<b>Oligonucleotides</b>		
D4Z4 bisulphite PCR primer F- GGGTTGAGGGTTGGGTTTATA	Ref. <sup>47</sup>	N/A
D4Z4 bisulphite PCR primer Seq F-GGGTTGGGTTTATAGT	Ref. <sup>47</sup>	N/A
D4Z4 bisulphite PCR primer BioR-ACAAAACCTAACCTAAAAATATAC	Ref. <sup>47</sup>	N/A
LINE-1 bisulphite PCR primer F- TTTTGAGTTAGGTGTGGGATATA	Ref. <sup>48</sup>	N/A
LINE-1 bisulphite PCR primer Seq F- AGTTAGTGTGGGATATAGT	Ref. <sup>48</sup>	N/A

(Continued on next page)

<b>Continued</b>		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
LINE-1 bisulphite PCR primer BioR- AAAATCAAAAATTCCT	Ref. <sup>48</sup>	N/A
<b>Software and algorithms</b>		
R	<a href="https://www.r-project.org/">https://www.r-project.org/</a>	V4.0.0
Pyro Q-CpG1.0.9 software	Biotage	V0.3.8
Trim Galore	<a href="https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/">https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/</a>	V0.6.5
Samtools	<a href="http://www.htslib.org/">http://www.htslib.org/</a>	V1.11
Bowtie2	<a href="https://bowtie-bio.sourceforge.net/bowtie2/index.shtml">https://bowtie-bio.sourceforge.net/bowtie2/index.shtml</a>	2.3.5.1
Bismark	<a href="https://www.bioinformatics.babraham.ac.uk/projects/bismark/">https://www.bioinformatics.babraham.ac.uk/projects/bismark/</a>	V0.22.3
Scater	Bioconductor	N/A
Scran	Bioconductor	N/A
SC3	Bioconductor	N/A
Monocle	Bioconductor	N/A
Seurat	Bioconductor	N/A
DESeq2	Bioconductor	N/A
Ginkgo		<a href="https://github.com/robertaboukhalil/ginkgo">https://github.com/robertaboukhalil/ginkgo</a>
scMethyl-SEQ and scRNA-SEQ pipelines	This paper; Github	<a href="https://github.com/ClaudiaBN/SingleCellPipelines/">https://github.com/ClaudiaBN/SingleCellPipelines/</a>

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests should be directed to the Lead Contact, David Monk ([d.monk@uea.ac.uk](mailto:d.monk@uea.ac.uk)).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

The scRNA-seq and scBS-seq datasets for oocyte, reference and arrested embryos have been deposited as the NCBI-BioProject repository and are publicly available as the date of publication. Accession numbers are listed in the [key resources table](#). This paper also analyses existing, publicly available data. The accession numbers for these datasets are listed in the [key resources table](#). This paper does not report original code. Source code for the pipelines used can be found in the [key resources table](#). Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Human oocytes and pre-implantation embryos

The use of surplus human oocytes and embryos for this study was evaluated and approved by the scientific and ethic committee of the Instituto Valenciano de Infertilidad (IVI) for two research protocols (1310-FIVI- 131-CS and 1710-VLC-103-MM), Bellvitge Institute of Biomedical Research, Barcelona (PR292/14), the Centro de Medicina Regenerativa de Barcelona (CMRB), the National Committee for Human Reproduction (CNRHA) and the Regional Health Departments for Valencia and Catalunya (4/2014 & 10/2017).

## METHOD DETAILS

### Human oocytes and pre-implantation embryos

All MII oocytes were from women undergoing superovulation for the purpose of oocyte donation who had given informed consent for surplus cryopreserved oocytes to be used for research. All “good quality” embryos donated for this study were obtained following

informed consent from couples who had delivered a healthy baby either from an initial fertility cycle or from a later cycle with thawed frozen embryos, who wished to donate their remaining frozen embryos for scientific research (Table S1). The cleavage-stage embryos and blastocysts characterized in this study are all from frozen embryo cycles and were classified as suitable for transfer. Embryonic assessment was performed according to standard guidelines and the embryonic stages were defined as follows.

- Day 2 embryos (CL2): 2–4 cell-stage embryos collected and cryopreserved ~48 h after routine fertilization. Each blastomeres was of similar size, with limited fragmentation to one cell per embryo.
- Day 3 embryos (CL3): -5-8 cell-stage embryos collected and cryopreserved ~64 h after routine fertilization. 8A1: Embryos had defined nuclei, with limited cytoplasmic fragmentation or degeneration restricted to individual cells of an embryo.
- Day 5 hatching blastocysts had smooth trophectoderm, clearly visible blastocyst cavity and well-defined inner cell mass.

In addition to the “good quality” reference embryos used to define EGA and DNA methylation dynamics described above, cleavage embryos that had arrested at 4-cell ( $n = 4$ ), 5–8 cell ( $n = 15$ ) and 9–16 cell ( $n = 3$ ) stages were collected at day 5 having deemed to have undergone spontaneous cleavage arrest. This second cohort of aberrant embryos resulted from ICSI cycles using fresh oocytes and had accompanying time-lapse imaging which revealed good morphology with minimal fragmentation.

### Embryo culture and time-lapse imaging

MII oocytes and embryos were thawed using the Cryotop Method following manufacturer’s instructions.<sup>49</sup> ICSI was carried out in fertilization medium (Origio, Cooper Surgical®) at  $\times 400$  magnification with the aid of an Olympus 1 $\times$ 7 microscope. Finally, oocytes were placed in pre-equilibrated EmbryoSlides (EmbryoSlide®, Vitrolife) until the blastocyst stage with 28  $\mu$ L (for conventional EmbryoScope) or 180  $\mu$ L (for EmbryoScope Plus) single-step medium (Gems, Genea Biomedx®) and 1.6 mL mineral oil. Embryos were cultured individually (conventional EmbryoScope) or in groups of up to eight (EmbryoScope Plus) until the day 2 and day 3 cleavage embryos or the fifth/sixth day of development. Images of up to 11 multiple focal planes were taken automatically every 10–20 min. Embryo development was assessed on an external computer with analysis software (EmbryoViewer™ workstation, Vitrolife).

After warming, oocytes were cultured in Fertilization Medium, the day 2 and day 3 cleavage embryos were cultured in Gems Medium and maintained with 5% CO<sub>2</sub> at 37°C for at least 2 h prior to processing. The day 5 blastocysts were incubated in Gems medium under the same culture conditions for 6–12 h before processing to allow for full expansion.

Human Metaphase II oocytes were washed in 1% polyvinylpyrrolidone (PVP) (Sigma-Aldrich) and loaded individually into UV treated sterile PCR grade tubes containing 2.5  $\mu$ L of PBS (Cell Signaling Technology), that were immediately snap frozen at –80°C until proceeding with the scM&T-seq protocol.

### Single-cell isolation

Day 2 & 3 embryos were removed from the incubator and a laser incision that removed one-quarter of the zona pellucida was performed with the Hamilton-Thorne Lykos® laser. Blastomeres were spread by blastomere biopsy micropipets (Origio, USA) and then isolated individually by stripper using 120  $\mu$ m tips. Single blastomere was washed in 1% PVP, individually placed in sterile PCR grade tubes containing 2.5  $\mu$ L of PBS and immediately snap frozen at –80°C until downstream processing.

The Inner Cell Mass (ICM) and Trophectoderm (TE) of the blastocysts were separated by micromanipulation using a laser technology (OCTAX). Separated ICMs and TEs were individually incubated in Accutase medium (Chemicon) at room temperature for 10 min until single cells were released by gently pipetting. Resulting single cells were washed in 1% PVP and individually placed in sterile PCR grade tubes containing 2.5  $\mu$ L of PBS and snap frozen at –80°C until processed.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Single-cell sequencing

We utilized the scM&T-seq method described by Angermueller and colleagues.<sup>7</sup> Essentially, each isolated cell was added to RLT plus lysis buffer (Qiagen) and processed using the G&T-seq protocol,<sup>50</sup> resulting in the physical separation of mRNA and genomic DNA from single cells with the DNA fraction eluted into 10  $\mu$ L of H<sub>2</sub>O.

Single-cell bisulphite libraries were then prepared as previously described<sup>51</sup> with minor modifications. Conversion was carried out using EZ Methylation Direct bisulphite reagent (Zymo) on purified DNA in the presence of AMPure XP beads (Beckman Coulter) following G&T-seq. Purification and desulphonation of converted DNA was performed with magnetic beads (Zymo) on a Bravo Workstation (Agilent), eluting into the mastermix for the first strand synthesis. Primers for first and second strand synthesis contained a 3'-random hexamer and biotin capture of first strand products was omitted, however, an extra 0.8 $\times$  AMPure XP purification was performed between second strand synthesis and PCR. Each pre-PCR AMPure XP purification was carried out using a Bravo Workstation. To minimise batch effects all libraries were prepared in parallel in a 96 well plate. Purified scBS-seq libraries were sequenced in pools of 16–20 per lane of an Illumina HiSeq2500 using 101-bp paired-end reads.

RNA sequencing libraries were prepared from the single-cell cDNA libraries using the Nextera XT kit (Illumina) as per the manufacturer’s instructions but using one-fifth volumes. Multiplexed library pools were sequenced on one lane of an Illumina HiSeq2500 generating 50-bp single-end reads.

### Sequence data processing and raw data analysis

BS-seq read alignment: sequencing data was processed as previously described<sup>25,52</sup> with minor modifications. Briefly, raw sequence reads were trimmed to remove the 6N random priming sequences of the reads, with adapter contamination and poor-quality base called using Trim Galore (v0.3.8). Subsequently, trimmed reads were aligned to the human reference genome (build hg38/19) using Bismark (v0.16.3) in single-end non-directional mode. Methylation calls were extracted after duplicate alignments had been removed. Predicted methylation states were binarized with methylation lower than 0.1 set to 0 (unmethylated cytosine), methylation between 0.1 and 0.9 set to 0.5 (equivalent to allelic methylation) and all methylation sites greater than 0.9 set to 1 (methylated cytosine).

RNA-seq read alignment: reads were trimmed using Trim Galore with a minimal “phred” score of 13 and a minimum read length of 35 bp. The resulting sequence reads were mapped by Hisat2 against the hg38 genome build. Cells with <0.75 million reads or more than 15 million reads were eliminated from downstream analysis as they represented inefficiently processed cells or possible contaminated samples. Subsequent expression analyses were carried out by combining Scater,<sup>53</sup> Scrn,<sup>54</sup> SC3,<sup>55</sup> Monocle<sup>56</sup> and Seurat<sup>57</sup> R-packages using different in-house scripts.

### Sequence features

The coordinates of genomic features, including CpG islands and genes bodies, used for methylation profiling were obtained from the UCSC genome browser, while all repetitive element information for SINEs, LINEs and LTRs were downloaded from RepeatMasker. Promoters, defined as 1kb upstream and 0.5kb downstream of the TSS were classified as high, intermediate and low CpG content as previously described.<sup>58</sup> The location of solo-WCGW highly and partially methylated domains was from Zhou et al.<sup>27</sup> Enhancer regions were identified as active (class I; H3K4me1+H3K27ac) and poised (class II; H3K4me1+H3K27me3) enhancers in human stem cells.<sup>59</sup> The DNA methylation level of these genomic regions was calculated on the basis of the average methylation of all covered CpG sites within these regions. For the characterisation of methylation of gene promoters at different developmental stages, we compared our scBS-seq with published bulk datasets from brain, blood, preimplantation embryos and sperm. The accession numbers for these datasets are listed in the [key resources table](#).

### Data analysis

Gene expression levels were normalised in terms of transcripts per million (TPM) of mapped reads to the transcriptome of each cell. tSNE analysis was performed on the normalized counts. Gene pathway annotations were obtained from G.O.<sup>60</sup> Differentially expressed transcripts were identified using DESeq2, whilst global variance in transcription was called using kernel density estimations as previously described.<sup>18</sup> Correlation analyses were performed excluding non-expressed genes and using Spearman’s correlation as a measurement between variables.

For sex determination, embryos were classified according to the presence of Y chromosome expression and/or mappable bisulphite reads to sex chromosomes. Phasing cells according to cell cycle stages was performed using “Cyclone” function in the Scrn R-package.<sup>61</sup> Aneuploidy estimations for each cell was performed using BS-seq datasets in Ginkgo.<sup>62</sup>

### Locus-specific methylation analysis

*D4Z4* methylation was assessed by PCR<sup>47</sup> on either bisulphite converted DNA isolated from samples and individual blastocysts using EZ DNA Methylation-Direct kit (ZYMO) or excess PBAT libraries using Immolase Taq polymerase (Bioline) for 45 cycles and the resulting PCR product ligated into pGEM-T easy vector (Promega) and sub-clones sequenced with T7 primer.

For LINE-1 pyrosequencing, standard bisulphite PCR was performed except that the reverse primer was biotinylated.<sup>48</sup> The entire biotinylated PCR product (diluted to 40  $\mu$ L) was mixed with 38  $\mu$ L of Binding buffer and 2  $\mu$ L (10 mg/mL) streptavidin-coated polystyrene beads. After incubation at 65°C, DNA was denatured with 50  $\mu$ L 0.5M NaOH. The single-stranded DNA was hybridized to 40-pmol sequencing primers dissolved in 11  $\mu$ L annealing buffer at 90°C. For sequencing, a primer was designed to the opposite strand to the biotinylated primer was used. The pyrosequencing reaction was carried out on a PyroMark Q96 instrument. The peak heights were determined using Pyro Q-CpG1.0.9 software (Biotage).