

UC Davis

UC Davis Previously Published Works

Title

Histone remodeling reflects conserved mechanisms of bovine and human preimplantation development.

Permalink

<https://escholarship.org/uc/item/9t4824t8>

Journal

EMBO Reports, 24(3)

Authors

Zhou, Chuan

Halstead, Michelle

Bonnet-Garnier, Amélie

et al.

Publication Date


2023-03-06

DOI

10.15252/embr.202255726

Peer reviewed

Histone remodeling reflects conserved mechanisms of bovine and human preimplantation development

Chuan Zhou^{1,†} , Michelle M Halstead^{2,3,†}, Amélie Bonnet-Garnier^{2,3} , Richard M Schultz^{4,5}  & Pablo J Ross^{1,*} 

Abstract

How histone modifications regulate changes in gene expression during preimplantation development in any species remains poorly understood. Using CUT&Tag to overcome limiting amounts of biological material, we profiled two activating (H3K4me3 and H3K27ac) and two repressive (H3K9me3 and H3K27me3) marks in bovine oocytes, 2-, 4-, and 8-cell embryos, morula, blastocysts, inner cell mass, and trophectoderm. In oocytes, broad bivalent domains mark developmental genes, and prior to embryonic genome activation (EGA), H3K9me3 and H3K27me3 co-occupy gene bodies, suggesting a global mechanism for transcription repression. During EGA, chromatin accessibility is established before canonical H3K4me3 and H3K27ac signatures. Embryonic transcription is required for this remodeling, indicating that maternally provided products alone are insufficient for reprogramming. Last, H3K27me3 plays a major role in restriction of cellular potency, as blastocyst lineages are defined by differential polycomb repression and transcription factor activity. Notably, inferred regulators of EGA and blastocyst formation strongly resemble those described in humans, as opposed to mice. These similarities suggest that cattle are a better model than rodents to investigate the molecular basis of human preimplantation development.

Keywords cattle; epigenetics; genome activation; preimplantation embryo

Subject Categories Chromatin, Transcription, & Genomics; Development

DOI 10.15252/embr.202255726 | Received 5 July 2022 | Revised 5 January 2023 |

Accepted 9 January 2023 | Published online 13 February 2023

EMBO Reports (2023) 24: e55726

Introduction

Preimplantation development encompasses two fundamentally important, but poorly understood, cell identity transitions: the acquisition and subsequent loss of totipotency. Key to these transitions is the maternal-to-embryonic transition (MET), marked by embryonic genome activation (EGA) and clearance of oocyte-specific messages (Schultz *et al.*, 2018). As blastomeres continue to divide, they progressively transition from totipotency to pluripotency, and finally segregate into two cell types that comprise the blastocyst: the inner cell mass (ICM) and trophectoderm (TE). MET and blastocyst formation correspond to changes in gene expression and widespread epigenetic remodeling of DNA methylation, histone modifications, and chromatin structure (Frum & Ralston, 2015). Unclear, however, is the causal relationship between transcriptomic changes and chromatin remodeling.

Modification of histone tails can modulate gene expression by controlling accessibility of chromatin to transcription factors (TFs) and RNA polymerase (Eckersley-Maslin *et al.*, 2018; Jambhekar *et al.*, 2019; Schulz & Harrison, 2019). Of many possible histone modifications, a select few are commonly profiled to characterize active and repressed genomic regions. These modifications include trimethylation of histone 3 lysine 4 (H3K4me3), which marks active promoters, acetylation of histone 3 lysine 27 (H3K27ac), which marks both active promoters and enhancers, trimethylation of histone 3 lysine 27 (H3K27me3), a mark of polycomb repression and facultative heterochromatin, and trimethylation of histone 3 lysine 9 (H3K9me3), associated with constitutive heterochromatin. Cross-talk between these modifications and their distribution throughout the genome collectively allows for fine-tuning of spatial and temporal gene expression. Unsurprisingly, disruption of this dynamic has catastrophic effects on transcription and development (Jambhekar *et al.*, 2019).

1 Department of Animal Science University of California, Davis, CA, USA

2 Université Paris-Saclay, UVSQ, INRAE, BREED, Jouy-en-Josas, France

3 Ecole Nationale Vétérinaire d'Alfort, BREED, Maisons-Alfort, France

4 Department of Anatomy, Physiology and Cell Biology, School of Veterinary Medicine, University of California, Davis, CA, USA

5 Department of Biology, University of Pennsylvania, Philadelphia, PA, USA

Corresponding author. Tel: +1 530 771 7225; E-mail: ross@ucdavis.edu

Preprint Servers: This article has been deposited as a preprint on *bioRxiv* ([10.1101/2022.04.07.486777](https://doi.org/10.1101/2022.04.07.486777)), with permission to display the preprint in perpetuity, under a CC-BY-ND 4.0 International license.

[†]These authors contributed equally to this work

Although several recent reports have shed light on epigenetic remodeling during preimplantation development, they either do not capture the entire continuum of stages or focus on discrete levels of regulation, e.g., chromatin accessibility (Lu *et al*, 2016; Wu *et al*, 2016, 2018; Gao *et al*, 2018; Li *et al*, 2018; Liu *et al*, 2019; Halstead *et al*, 2020a; Ming *et al*, 2021) or specific histone modifications (Dahl *et al*, 2016; Liu *et al*, 2016; Zhang *et al*, 2016, 2019; Zheng *et al*, 2016; Wang *et al*, 2018; Xia *et al*, 2019; Lu *et al*, 2021). Efforts to integrate such datasets *post hoc* can be problematic because technical and methodological differences, such as missing biological timepoints and differing sample composition, can confound downstream analyses. Thus, the collective dynamics of histone modifications, chromatin structure, and DNA methylation during preimplantation development and how these changes relate to changes in gene expression—and therefore cell identity and developmental potential—remain largely unknown, especially in mammals other than mouse. This gap in knowledge is particularly conspicuous because accumulating evidence suggests that larger mammals, e.g., cattle, constitute a better model system than rodents to understand human preimplantation development (Halstead *et al*, 2020a; Lu *et al*, 2021).

We report here a comprehensive analysis of both active (H3K4me3, H3K27ac) and repressive histone modifications (H3K9me3, H3K27me3) that bridges key developmental transitions from fertilization to blastocyst formation in bovine embryos. In addition, we integrate these results with previously published DNA methylation, chromatin accessibility, and gene expression datasets. We find a dramatic shift in the overall structure of the epigenome between the 8-cell and morula stages. Prior to EGA, embryos are distinguished by broad domains of all profiled histone marks, as well as enrichment for repressive marks in gene bodies. Post-EGA, this naïve chromatin organization resolves into a more typical structure, with chromatin accessibility appearing earlier than canonical H3K4me3 and H3K27ac signatures. Inhibiting embryonic transcription severely impacts H3K27ac remodeling, leading to the retention of a 4-cell-like epigenetic state in 8-cell embryos. Finally, the two cell lineages in blastocysts—ICM and TE—are defined by accumulation of differential polycomb repression and distinct regulatory circuitry, which is notably similar to that observed in human blastocysts, further supporting use of cattle as a relevant model for human preimplantation development.

Results

Histone modification profiles during bovine preimplantation development

Using low-input CUT&Tag (cleavage under targets & tagmentation; Kaya-Okur *et al*, 2019), H3K4me3, H3K27ac, H3K9me3, and H3K27me3 were profiled in bovine germinal-vesicle (GV) oocytes, 2-, 4-, and 8-cell embryos (that developed in the presence or absence of the transcription inhibitor α -amanitin), morula, blastocysts, ICM, and TE ($n = 2$ biological replicates). In total, 72 CUT&Tag libraries were generated and sequenced, producing 6,858,594,136 raw reads, from which 1,387,865,654 informative alignments were obtained and used for downstream analyses (Appendix Tables S1–S4). To complement data previously published

by our group (Halstead *et al*, 2020a; Data ref: Halstead & Ross, 2020), chromatin accessibility was profiled in blastocysts, ICM, and TE ($n = 3$ biological replicates) using ATAC-seq (assay for transposase-accessible chromatin by sequencing; Buenrostro *et al*, 2013; Appendix Fig S1A and B, and Table S5). Genome-wide profiles of CUT&Tag biological replicates were highly correlated (average Pearson's $r = 0.96$; Appendix Fig S1C). For each histone mark, samples clustered into two general groupings: pre-EGA stages (2-, 4-, and 8-cell embryos) and post-EGA stages (morula, blastocysts, ICM, and TE; Fig 1A). For each stage and mark, regions with significant enrichment in both biological replicates, *i.e.*, peaks, were identified. Generally, when calling peaks, H3K4me3 and H3K27ac were treated as “narrow” marks and H3K27me3 and H3K9me3 as “broad” marks. However, considering recent findings that broad domains of noncanonical H3K4me3 (ncH3K4me3) are present in bovine oocytes (Lu *et al*, 2021), for peak calling purposes, H3K4me3 was treated as a broad mark up until the 8-cell stage (see [Materials and Methods](#)). Similar to H3K4me3, H3K27ac was also broadly distributed during early stages and was treated as a broad mark until the 8-cell stage. Because we previously showed that IgG controls for CUT&Tag produced negligible reads after sequencing (Navarro *et al*, 2022), IgG libraries were not included in the current study. To confirm that CUT&Tag did not produce a nonspecific signal, H3K27me3 and H3K27ac peaks were compared across development to show that these mutually exclusive marks did not colocalize (Appendix Fig S1D).

Global enrichment of a given histone modification, measured as percent of the genome covered by peaks, changed dramatically throughout the course of preimplantation development (Fig 1B). In oocytes and pre-EGA embryos, H3K4me3 was widespread, covering 11% of the genome in GV oocytes (Fig 1B). These broad domains of H3K4me3 were clearly evident up until the 8-cell stage, at which point they resolved into narrow peaks at promoters (Fig 1C). H3K27ac also occupied broad domains in pre-EGA embryos, but the H3K27ac signal was generally weaker than that observed for H3K4me3 (Fig 1C). Indeed, genomic coverage of H3K27ac decreased substantially after fertilization, with partial re-establishment in 8-cell embryos (Fig 1B), although canonical narrow peaks of H3K27ac were not abundantly evident until the morula stage (Fig 1C). Compared to these activating marks, H3K9me3 abundance was relatively stable throughout preimplantation development (Fig 1B). However, localization of this mark underwent a remarkable shift after fertilization, with H3K9me3 in oocytes and pre-EGA embryos occurring primarily at gene bodies (Fig 1C and D) and then switching to an intergenic distribution between the 8-cell and morula stages (Fig 1B).

Similar to H3K9me3, H3K27me3 coverage was relatively constant in oocytes and pre-EGA embryos (Fig 1B). However, unlike H3K9me3, the localization of this mark changed several times, from primarily intergenic in oocytes to intragenic in pre-EGA embryos, and then back to intergenic after EGA (Fig 1C and D). This profile contradicts immunofluorescence studies, which found that H3K27me3 is progressively and globally erased after fertilization, and almost absent in 8-cell embryos (Ross *et al*, 2008; Canovas *et al*, 2012; Zhou *et al*, 2019). In contrast to our findings, a recent study using CUT&RUN (cleavage under targets and release using nuclease; preprint: Skene & Henikoff, 2017) indicated global loss of H3K27me3 following fertilization from the 4- to 16-cell stage in

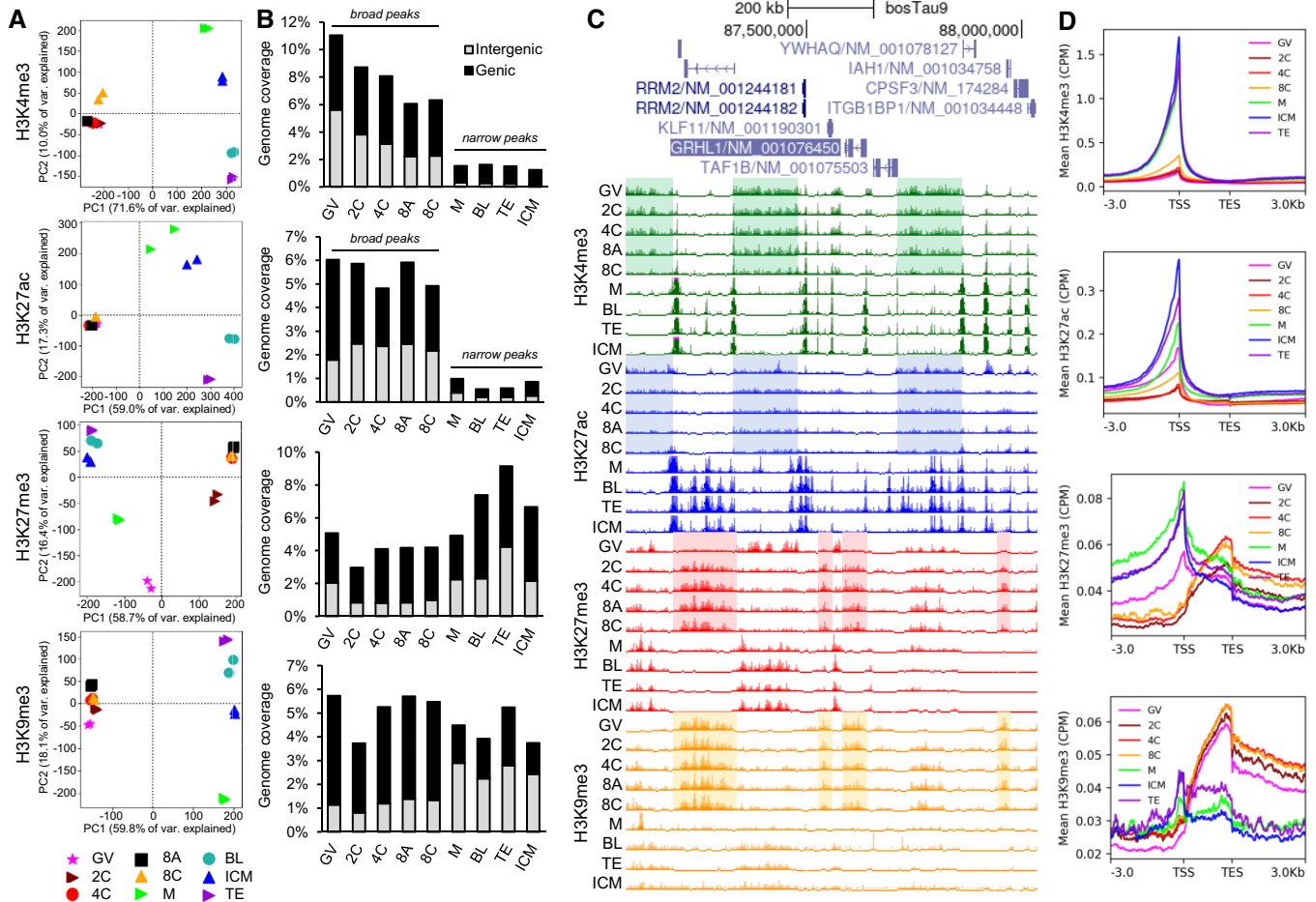


Figure 1. Profiles of histone modifications in bovine oocytes and preimplantation embryos.

A Principal components analysis (PCA) of CUT&Tag libraries.

B Genome coverage of peaks identified in both biological replicates. Peaks classified as genic (overlapping 2 Kb promoter or gene bodies) or intergenic. For H3K4me3 and H3K27ac, genome coverage calculated for broad peaks until the 8-cell stage, then for narrow peaks.

C Normalized signal (counts per million; CPM) of one biological replicate per developmental stage and histone mark. Shaded regions correspond to noncanonical broad distributions (green, blue) and intragenic enrichment (red, orange). Viewing range from 0 to 1.5 CPM. Values exceeding maximum range indicated by pink bars.

D Average normalized signal (CPM) for each histone mark at transcription start sites (TSS) and end sites (TES). Transcripts scaled to 2 Kb, 3 Kb upstream and downstream regions shown.

Data information: GV oocytes (GV), 2-cell (2C), 4-cell (4C), 8-cell embryos (8C), α -amanitin-treated 8-cell embryos (8A), morula (M), blastocyst (BL), trophectoderm (TE), and inner cell mass (ICM).

bovine embryos (Lu *et al.*, 2021). To determine whether these differences were due to technique, H3K4me3 and H3K27me3 CUT&Tag profiles were compared to those obtained through CUT&RUN (Lu *et al.*, 2021). In general, the profiles were highly comparable, with the exception of H3K27me3 at the 4- and 8-cell stages (Appendix Fig S2). For these stages, the CUT&RUN H3K27me3 data suffered from low sequencing depth, resulting in low replicability between samples (Appendix Table S6). Notably, the CUT&RUN data for these stages were generated from 10-fold fewer cells than the current study, suggesting that for certain histone modifications (e.g., H3K27me3) at certain stages, it may be advisable to use more cells as input to maximize library complexity and capture the true biological signal.

Overall, our data indicate that the maternal profiles of H3K4me3 and H3K9me3 were highly similar to that observed in 2-, 4-, and 8-cell embryos, suggesting substantial inheritance from the oocyte, both at intergenic regions marked by ncH3K4me3 and at gene bodies marked by H3K9me3. On the other hand, H3K27ac underwent substantial depletion after fertilization, and although the oocyte-specific H3K27me3 is largely erased in 2-cell embryos, *de novo* establishment of this mark was evident at gene bodies.

Broad bivalent domains in oocytes mark developmental genes

The results described above raise the question about the relationship between remodeling of histone modifications and the

repression and activation of gene expression programs that underlie changes in cell identity. Chromatin structure of pre-EGA embryos is markedly immature, characterized by high histone mobility (Ooga *et al*, 2016), dispersed chromatin (Ahmed *et al*, 2010), and lack of canonical 3-D chromatin architecture (Du *et al*, 2017). With developmental time, the chromatin adopts a more mature, condensed structure, at which point enhancer activity once again becomes necessary to relieve repression and activate gene expression (Wiekowski *et al*, 1993). Consequently, the genomic distribution of histone modifications during preimplantation development, especially prior to EGA, does not necessarily follow canonical rules and may not serve typical functions. For instance, H3K4me3, which generally occurs as sharp peaks at promoters in differentiated cells, occupies broad intergenic regions (> 5 Kb) in most mammalian oocytes and pre-EGA embryos (Dahl *et al*, 2016; Lu *et al*, 2021). To account for this finding, H3K4me3 was treated as a “broad” mark for peak calling purposes, which improved reproducibility between biological replicates in GV oocytes, 2-, 4-, and 8-cell embryos (Fig EV1A). Treating H3K27ac as a broad mark also improved reproducibility at these stages (Fig EV1A), indicating that this mark also has a broad noncanonical distribution prior to EGA.

Noncanonical H3K4me3 is anticorrelated with DNA methylation and occurs almost exclusively in partially methylated domains (PMDs; Dahl *et al*, 2016; Zhang *et al*, 2016; Zheng *et al*, 2016; Lu *et al*, 2021). To determine whether this pattern was recapitulated in bovine oocytes and embryos, and how PMDs relate to other histone modifications, we first identified PMDs in GV oocytes from published DNA methylation data (Data ref: Ivanova *et al*, 2020b; Ivanova *et al*, 2020a) using previously described methods (Lu *et al*, 2021; see Materials and Methods). We found that PMDs covered 25.3% of the genome, consistent with previous reports (Lu *et al*, 2021). Noncanonical H3K4me3 and H3K27ac were apparent within PMDs (Fig EV1B and C). In GV oocytes, H3K4me3 was found almost exclusively in PMDs (90% of broad H3K4me3 peaks overlapped PMDs), although only 34% of PMDs were marked by H3K4me3 (Fig EV1D and E). Enrichment of H3K4me3 in PMDs was maintained until the 8-cell stage. H3K27ac demonstrated a similar pattern. In GV oocytes, about 80% of broad H3K27ac domains occurred within PMDs, and this enrichment was maintained until the 4-cell stage (Fig EV1E). H3K27me3 was also enriched at PMDs in GV oocytes, with more than 80% of H3K27me3 peaks overlapping PMDs. However, this enrichment was rapidly lost after fertilization, and H3K27me3 was not re-established at these regions until after EGA (Fig EV1E). In contrast to the other three marks, H3K9me3 was notably depleted at PMDs in GV oocytes (only 10% of peaks overlapped with PMDs) and was gradually established in PMDs as development progressed (Fig EV1C–E). This lack of H3K9me3 enrichment at PMDs was anticipated given global demethylation of the embryonic genome after fertilization (Greenberg & Bourc’his, 2019; Ivanova *et al*, 2020a) and that H3K9me3 deposition can depend on DNA methylation, and vice versa (Rose & Klose, 2014).

H3K4me3, H3K27me3, and H3K27ac were all enriched at PMDs. Because H3K27ac and H3K27me3 are mutually exclusive modifications, this enrichment suggested that different PMDs are probably marked by different combinations of histone modifications, which could entail distinct functions. To identify these patterns, PMDs were grouped according to histone modifications and chromatin

accessibility signal in GV oocytes. The analysis revealed four clusters of PMDs (Fig 2A), each exhibiting unique developmental changes (Appendix Fig S3A).

PMDs belonging to “cluster_1”, hereafter referred to as bivalent PMDs, were co-marked by H3K4me3 and H3K27me3 in GV oocytes (Fig 2A and B), and covered 2% of the genome. Although H3K4me3 was maintained at bivalent PMDs up to the 8-cell stage, H3K27me3 was rapidly erased following fertilization before re-establishment in morula, and accessibility remained consistently low (Fig 2C). We found that bivalent PMDs in bovine GV oocytes mark genes related to transcriptional activation and development, and especially conserved homeobox genes (Appendix Table S7). Moreover, among all PMD clusters, genes marked by bivalent PMDs were the least expressed throughout preimplantation development (Appendix Fig S3B), suggesting a role in poisoning these genes for future expression, similar to the role of bivalency in ESC.

Both “cluster_2” and “cluster_3” demonstrated enrichment for activating marks at early stages, but “cluster_2” PMDs, hereafter referred to as highly active PMDs, were strongly and briefly enriched for H3K27ac, demonstrated higher accessibility in GV oocytes, and marked genes related to transcription silencing (Fig 2C and Appendix Table S7). These strongly active PMDs might therefore play a role in maintaining transcriptional quiescence in full-grown oocytes (Schultz *et al*, 2018). On the other hand, moderately active PMDs (“cluster_3”) were not as enriched for activating marks as strongly active PMDs, and rather than occurring near genes related to transcriptional repression, instead occurred near genes related to transcription activation (Fig 2C and Appendix Table S7).

Finally, the most abundant category of PMDs (“cluster_4,” termed quiescent PMDs), demonstrated low enrichment for most marks, with slightly higher levels of H3K27me3 in GV oocytes (Fig 2A and C). Genes marked by these regions demonstrated little functional enrichment. As such, these PMDs likely represent heterochromatin compartments in differentiated cells, especially given their stable inaccessibility and accumulation of H3K9me3 during later stages of development (Fig 2C), which was also maintained in fetal fibroblasts (Appendix Fig S3C). In summary, these broad domains, devoid of DNA methylation and precisely marked by specific combinations of histone modifications, likely function to both repress transcription and poise developmental genes for future expression.

H3K9me3 and H3K27me3 co-occupy gene bodies in pre-EGA embryos

Given the extensive changes to chromatin structure at intergenic regions, we next focused on the genes themselves. H3K4me3 and H3K27ac signal at promoters increased progressively after fertilization, whereas H3K27me3 and H3K9me3 underwent a remarkable shift in distribution from gene bodies (“intragenic”) in 2-, 4-, and 8-cell embryos, to promoters from the morula stage onward (Fig 1D). Notably, intragenic H3K9me3 was already evident in GV oocytes, but intragenic H3K27me3 was not, indicating that maternal H3K9me3 is inherited, whereas intragenic H3K27me3 is deposited *de novo* following fertilization. To determine if this pattern was characteristic of all genes or only a subset, transcript isoforms were clustered based on their H3K27me3 and H3K9me3 signatures (Appendix Fig S4A).

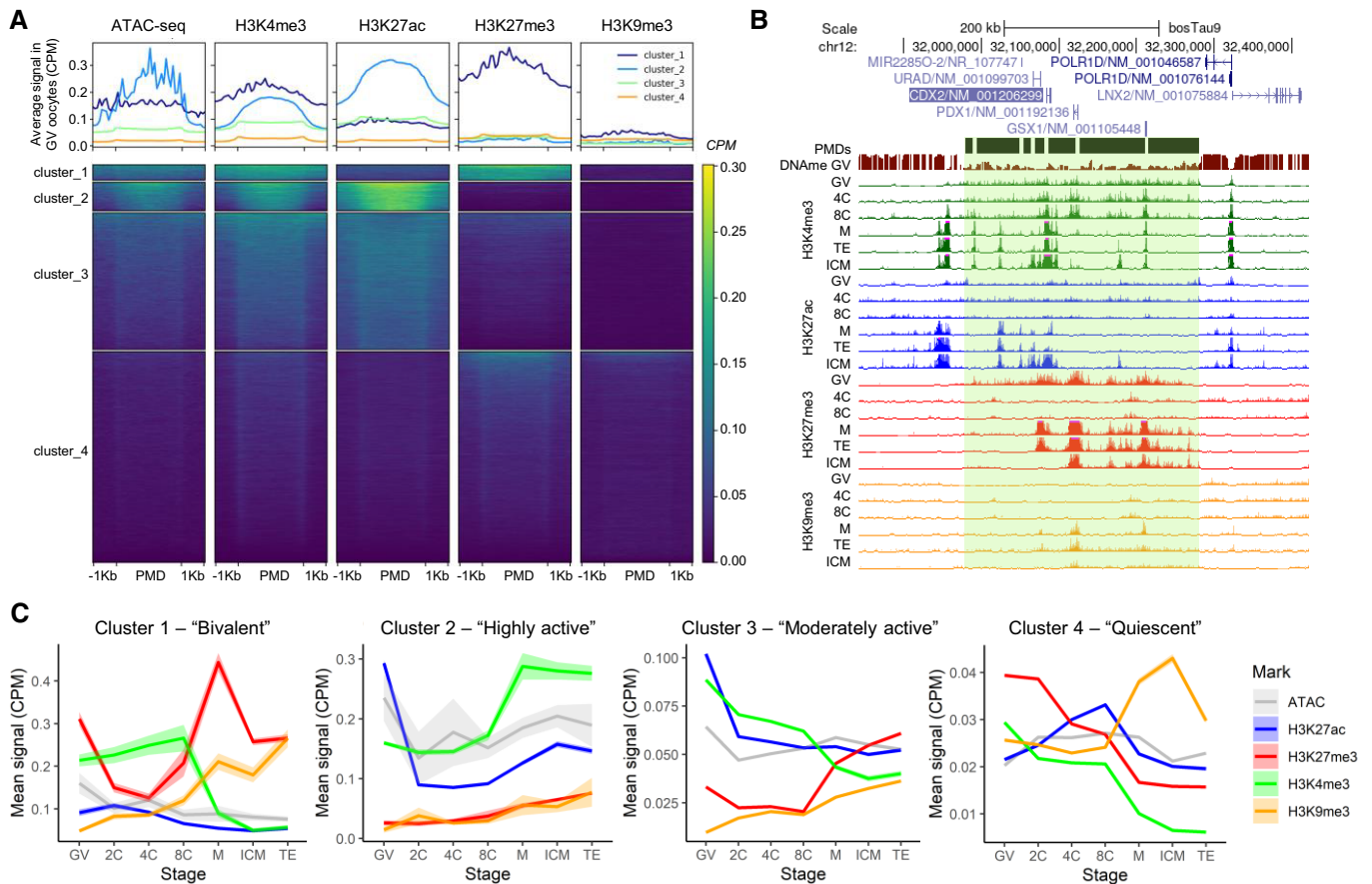


Figure 2. Characterization of partially methylated domains (PMDs) in GV oocytes.

- A Clustering of PMDs based on normalized signal of H3K4me3, H3K27ac, H3K27me3, H3K9me3, and ATAC-seq in GV oocytes. Signal normalized by CPM in 100 bp windows, with PMDs scaled to 3 Kb, and showing regions 1 Kb up- and downstream.
- B Representative gene track image of a bivalent PMD (cluster_1), overlapping three homeobox genes: *CDX2*, *GSX1*, and *PDX1*. One biological replicate shown per stage and mark. Viewing range from 0 to 1.5 CPM. Values exceeding maximum range indicated by pink bars.
- C Average epigenetic signal (CPM \pm standard error) for PMDs belonging to each cluster across development.

A small subset of transcripts (“cluster_1”; $n = 1,639$) was depleted for both repressive marks in pre-EGA embryos, but strongly marked from the morula stage onward; these transcripts also exhibited a strong H3K27me3 signal in GV oocytes, which was promptly erased after fertilization (Fig 3A and B, and Appendix Fig S5A). Moreover, they were broadly marked by H3K4me3 and H3K27ac, rather than demonstrating sharp peaks at their transcription start sites (TSS, Fig 3A and Appendix Fig S4B), and were essentially not expressed throughout development (Appendix Fig S4C). Therefore, “cluster_1” transcripts are hereafter referred to as repressed loci. Repressed loci demonstrated similar patterns of histone remodeling to bivalent PMDs and were also enriched for homeobox genes (Appendix Fig S5B). Notably, not all homeobox genes belonged to this cluster. For instance, several homeobox genes implicated in early development and pluripotency (e.g., *NANOG*, *POU5F1*, *DUX4*) corresponded instead to “cluster_3”, suggesting that this pattern of strong repression is limited to genes required later in development.

The majority of transcripts belonged to either “cluster_2” ($n = 6,469$) or “cluster_3” ($n = 35,754$), both of which

demonstrated intragenic H3K9me3 and H3K27me3 in pre-EGA embryos, although only intragenic H3K9me3 was inherited from the oocyte (Fig 3A and C, and Appendix Fig S4A). In contrast to repressed loci, the TSS of these transcripts were markedly enriched for H3K4me3, H3K27ac, and accessibility, albeit with varying intensity depending on the developmental stage (Fig 3A and Appendix Fig S4B). Consequently “cluster_2” and “cluster_3” transcripts are hereafter referred to as “poised-1” and “poised-2” loci, respectively. “Poised-1” loci exhibited more intense intragenic H3K9me3 and H3K27me3 and were functionally enriched for ATP-binding and kinase activity (Appendix Fig S5C and D), whereas “poised-2” loci demonstrated subtler intragenic H3K9me3 and H3K27me3 and were enriched for functions related to transcript processing, such as poly(A) binding, ribosomes, and spliceosomes (Appendix Fig S5E and F).

These results suggest that genic H3K9me3 and H3K27me3 represses expression of oocyte-specific programs. Based on maternal and embryonic genes identified from RNA-seq datasets (Graf *et al*, 2014b; Data ref: Graf *et al*, 2014a; Data ref: Bogliotti *et al*, 2018; Bogliotti *et al*, 2020; Appendix Fig S6A–C), genes

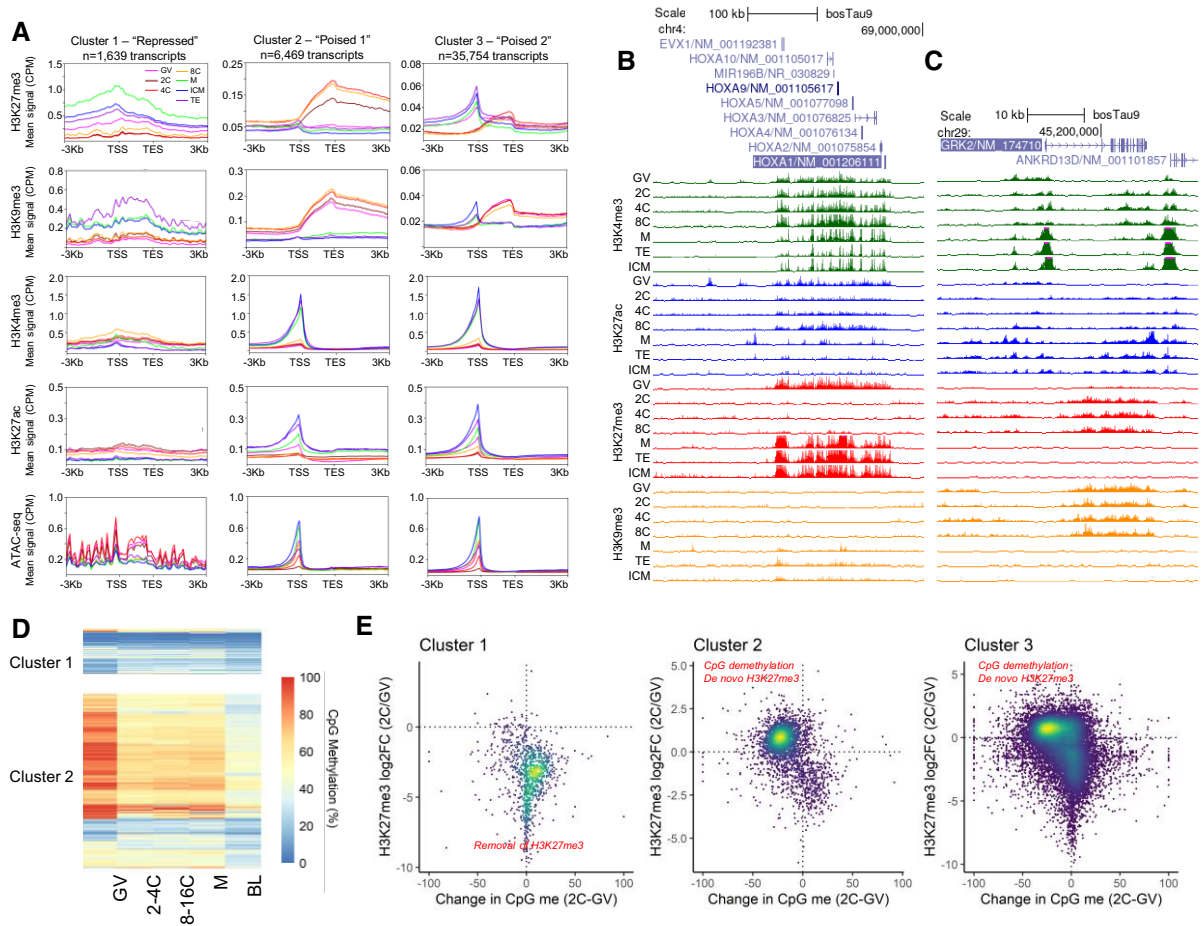


Figure 3. H3K9me3 and H3K27me3 mark gene bodies in pre-EGA embryos.

- A** Average normalized signal (CPM) of histone marks and chromatin accessibility for each transcript cluster, defined by *k*-means clustering (*k* = 3) based on normalized H3K27me3 and H3K9me3 signal (CPM) across development. Loci scaled to 2 Kb, regions 3 Kb upstream and downstream shown.
- B** An example of “cluster_1” loci: the HOX cluster.
- C** An example of “cluster_2” loci: *GRK2* (a kinase). One biological replicate shown per stage and mark. Viewing range from 0 to 1.5 CPM. Values exceeding maximum range indicated by pink bars.
- D** Gene body CpG methylation (%) of “cluster_1” and “cluster_2” genes throughout development.
- E** For each cluster, comparison of change in gene body CpG methylation (GV oocytes versus 2- to 4-cell embryos) to change in intragenic H3K27me3 signal (GV oocytes versus 2-cell embryos). Trends summarized in red text.

upregulated during minor EGA were enriched in “cluster_2” loci, which demonstrated the strongest genic repression, and maternal genes were notably depleted in “cluster_1” loci, which lacked genic H3K9me3 and H3K27me3 (Appendix Fig S6D and E). Maternal genes were also depleted at bivalent PMDs (Appendix Fig S6F), which, similarly to “cluster_1” loci, were enriched for homeobox genes.

In GV oocytes, most genes (poised loci) were marked by high gene body H3K9me3, but still retained some H3K4me3 at their promoters (Fig 3A). This situation is highly reminiscent of the poised state of lineage-specific genes in mesenchymal stem cells and preadipocytes, in which promoters are marked by H3K4me3 and gene bodies are marked by DNA methylation, which recruits SETDB1, leading to deposition of H3K9me3 downstream of TSS (Matsumura *et al*, 2015). Intragenic H3K9me3 then prevents histone acetylation and spreading of H3K4me3, resulting in pausing of RNA

polymerase and very low expression (Matsumura *et al*, 2015). Consistent with the poised state observed in these cells, genes with the highest levels of intragenic H3K9me3 in GV oocytes (“poised-1” loci) demonstrated significantly higher gene body methylation than those lacking intragenic H3K9me3 in GV oocytes (“repressed” loci; unpaired Wilcoxon signed-rank test; $P < 2.2 \times 10^{-16}$; Fig 3D and Appendix Fig S4D). Thus, in conjunction with DNA methylation, intragenic H3K9me3 may help prevent aberrant gene transcription in the transcriptionally quiescent oocyte. After fertilization, however, the embryonic genome undergoes global DNA demethylation (Greenberg & Bourc’his, 2019; Ivanova *et al*, 2020a), including gene bodies. For both “poised-1” and “poised-2” loci, methylation levels dropped after fertilization and declined further upon blastocyst formation (Appendix Fig S4D). This loss of gene body methylation was concomitant with *de novo* deposition of intragenic H3K27me3 (Fig 3E). The exchange of gene body methylation for

H3K27me3 is not surprising, given that DNA methylation antagonizes deposition of H3K27me3 (Hagarman *et al*, 2013).

Overall, the presence of intragenic H3K9me3 in oocytes suggests a mechanism for general transcriptional repression until major EGA. As *de novo* deposition of intragenic H3K27me3 coincides with the loss of gene body methylation, this mark may serve a compensatory repressive role during global DNA demethylation, thus preventing premature transcription of the embryonic genome before the epigenome has been sufficiently reprogrammed to permit appropriate expression of developmental programs.

Chromatin accessibility precedes establishment of canonical H3K4me3 and H3K27ac

Whereas all profiled histone marks underwent a dramatic switch in distribution during the 8-cell to morula transition (Fig 1A–C), the ATAC-seq signal displayed a more gradual change, with a weaker separation of pre- and post-EGA stages (Appendix Fig S1B). In fact, the most notable shift in chromatin accessibility occurred between the 4- and 8-cell stage (Appendix Fig S1A), when more than 68,000 ATAC-seq peaks were gained. This timing difference could reflect a priming mechanism, wherein a region becomes accessible before accumulating activating histone marks. Indeed, H3K27ac peaks established at the morula stage were already accessible in 8-cell embryos (Fig 4A). The “accessibility first” pattern was even more dramatic for transcription start sites (TSS), which often gained accessibility as early as the 4-cell stage, but did not exhibit canonical narrow H3K4me3 peaks until the morula stage (Fig 4B and C).

It is possible that pioneer TFs contribute to this accessibility, given their ability to bind to inaccessible chromatin and subsequently recruit remodeling complexes, histone modifiers, and other TFs (Zaret, 2020). To address this hypothesis, motif enrichment analyses were conducted on intergenic ATAC-seq and H3K27ac peaks to determine whether the motifs of certain pioneer factors were overrepresented in active regions prior to EGA. Although there was little enrichment for pioneer factor motifs in open chromatin at the 4-cell stage, H3K27ac peaks were already significantly enriched for DUX4 motifs (Fig 4D), supporting the hypothesis that DUX4 can bind initially inaccessible regions and induce H3K27ac by recruiting histone acetyltransferases (Choi *et al*, 2016). By the 8-cell stage, DUX4 motifs were also enriched in open chromatin, along with the motifs of several other pioneer factors, including NFY, PAX7, KLF4, and GATA factors (Fig 4D). Of note, whereas DUX4 motifs first appeared in H3K27ac peaks and then in accessible regions, the opposite was true for NFY, PAX7, KLF4, and GATA factor motifs, which first appeared in open chromatin and then in H3K27ac peaks. The latter pattern—“accessibility first”—is consistent with the pre-establishment of accessibility at sites of future H3K27ac peaks (Fig 4A), suggesting that pioneer factors contribute to the establishment of chromatin accessibility at the 8-cell stage prior to deposition of H3K27ac at these sites in morula.

Transcription is required for histone remodeling but not maintenance

As described above, temporal changes in histone modification distribution during preimplantation development occur on both a large scale (e.g., at PMDs) and at specific loci (e.g., genes and enhancers).

The identity of factors responsible for these changes, and whether they are maternally derived and/or products of embryonic transcription, remains unknown. To assess the contribution of embryonic transcription to epigenetic reprogramming, histone modifications were profiled in 8-cell embryos cultured from the 1-cell stage in the presence of α -amanitin. We focused on regions that gained, lost, or maintained enrichment (e.g., peaks) for a given histone modification during the 4- to 8-cell transition, and whether these regions corresponded to peaks in transcription-inhibited 8-cell embryos. Among H3K4me3, H3K27ac, H3K27me3, and H3K9me3, 73, 77, 52, and 63% of gained peaks were not established, respectively, and 36, 73, 50, and 59% of 4-cell peaks normally erased by the 8-cell stage were still present in 8-cell embryos treated with α -amanitin, respectively. In contrast, for all marks, about 90% of peaks found in both 4- and 8-cell embryos were also present in transcription-inhibited 8-cell embryos (Appendix Table S8). Signal at these gained, lost, and retained peaks revealed that the epigenetic landscape of transcription-inhibited 8-cell embryos more closely resembled that of 4-cell embryos than 8-cell embryos (Fig 5A), particularly for H3K27ac, which gained and lost the most peaks of any mark during the 4- to 8-cell transition (Appendix Table S8).

This disruption of histone remodeling is mirrored by the α -amanitin-sensitive expression of the respective histone modifiers. For instance, erasers of H3K4me3 were highly expressed in 8-cell embryos relative to oocytes and transcription-inhibited embryos (Fig EV2A). In particular, *KDM5B*, which removes ncH3K4me3 in mouse embryos (Dahl *et al*, 2016; Liu *et al*, 2016; Zhang *et al*, 2016), was 30-fold more expressed in control relative to transcriptionally inhibited 8-cell embryos. The time it would take to synthesize sufficient amounts of KDM5B protein from zygotically generated *KDM5B* transcripts may account for the observation that ncH3K4me3 is still present at the 8-cell stage. Regarding H3K27me3 modifiers, components of both polycomb repressive complexes (PRC1 and 2) and the H3K27me3 demethylase *KDM6A* are embryonically expressed in 8-cell embryos (Fig EV2B). This expression could explain the erasure of intragenic H3K27me3 and establishment of intergenic H3K27me3 domains during the 8-cell to morula transition. In contrast, transcript abundance of H3K9me3 remodelers was generally insensitive to transcription inhibition, although *CBX3/HP1-gamma*, which binds to H3K9me3 and excludes H3K27me3 from the same loci (Tardat *et al*, 2015), was upregulated in 8-cell embryos (Fig EV2C), again reflecting the imminent loss of co-occupancy of H3K9me3 and H3K27me3 after the 8-cell stage. Embryonic expression of these key histone modifiers likely underpins the transition from a pre-EGA state, characterized by ncH3K4me3 and intragenic H3K9me3 and H3K27me3, to a post-EGA state, characterized by canonical distributions of these marks.

In contrast to H3K4me3, H3K9me3, and H3K27me3, the distribution of H3K27ac, similar to chromatin accessibility (Halstead *et al*, 2020a), was markedly perturbed in transcriptionally-inhibited 8-cell embryos. Transcription inhibition prevented 74% of regions that gained H3K27ac during the 4- to 8-cell transition from becoming marked and resulted in incorrect retention of H3K27ac at 78% of regions that should have been deacetylated (Fig 5B). These failures to deposit and remove H3K27ac at pertinent loci were reflected by downregulation of both histone acetyltransferases (*p300*, *PCAF*, and *KAT14*) and deacetylases (*HDAC1* and *HDAC2*) in response to α -amanitin treatment (Fig EV2D). Moreover, 6,223 regions

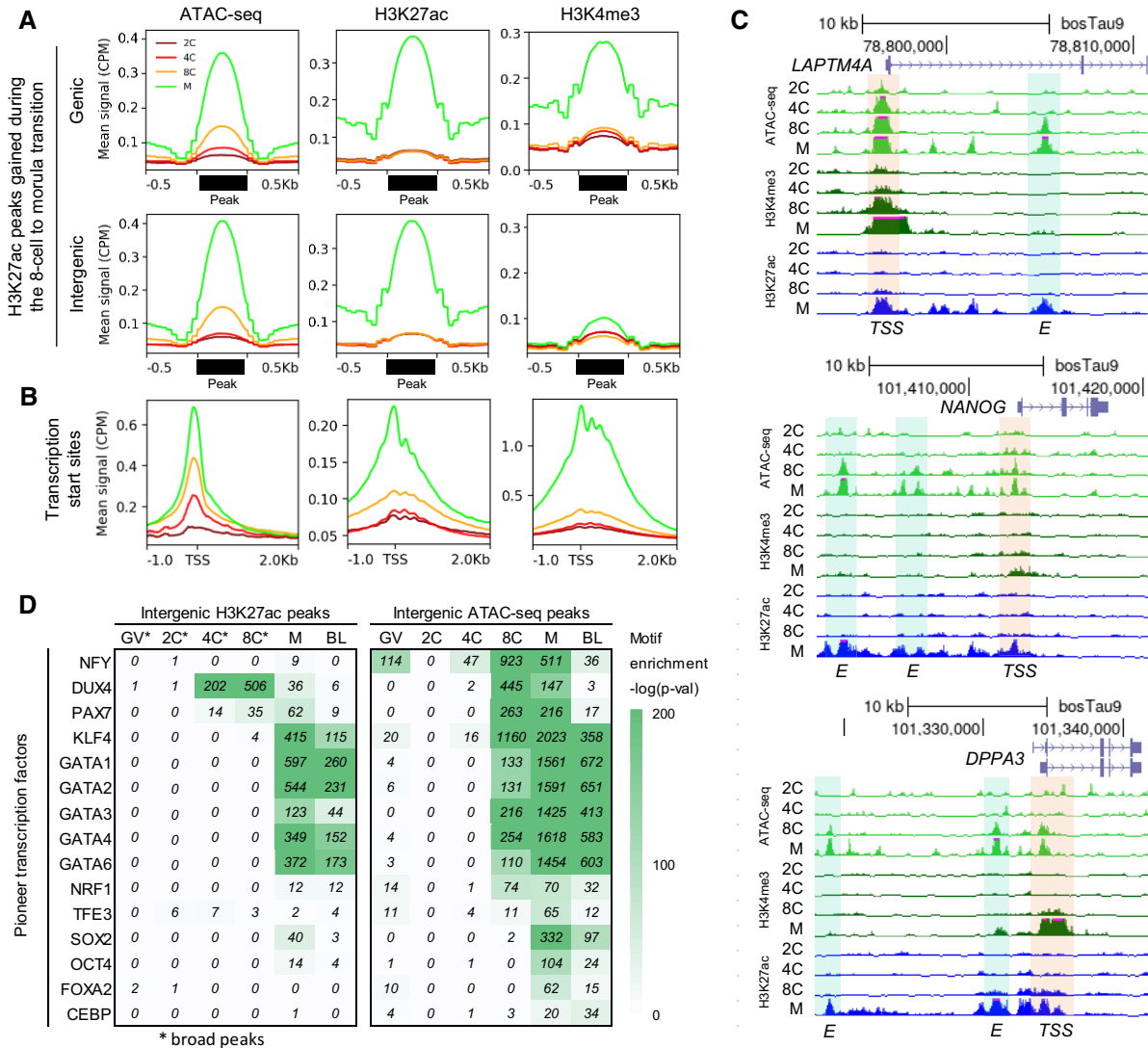


Figure 4. Chromatin accessibility precedes establishment of canonical H3K4me3 and H3K27ac in morula.

A Average ATAC-seq, H3K27ac, and H3K4me3 signal (CPM) at H3K27ac peaks which appeared during the 8-cell to morula transition. Peaks classified as genic if they overlapped gene bodies or promoters (2 Kb upstream of TSS).
 B Average signal at TSS.
 C Normalized ATAC-seq, H3K27ac, and H3K4me3 signal for one biological replicate per developmental stage at the *LAPT4A*, *NANOG*, and *DPPA3* loci. Shaded regions correspond to putative enhancers (E; blue) and promoters (TSS; orange). Viewing range from 0 to 1.5 CPM. Values exceeding maximum range indicated by pink bars.
 D Enrichment of pioneer transcription factor motifs in H3K27ac and ATAC-seq peaks at each stage. From the GV to 8-cell stage, broad H3K27ac peaks (*) were used for motif enrichment analysis, then narrow peaks were used.

accumulated aberrant H3K27ac signal not present in either 4- or 8-cell embryos (Fig EV2E), suggesting that factors directing HATs to target loci are also dysregulated. This aberrant H3K27ac signal was especially evident at TSS, despite reduced H3K4me3 (Fig 5C). Clustering of TSS based on H3K27ac signal revealed that transcription inhibition substantially increased acetylation signal downstream of “cluster_3” loci (cluster 3; $n = 4,137$ TSS; Fig EV2F). This group was again enriched for homeobox genes (Appendix Table S9), e.g., *CDX2* (Fig 5D). Of note, another TSS cluster with high H3K27ac signal in both control and transcriptionally inhibited 8-cell embryos (“cluster_1”; $n = 1,324$ TSS) was also enriched for homeobox genes

(Appendix Table S9), e.g., *OTX2* (Fig 5E). These two distinct patterns at homeobox genes suggest separate modes of regulation for different subsets of developmentally related genes. Overall, a significant portion of TSS (clusters 1, 3, and 5) demonstrated increased H3K27ac signal in 8-cell embryos that developed in the presence of α -amanitin. The corresponding genes were functionally enriched for processes involved in transcription regulation (Appendix Table S9). This generalized increase in H3K27ac at promoters is puzzling given that this mark is normally a hallmark of transcriptional activation. However, because α -amanitin treatment prevents transcription elongation, but not binding of RNA

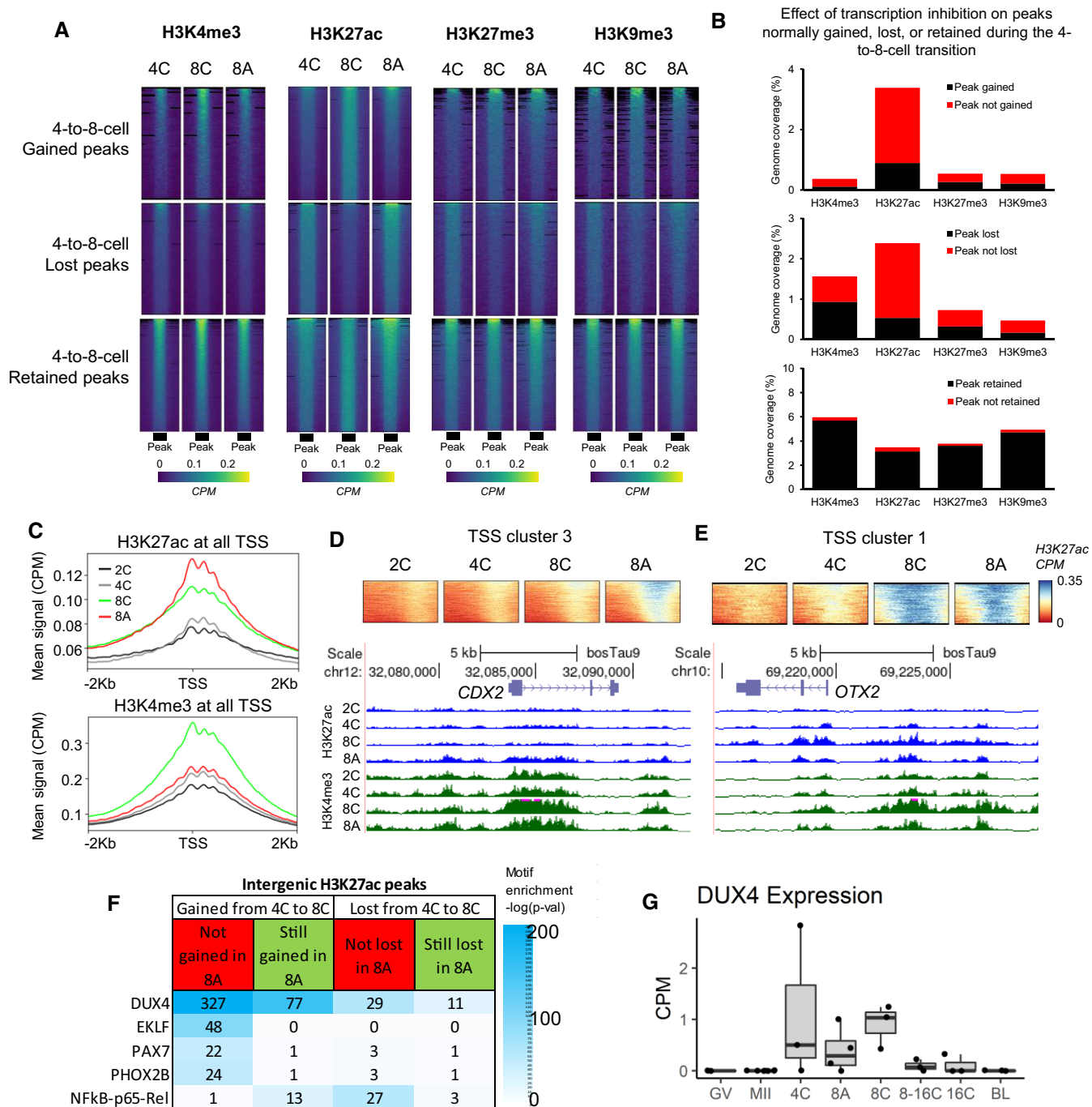


Figure 5. Effect of transcription inhibition on the epigenetic profile of 8-cell embryos.

A Normalized signal (CPM) in 4-cell (4C), 8-cell (8C), and 8-cell embryos cultured in the presence of α -amanitin (8A) at regions that gained, lost, or retained peaks during the 4- to 8-cell transition. Peaks scaled to 500 bp (± 500 bp).

B Genomic coverage of regions that gained, lost, or retained peaks. Shading reflects the impact of transcription inhibition on peak status. Red indicates regions with a change in peak status in 8A embryos compared to 8C.

C Average signal (CPM) of active epigenetic marks at all TSS in 2-cell (2C), 4C, 8C, and 8A embryos.

D Normalized H3K27ac signal at “cluster 3” TSS (± 2 Kb), identified based on H3K27ac signal at TSS in 2C, 4C, 8C, and 8A embryos (Fig EV2), and a gene track view of representative locus, *CDX2*.

E Normalized H3K27ac signal (CPM) at “cluster 1” TSS (± 2 Kb) and a gene track view of a representative locus, *OTX2*. One biological replicate shown per stage and mark. Viewing range of gene tracks from 0 to 1.5 CPM. Values exceeding maximum range indicated by pink bars.

F Enrichment of select transcription factor motifs in intergenic H3K27ac peaks that were gained or lost during the 4- to 8-cell transition, and which were either sensitive or robust to inhibition of embryonic transcription.

G Normalized expression of *DUX4* (ENSBTAG0000049205) across development. Boxplots indicate the median and interquartile range (IQR), and whiskers span 1.5 times the IQR. Data points indicate biological replicates ($n = 3$ for GV, 4C, 8C, 8-16C, 16C, BL; $n = 4$ for MII and 8A).

polymerase II (Bushnell, 2002), polymerase stalling at TSS may be a contributing factor.

Regarding regions that were normally acetylated during the 4- to 8-cell transition, those which gained H3K27ac peaks regardless of α -amanitin treatment were primarily genic (68%; Fig EV2G), and marked genes related to cell division, e.g., actin and microtubule binding (Appendix Table S10). In contrast, regions that failed to gain H3K27ac in transcriptionally inhibited embryos marked genes related to RNA and chromatin binding (Appendix Table S10) and were more often intergenic (57%; Fig EV2G). H3K27ac peaks lost during this transition, regardless whether they were also erased in transcription inhibited embryos, primarily marked genes involved in calcium signaling (Appendix Table S10). Overall, these results suggest that H3K27ac remodeling mediated by maternal factors acts to (i) repress transcriptional programs related to oogenesis and fertilization and (ii) reinforce transcriptional programs related to cleavage. On the other hand, H3K27ac remodeling that depends on embryonic transcription seems to regulate programs related to epigenetic regulation and EGA.

Differential motif enrichment was also evident in these different sets of H3K27ac peaks. Peaks gained from the 4- to 8-cell stage that depended on embryonic transcription were enriched for KLF, PAX7, and PHOX2B motifs, which were not found in peaks gained regardless of α -amanitin treatment, nor in peaks lost during this transition (Fig 5F). The opposite pattern was true for NFkB factors, which can participate in cross-talk with chromatin remodelers, e.g., HAT p300 and acetylation readers such as BRD4 (Zhong et al, 2002; Huang et al, 2009a). NFkB motifs were uniquely present in gained H3K27ac peaks that did not depend on embryonic transcription, as well as regions that incorrectly retained H3K27ac when transcription was inhibited (Fig 5F). As NFkB factors are maternally provided in mice (Nishikimi et al, 1999) and transcripts are present in bovine oocytes (Halstead et al, 2020a), these results suggest dysregulation of a maternal system that directs HATs to target loci. Notably, in all peak sets, the top enriched motif was that of DUX4 (Fig 5F), which is a pioneer factor that induces acetylation by recruiting the HAT p300, and which is increasingly implicated in establishing totipotency (De Iaco et al, 2017; Hendrickson et al, 2017). The DUX4 motif was even enriched in peaks that were gained regardless of α -amanitin treatment (Fig 5F), suggesting that functional DUX4 is present in transcriptionally inhibited embryos. Indeed, DUX4 appears to escape transcription inhibition (Fig 5G). Because DUX4 is one of only a few genes activated during minor EGA (Appendix Fig S6C), and RNA-seq data are unavailable for embryos before the 4-cell stage, DUX4 expression may have initiated before embryos were transferred to culture medium containing α -amanitin. Zygotic expression of DUX4 would be consistent with preliminary studies of human embryos, which indicate that DUX4 expression initiates during the 1-cell stage (Vuoristo et al, 2022).

In summary, the impact of transcription inhibition on the epigenome is most severe for marks that began to resolve earliest, namely, H3K27ac. For marks that do not resolve to a canonical form until after the 8-cell stage (e.g., H3K4me3, H3K27me3, and H3K9me3), the impact of transcription inhibition was less evident at the 8-cell stage. The lack of embryonic expression of the appropriate modifiers, and therefore the inability to resolve the epigenome to a post-EGA state, may underlie the failure of α -amanitin treated embryos to progress past the 8- to 16-cell stage (Memili & First, 1998).

Blastocyst lineages are defined by differential polycomb repression

Following EGA, blastomeres progressively transition from totipotency to pluripotency, and by the blastocyst stage two distinct cell lineages are present: the ICM, a pluripotent lineage that forms the embryo proper, and the TE, a differentiated cell type that contributes to extra-embryonic structures (for detailed reviews, see Simmet et al, 2018; Gerri et al, 2020b). TE specification initiates in morula via a conserved mechanism in which outer blastomeres express GATA3, which upregulates TE-specific programs (Gerri et al, 2020a; Fig 6A). TE specification progresses more quickly in mouse, as the TE-specific factor CDX2 is already expressed in the outer cells of the morula, whereas CDX2 is not detected until blastocyst formation in human and cattle (Fig 6B). Moreover, OCT4/POU5F1 expression remains widespread in human and bovine blastocysts, but is already restricted to the ICM in mouse (Gerri et al, 2020b; Pérez-Gómez et al, 2021; Fig 6B).

We examined our data to identify potential molecular mechanisms responsible for these differences and observed that the *CDX2* locus was polycomb repressed in morula and ICM, whereas *POU5F1* demonstrated activating signatures in all three cell types (Fig 6C). Expected histone modification profiles were also found at other classical markers of ICM (e.g., NANOG and SOX2) and TE (e.g., GATA2; Fig 6C). Moreover, the murine-specific TE marker EOMES, which is not detected in human blastocysts, was polycomb repressed and downregulated in blastocysts, whereas the human-specific TE marker PLAC8, which is not detected in mouse TE, demonstrated stronger H3K27ac in TE and increased expression in blastocysts relative to 16-cell embryos (Fig EV3A and B). Notably, the locus coding for pregnancy associated glycoproteins (PAGs)—a group of ruminant-specific genes expressed in the placenta (Green et al, 2000)—was specifically repressed by H3K9me3 in morula and ICM, but not in TE (Fig EV3A).

Across all of these key loci, both activating and repressive signatures in ICM were very similar to those observed in morula, whereas TE demonstrated a distinct epigenetic profile (Figs 6C and EV3A). The genome-wide distribution of all four histone marks recapitulated this trend, with ICM always falling between morula and TE according to PCA (Fig 1A). The resemblance of morula and ICM was anticipated because the transition from morula to ICM entails a restriction of pluripotency potential rather than differentiation. Therefore, we hypothesized that the distinct TE epigenome resulted from *de novo* deposition of histone modifications during lineage specification. Among H3K4me3, H3K27ac, H3K9me3, and H3K27me3 peaks that were TE-specific (e.g., not present in ICM), 68, 75, 82, and 88%, respectively, were not detected in morula (Figs 6D and EV3C). Of note, 79% of ICM-specific H3K27me3 peaks were also newly established during blastocyst formation (Figs 6D and EV3C), suggesting that polycomb repression plays a major role in specification of both the ICM and TE.

Increased H3K27me3 suggests that an increasingly repressive chromatin landscape is formed as potency declines. This restriction would require reciprocal establishment of enhancer activity to selectively relieve repression and activate cell-specific expression programs. To address whether lineage segregation in the blastocyst follows this model, histone modification and chromatin accessibility data sets were integrated to identify unique chromatin states. The

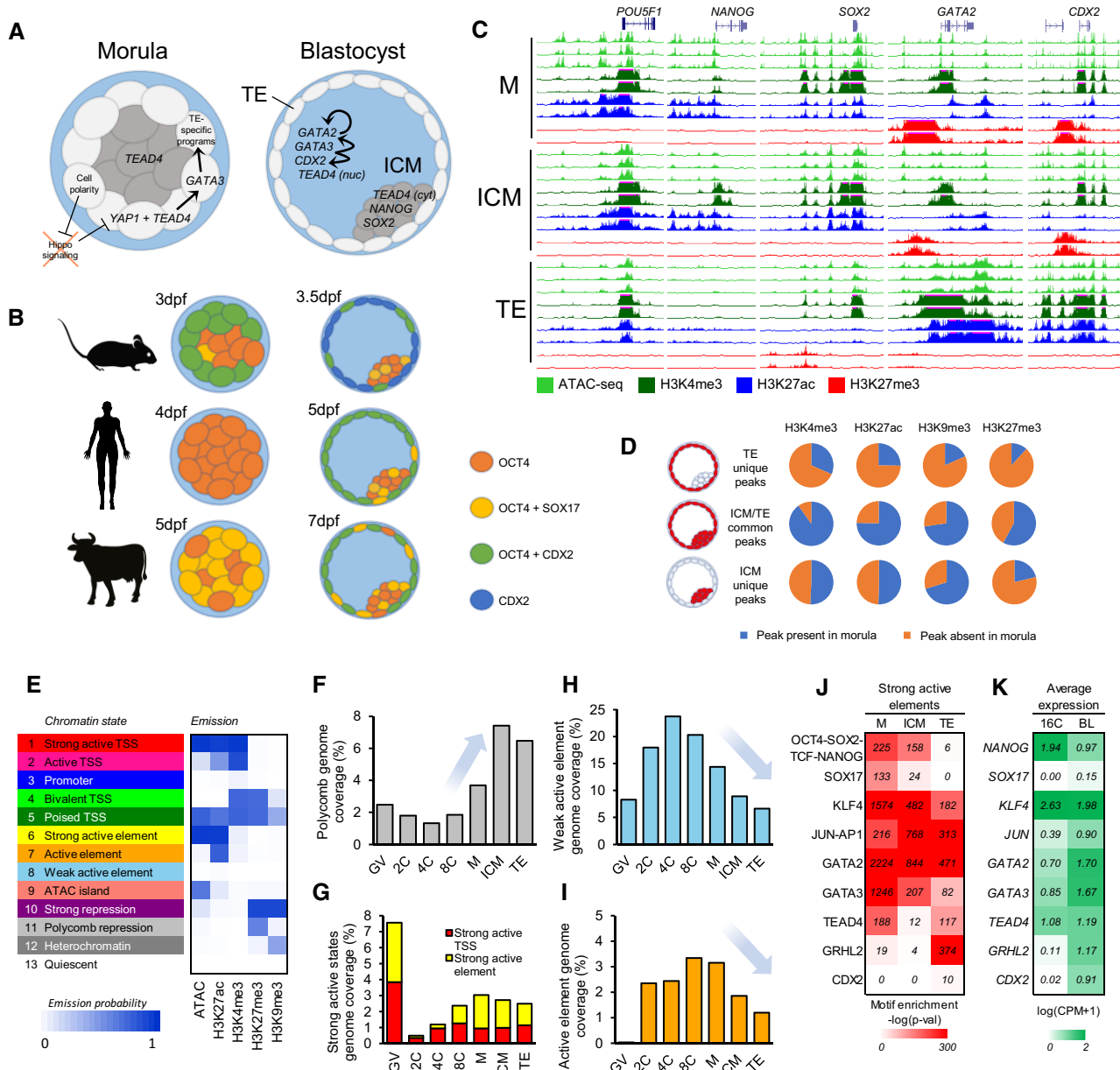


Figure 6. Epigenetic shifts that underscore lineage segregation in the blastocyst.

- A Conserved mechanisms of TE and ICM specification and maintenance in mammals.
- B Species-specific differences in TE and ICM markers in morula and blastocysts. The timing of each stage is indicated by days post fertilization (dpf).
- C Epigenetic profiles of key pluripotency genes and TE-specific markers. Viewing range from 0 to 1.5 CPM. Values exceeding maximum range indicated by pink bars. All biological replicates shown.
- D For each histone modification, proportion of peaks unique to the ICM, unique to the TE, or shared in common between ICM and TE that were already present in morula.
- E Chromatin state predictions based on chromatin accessibility and histone modification data. Emission probabilities indicate the likelihood of a given mark occurring in a given state.
- F–I Genome coverage of (F) polycomb repression, (G) strong active TSS and strong active elements, (H) weak active elements, and (I), active elements by developmental stage.
- J Motif enrichment of selected regulators in strong active elements in M, ICM, and TE.
- K Average normalized expression of selected regulators in 16-cell embryos (16C) and blastocysts.

resulting 13 states were annotated based on genome coverage, enrichment at genes and repetitive elements, CpG content, methylation status, and sequence conservation (Appendix Figs S7 and S8).

Overall, chromatin states represented promoter and enhancer activity, constitutive and facultative heterochromatin, bivalency, and quiescence (Fig 6E). Notably, this analysis revealed that bovine-

specific long terminal repeats (LTRs) were more epigenetically and transcriptionally dynamic during preimplantation development than their conserved counterparts (Fig EV4A and B, and Appendix Fig S9), which is consistent with findings that species-specific repetitive elements play roles in preimplantation development in human and mouse (Kigami *et al*, 2003; Peaston *et al*, 2004; Macfarlan *et al*, 2012; Hendrickson *et al*, 2017). ERVK elements, in particular, were substantially enriched for H3K9me3 in the blastocyst (Fig EV4A and C), with TE- and ICM-specific domains enriched for pluripotency and homeobox motifs, respectively (Appendix Table S11). These results suggest that constitutive heterochromatin contributes to repression of ICM-specific factors in the TE and morula-specific factors in the ICM.

Overall, repression increased substantially during the morula to blastocyst transition and was largely due to an increase in polycomb repression (Fig 6F). Both ICM and TE reached similar levels of repression to that observed in GV oocytes, although oocytes were markedly more enriched for constitutive heterochromatin (H3K9me3), whereas facultative heterochromatin (H3K27me3) was more prevalent in blastocyst lineages (Appendix Fig S9D). Concurrently, the active compartment of the genome steadily decreased after EGA, reaching a minimum in TE (Appendix Fig S9D). However, the most active regions (e.g., “Strong active elements” and “Strong active TSS”) occupied comparable portions of the genome in morula, ICM, and TE (Fig 6G), whereas regions of intermediate activity (e.g., “Weak active elements” and “Active TSS”) were progressively eliminated, eventually becoming the least prevalent in TE (Fig 6H and I). Overall, these trends suggest that lineage specification in the blastocyst entails an overall refinement of chromatin structure, with targeted repression at cell-specific regions and restriction of active regions.

Despite the comparable genome occupancy of polycomb repression and highly active elements (e.g., enhancers and promoters) in ICM and TE, the differential distribution of these chromatin states, combined with TF activity, distinguished these two cell lineages. To identify genes specifically repressed or activated in ICM or TE, H3K27me3 and H3K27ac signal were quantified in promoters (2 Kb upstream of TSS) and compared between the two cell types to identify differentially repressed genes (DRG) and differentially activated genes (DAG; Fig EV5A). In TE, DRG ($n = 1,065$ genes) were functionally enriched for calcium ion binding, and DAG ($n = 1,611$ genes) were enriched for cholesterol biosynthesis (Appendix Table S12). In ICM, DRG ($n = 949$ genes) were enriched for homeobox genes, TFs, and developmental pathways, and DAG ($n = 1,179$ genes) were enriched for TF activity (Appendix Table S12). Notably, although both DRG and DAG in ICM were enriched for TF activity, different regulators were represented in each gene set (Appendix Table S13). Moreover, key genes demonstrated both repression and activation in respective cell types. Loci that were repressed in TE and activated in ICM ($n = 151$ genes) were enriched for cancer pathways and RNA polymerase II core promoter proximal region sequence-specific DNA binding, and loci that were activated in TE and repressed in ICM ($n = 309$ genes) were enriched for conserved homeoboxes, including CDX2 (Fig EV5A).

Considering the overrepresentation of TFs among genes differentially regulated by H3K27me3 and H3K27ac, motif enrichment analysis was conducted on strong active elements—the chromatin state that likely represents active enhancers—to infer key regulators of

ICM and TE identity (Fig 6J). Overall, the motifs of ICM- and TE-specific markers demonstrated the expected patterns of enrichment. Key pluripotency factor-binding motifs (OCT4-SOX2-TCF-NANOG) were differentially enriched in active ICM enhancers relative to TE. Recognition motifs of SOX17, which is heterogeneously expressed in morula and then restricted to the ICM (Gerri *et al*, 2020b), were also enriched in active enhancers in morula and ICM. In contrast, TEAD4, which is involved in TE-specification in morula (Gerri *et al*, 2020a), was specifically enriched in active enhancers in morula and TE, but not in the ICM. These results are consistent with protein localization, as TEAD is markedly cytoplasmic in the ICM and nuclear in the TE (Home *et al*, 2012). Enrichment of the GRHL2 motif was highly specific to active enhancers in the TE (14% of enhancers contained this motif). Notably, in humans, GRHL2 is widespread in both morula and TE (Gerri *et al*, 2020a), but in bovine there was little evidence of GRHL2 activity at the morula stage, potentially indicating a delay in expression of TE-specific genes in cattle, relative to human. In a similar manner, active enhancers in TE demonstrated very little enrichment for CDX2 motifs, despite notable expression in the blastocyst (Fig 6K) and presence of CDX2 protein (Gerri *et al*, 2020b; Pérez-Gómez *et al*, 2021). Thus, in terms of genome regulation, CDX2 may play a more prevalent role during later stages of bovine blastocyst development.

In both morula and blastocyst lineages, motifs of KLFs and GATA factors were heavily overrepresented in active enhancers. However, because of similarities in binding motifs between family members, it remains unclear which specific members are active in a given cell type. For example, whereas GATA3 induces TE fate (Ralston *et al*, 2010; Deglincerti *et al*, 2016; Gerri *et al*, 2020a), GATA6 and GATA4 are implicated in primitive endoderm specification and maintenance (Chazaud *et al*, 2006; Plusa *et al*, 2008; Roode *et al*, 2012). Among KLFs, which generally play roles in pluripotency maintenance, KLF2 marks murine epiblast but is absent in humans, whereas KLF17 is widespread in human blastocysts but absent in mouse (Yeo *et al*, 2014; Blakeley *et al*, 2015; Lea *et al*, 2021). Enrichment of KLF motifs in bovine morula, ICM, and TE is reminiscent of the widespread expression of KLF17 in human embryos, but it remains to be determined which member(s) of the KLF family are active in bovine; several candidates are expressed at the blastocyst stage, including *KLF6*, which undergoes a 14-fold increase in expression between the 16-cell and blastocyst stages (Fig EV5B).

Similar to KLF and GATA factors, the motif of JUN-AP1, a factor recently implicated as a gatekeeper to reprogramming of pluripotency (Markov *et al*, 2021), was also enriched in active enhancers in morula, ICM, and TE. However, based on TF footprinting analysis of chromatin accessibility data, which measures TF activity based on reduced DNA cleavage at motifs actively bound by TFs, the activity of JUN and FOS factors—which together form the AP-1 complex—was markedly increased in ICM relative to morula (Fig EV5C and D) and was accompanied by an increase in expression of both JUN and FOS factors (Fig EV5E). Moreover, the morula to ICM transition entailed a loss in homeobox activity (Fig EV5C and F). These shifts in TF activity distinguished these two pluripotent cell types, which otherwise demonstrated strong epigenetic similarities, both in terms of histone modification profiles (Fig 6A) and motif enrichment in active enhancers (Fig 6J).

Overall, we find that lineage specification entails further refinement of the epigenome, including increased polycomb repression at specific loci, especially those related to future developmental pathways, and resolution of regions with intermediate or weak activity. Because the transition from pluripotency to a differentiated state is generally irreversible, it is possible that by restricting activity to certain genomic regions, the potential to activate others is consequently lost.

Discussion

Accumulating evidence suggests that the molecular underpinnings of preimplantation development are not necessarily conserved within mammals, despite the fundamental nature of this period. Developmental timing, H3K27me3-mediated imprinting, X chromosome inactivation, and TF function during EGA and blastocyst formation all differ considerably between humans and mice (Ross & Sampaio, 2018; Chen & Zhang, 2020; Halstead et al, 2020a; Gerri et al, 2020b; Lu et al, 2021; Pérez-Gómez et al, 2021). These divergences highlight the need to establish additional animal models, such as bovine, to investigate the molecular basis of key processes, such as the MET and blastocyst formation. By profiling core histone modifications and chromatin accessibility during bovine preimplantation development, we identify both conserved and species-specific mechanisms that govern the reprogramming of cell identity in the oocyte and developing embryo.

For instance, although broad, nontranscribed partially DNA methylated domains (PMDs) are present in all mammalian oocytes examined to date, the histone modification profiles associated with these domains are inconsistent between species. In rodent and ungulate oocytes—but not humans—PMDs are carpeted by broad domains of H3K4me3 (Dahl et al, 2016; Liu et al, 2016; Zhang et al, 2016; Zheng et al, 2016)—a mark usually associated with TSS activity, but which in this context is essential for gene silencing (Zhang et al, 2016). Unexpectedly, we find that in bovine oocytes, PMDs are additionally marked by either H3K27me3 or H3K27ac, suggesting at least two subtypes of PMDs, which likely serve distinct functions. As suggested by Zhang et al (2016), acetylated PMDs may act as “molecular sponges” for TFs, thereby globally suppressing transcription, although it remains unknown whether PMDs are acetylated in oocytes of other species. Conversely, PMDs bivalently marked by H3K4me3 and H3K27me3 may poise developmentally related genes for future expression, similar to bivalent domains in stem cells (Bernstein et al, 2006). Notably, bivalency at PMDs has only been observed in cattle and pig oocytes (Lu et al, 2021). In rodent oocytes, H3K4me3 and H3K27me3 instead occupy distinct domains within PMDs, and in human oocytes, only broad H3K27me3 has been observed at PMDs (Xia et al, 2019; Lu et al, 2021). Curiously, although H3K4me3 at PMDs is maintained until EGA in all species (in humans, H3K4me3 is deposited *de novo* at PMDs in embryos), H3K27me3 is erased from these domains in humans and ungulates after fertilization, but retained in rodents until the blastocyst stage (Xia et al, 2019; Lu et al, 2021). These discrepancies strongly suggest that the functional role of PMDs in oocytes and pre-EGA embryos is species specific and warrants investigation in additional animal models.

Even beyond PMDs, substantial portions of oocyte epigenetic signatures are briefly inherited by the embryo after fertilization,

suggesting regulatory roles beyond oogenesis. For instance, gene bodies in bovine oocytes and pre-EGA embryos are pervasively marked by H3K9me3, which usually localizes to repetitive elements and chromosomal periphery, rather than coding regions. Interestingly, in other cell types, genic H3K9me3 stalls RNA polymerase, effectively poisoning lineage-specific genes for future expression (Matsumura et al, 2015). Because H3K9me3 is also enriched at gene bodies in human and murine oocytes (Wang et al, 2018; Xu et al, 2022), genic H3K9me3 may be an additional conserved mechanism to globally suppress transcription in the oocyte, complementing the repressive nature of PMDs. Moreover, genic H3K9me3 is also present in cattle and murine embryos prior to EGA (Wang et al, 2018; Xu et al, 2022), suggesting this mark continues to repress transcription until the epigenome has been sufficiently reprogrammed to permit appropriate expression of developmental programs. Of note, bovine pre-EGA embryos also demonstrate widespread H3K27me3 in gene bodies, which appears to be deposited *de novo* after fertilization, and in a similar fashion to H3K9me3, is maintained up until major EGA. As such, polycomb repression may also contribute to transcription repression—similar to H3K27me3-mediated noncanonical imprinting in mice (Chen & Zhang, 2020; Lu et al, 2021). However, it remains unclear if genic H3K27me3 is present in other mammalian embryos, as a previous study failed to capture the H3K27me3 signal in pre-EGA bovine embryos (Lu et al, 2021), and H3K27me3 profiles in humans were obtained using similar methodology (Xia et al, 2019).

Between PMDs and repression of gene bodies, there appear to be several conserved or semi-conserved mechanisms to suppress transcription in the oocyte and embryo prior to genome activation. The extended pre-EGA period in bovine embryos, as compared to mice, may underlie the observation that this repressive epigenetic landscape is remodeled in a two-step fashion, wherein discrete regions of accessibility are established around the 8-cell stage (beginning of EGA), whereas a classical histone modification landscape is not evident until the morula stage (post-EGA). This remodeling sequence could be attributed to increased pioneer factor activity, leading to establishment of accessible sites in 4- and 8-cell embryos, with later modifications (e.g., H3K27ac) serving to reinforce that accessibility. The specific factors that drive this transition remain unknown and are the subject of ongoing investigation.

Transcription inhibition reduced the activity of certain factors (e.g., DUX4 and PAX7), and ablated histone acetylation at their binding sites in 8-cell embryos; however, binding of other pioneer factors (e.g., KLF4 and GATA factors) was not impaired (Halstead et al, 2020a), suggesting complementary action by both embryonic and maternal factors. Although the activity of some pioneer factors may be conserved, DUX proteins, for example, are implicated in both mouse and human development (De Iaco et al, 2017; Hendrickson et al, 2017; Chen & Zhang, 2019; De Iaco et al, 2020; Bosnakovski et al, 2021; Xu et al, 2022), the TFs active during human and bovine EGA are distinct from mouse (Halstead et al, 2020a). For instance, NFY and GATA factors do not appear active during murine EGA, but are markedly enriched at the 8-cell stage in cattle and humans. Conversely, the pioneer factors DUX4 and KLF4 are enriched in all three species during major EGA. Altogether, this finding suggests that although establishment of accessibility by pioneer factors is a conserved mechanism within mammals, the specific factors may differ. Moreover, maternally inherited factors

(NFY and GATA factors, for example) may play a larger role in species with extended pre-EGA periods, such as humans and cattle, as opposed to rodents, which rapidly initiate EGA.

Transcription factor utilization continues to differ between species during blastocyst formation. For example, the key pluripotency factor OCT4 localizes uniquely to the ICM in mouse blastocysts, but OCT4 is not restricted to the ICM in humans, cattle, pigs, or rabbits (van Eijk *et al*, 1999; Kirchhof *et al*, 2000; Cauffman *et al*, 2004; Chen *et al*, 2009; Hall *et al*, 2009; Kobolak *et al*, 2009; Berg *et al*, 2011; Gerri *et al*, 2020b; Pérez-Gómez *et al*, 2021). Widespread expression of OCT4 may delay TE lineage commitment, such that in both cattle and humans, ICM and TE maintain developmental plasticity and can be readily interconverted (Berg *et al*, 2011; De Paepe *et al*, 2013; Guo *et al*, 2021). In keeping with this plasticity, epigenetic repression in bovine ICM and TE is primarily driven by H3K27me₃, a marker of facultative heterochromatin, as opposed to constitutive heterochromatin, which acts as a more permanent barrier to cell fate changes (Becker *et al*, 2016). In humans, H3K27me₃ is also differentially distributed at ICM- and TE-specific genes (Xia *et al*, 2019), and in mice, H3K9me₃ is not asymmetrically distributed at promoters until after blastocyst implantation (Wang *et al*, 2018). Overall, these results suggest that in mammals polycomb repression is an early contributor to ICM/TE specification, whereas lineage-specific H3K9me₃, especially at promoters, is mostly established after implantation, signaling commitment of the ICM and TE cell identities. Additional studies are needed to determine if asymmetric distribution of H3K9me₃ at promoters becomes apparent at later stages of bovine development.

Concurrently, intergenic H3K9me₃ appears to regulate activity at species-specific transposons in the ICM and TE, entailing species-specific differences in expression of repetitive elements and TF activity. In human TE, H3K9me₃ represses hominoid-specific retrotransposons that harbor pluripotency factor motifs, but no such pattern was detected in ICM (Yu *et al*, 2022). Similarly, H3K9me₃ in bovine TE is enriched at bovine-specific ERVK elements and demonstrates moderate enrichment for OCT4-binding sites. However, H3K9me₃ is also enriched at ERVK elements in bovine ICM, and instead marked homeobox factor motifs. Altogether, this suggests that in both humans and cattle, H3K9me₃ deposition at species-specific LTRs represses ICM-related signaling in the TE, whereas in cattle, H3K9me₃ repression also contributes to the suppression of morula-specific homeobox activity, potentially facilitating the transition from totipotency to pluripotency.

In light of the difficulties entailed by directly studying human development, establishing appropriate models is crucial to push forward our understanding of the preimplantation period. Human oocytes, in particular, remain an especially challenging system, as their epigenetic signature appears distinct from both rodents and ungulates. However, epigenetic remodeling and TF activity during the MET and EGA indicate that bovine embryos are a more appropriate model for human than mouse, at least for this specific window of development. Regarding blastocyst formation, although bovine and human share several key markers of ICM and TE, differences in specific regulators suggest that cell fate commitment in the bovine blastocyst is delayed relative to human. In bovine TE, CDX2 activity is very low, and GRHL2, which is widespread in both morula and TE in humans (Gerri *et al*, 2020a), is only active in bovine TE. These differences suggest a delay in TE commitment,

potentially reflecting the timing of implantation: cattle blastocysts spend two weeks in the uterus before implanting (Berg *et al*, 2011), whereas human blastocysts implant after about two days, and mice implant one day after blastocyst formation (Niakan *et al*, 2012). Taken together, these results suggest that bovine could be a better model for human blastocysts than mouse, but more appropriate models may exist. Rabbit, for instance, has not yet been sufficiently characterized, although rabbit embryos initiate EGA at the 8-cell stage (Manes, 1973), similarly to humans, and broadly express OCT4 in the blastocyst (Kobolak *et al*, 2009), which implants shortly (3–4 days) after formation (Orsini, 1962).

To date, it appears that several mechanisms underlying the MET and blastocyst formation in mice are distinct from those employed by larger mammals and perhaps evolved to allow rodents to progress more quickly through development. Consequently, bovine embryos could constitute an appropriate alternative to study the biology of human development. To this end, the atlas of chromatin states produced by this study will be an invaluable resource for future research in mammalian embryos, ranging from functional studies of specific genes to broadening our understanding of genome reprogramming during the preimplantation period.

Materials and Methods

Collection of bovine oocytes and preimplantation embryos

Bovine ovaries were collected from a local slaughterhouse and transported to the laboratory in 0.9% NaCl solution, complying with the guidelines of University of California Davis relevant ethical regulation for animal research. Cumulus-oocyte complexes (COCs) were aspirated from antral follicles (2–10 mm in diameter) with a 10 ml syringe and then washed in M199 (Sigma, M7653) containing 2% (v/v) fetal bovine serum (FBS, Hyclone). Only COCs with at least three layers of compact cumulus cells and evenly granulated cytoplasm were selected for maturation *in vitro*. After washing three times in M199 supplemented with 2% (v/v) FBS, COCs were cultured for 22–24 h in BO-IVM media (IVF Bioscience, 71001) in an atmosphere of 5% CO₂ in air at 38.5°C. After maturation *in vitro*, MII oocytes were washed three times in SOF-IVF medium (107.7 mM NaCl, 25.07 mM NaHCO₃, 7.16 mM KCl, 1.19 mM KH₂PO₄, 1.17 mM CaCl₂, 0.49 mM MgCl₂, 5.3 mM sodium lactate, 0.20 mM sodium pyruvate, 10 µg/ml heparin, 0.5 mM fructose, 1× nonessential amino acids, 6 mg/ml BSA, 5 µg/ml gentamicin) and then transferred to 90 µl drops (50 COCs per drop) of SOF-IVF covered with mineral oil (Vitrolife, 10029). Frozen semen was thawed in water at 37°C for 1 min. Sperm were selected using density gradient centrifugation, washed once using TALP-Sperm medium (100 mM NaCl, 25 mM NaH₂CO₃, 3.1 mM KCl, 2.1 mM CaCl₂, 0.29 mM NaH₂PO₄, 0.4 mM MgCl₂, 21.6 mM sodium lactate, 10 mM Hepes, 6 mg/ml BSA, 5 µg/ml gentamicin), counted, and 10 µl of 2 × 10⁶ sperm/ml were added to each drop containing matured oocytes. After incubating for 6 h in an atmosphere of 5% CO₂ in air at 38.5°C, cumulus cells were removed by vortexing for 5 min and zygotes were washed three times in BO-IVC media (IVF Bioscience, 71005) and then transferred to 100 µl drops (50 zygotes per drop) of BO-IVC media covered with mineral oil. For each collection, 50 zygotes were set aside and cultured to the blastocyst

stage and only collections for which the incidence of blastocyst formation was at least 20% were used.

Cumulus cells were removed from COCs by vortexing for 5 min, and GV oocytes were collected for CUT&Tag library construction. Embryos at the 2-, 4-, 8-cell (cultured in the presence or absence of 50 µg/ml α -amanitin (Sigma, A2263)), morula, and blastocyst stages were collected at 32, 44, 56, 122, and 172 h post-insemination, respectively. For ICM collection, blastocysts were collected at day 7 and subjected to immunosurgery with antiovine serum antibody (Sigma, B8270) and guinea pig complement serum (Innovative Research, IGP-COMPL-21249) as previously described (Halstead *et al.*, 2020a). For TE separation, blastocysts were collected at day 7 and subjected to surgery under a microscope, cutting with a micro scalpel (Feather, 72046-30). GV oocytes and embryos were treated with 0.5% pronase to remove the zona pellucida and then used for library construction. A minimum of 500 oocytes/blastomeres were used for each CUT&Tag or ATAC-seq library.

CUT&Tag library construction and sequencing

CUT&Tag was performed following the manufacturer's instructions (Hyperactive In-Situ ChIP Library Prep Kit, TD901, Vazyme) with modifications. Individual CUT&Tag libraries were constructed from about 500 GV oocytes, 200 2-cell embryos, 150 4-cell embryos, 150 8-cell embryos, 50 morula, 50 day 7 blastocysts, 150 isolated ICM, or 150 isolated TE. In brief, bovine oocytes and embryos were incubated with concanavalin-coated magnetic beads for 20 min on a thermomixer at 400 rpm at room temperature (RT). The samples then were incubated with a primary antibody (1:50 dilution of a rabbit polyclonal anti-H3K4me3 (Cat# C15410003), H3K9me3 (Cat# C15410193), H3K27me3 (Cat# C15410195), or 1:100 dilution of H3K27ac (Cat# C15410196) from Diagenode) overnight at 4°C on a nutator, then with the secondary antibody (1:100 dilution, ABIN101961, Antibodies online) for 1 h on a nutator at room temperature (RT), then with hypoactive pA-Tn5 transposon for 1 h on a nutator at RT. To perform targeted digestion, 300 µl tagmentation buffer was then added, and samples were incubated at 37°C for 1 h. The reaction was terminated by adding 10 µl of 0.5 M EDTA, 3 µl of 10% SDS, and 2.5 µl of 20 mg/ml proteinase K to each tube, which were then incubated at 50°C for 1 h. DNA extraction was performed by adding 300 µl of PCI (Phenol: Chloroform: Isoamyl alcohol = 25:24:1) to each tube and mixed by full-speed vortexing, after which the sample was transferred to a phase-lock tube (1038987, Qiagen) and centrifuged for 3 min at RT at 16,000 × g. Then, 300 µl chloroform was added followed by centrifugation for 3 min at RT at 16,000 × g. The aqueous phase was removed and transferred to a 1.5 ml tube to which 750 µl of 100% ethanol was added, the tube inverted 10 times, and then centrifuged for 30 min at 4°C at 16,000 × g. The supernatant was removed, and 1 ml of 100% ethanol added, followed by centrifugation for 5 min at 4°C at 16,000 × g. The supernatant was removed and the tube then allowed to dry in air. TE buffer (20 µl) was added and PCR was performed using NEBNext High-Fidelity 2X PCR Master mix (New England Biolabs, M0541) as follows: 58°C for 5 min, 7°C for 5 min, 98°C for 45 s, followed by 14 cycles of 98°C for 15 s and 60°C for 10 s, with a final extension at 72°C for 1 min. Purification of PCR products was performed using Ampure XP beads (Bechman Coulter, A63881), and libraries were pooled for

sequencing on the Illumina NextSeq platform to generate 40 bp paired-end reads.

ATAC-seq library construction and sequencing

Pools of zona-free blastocysts, ICM, and TE were collected for ATAC-seq library construction from three separate collections, respectively, and transferred to cold lysis buffer (10 mM Tris-HCl pH7.4, 10 mM NaCl, 3 mM MgCl₂, and 0.1% IGEPAL CA-630). Individual ATAC-seq libraries were constructed from about 50 7-day blastocysts, 150 isolated ICM, or 150 isolated TE. The samples were incubated on ice for 5 min and then centrifuged using a swinging bucket rotor for 10 min at 500 × g at 4°C. The supernatant was removed, and the pellet was resuspended in 50 µl of transposition reaction mix (25 µl 2× TD buffer (Nextera DNA Library Prep Kit, Illumina), 2.5 µl TDE1 enzyme (Nextera DNA Library Prep Kit, Illumina), 22.5 µl ddH₂O) by pipetting up and down three times. Transposition reactions were incubated at 37°C for 60 min in a thermomixer at 300 rpm. Reaction products were purified with a MinElute PCR purification kit (Qiagen, Germany) and eluted in 20 µl EB buffer. Transposed DNA was amplified with NEBNext High-Fidelity 2X PCR Master mix (New England Biolabs, M0541) as follows: 58°C for 5 min, 72°C for 5 min, 98°C for 45 s, followed by 12 cycles of 98°C for 15 s and 60°C for 10 s, with a final extension at 72°C for 1 min. Libraries were purified with Ampure XP beads (Bechman Coulter, A63881) and were pooled for sequencing on the Illumina NextSeq platform to generate 40 bp paired-end reads.

CUT&Tag and ATAC-seq data processing

The following published data sets were accessed through the NCBI Gene Expression Omnibus (GEO) repository: ATAC-seq data for bovine oocytes and embryos (accession GSE143658; Data ref: Halstead & Ross, 2020), CUT&Tag data for bovine fibroblasts (GSE171104; Data ref: Halstead *et al.*, 2020b), and CUT&RUN data for bovine oocytes and blastocysts (GSE163620; Data ref: Lu *et al.*, 2020). Raw reads were trimmed with Trim Galore (v0.4.0) and Cutadapt (v1.8.3; Martin, 2011) with options “-q 20 -stringency 1 -length 10” to remove adaptor sequences and low-quality reads ($q < 20$). Trimmed reads were then mapped to the cattle reference genome (ARS-UCD1.2), excluding the mitochondrial chromosome, using BWA (v0.7.9a; Li & Durbin, 2010) in “mem” mode and with default parameters. Duplicate reads were removed using Picard-Tools (v2.26.10) and low-quality mapped reads ($q < 15$) were removed using SAMtools (v1.9; Li *et al.*, 2009) to obtain the final informative reads used for downstream analyses. Analysis was not blinded.

RNA-seq data processing

RNA-seq data for bovine oocytes and embryos were accessed through the NCBI GEO repository (accession no. GSE52415 (Data ref: Graf *et al.*, 2014a) and GSE110040 (Data ref: Bogliotti *et al.*, 2018)). Raw reads were trimmed with Trimmomatic (v0.33; Bolger *et al.*, 2014) to remove low-quality leading and trailing bases (three bases) and adapter sequences, allowing for two seed mismatches, a palindrome clip threshold of 30, and simple clip threshold of 10. Trimmed reads were mapped to the cattle reference

genome (ARS-UCD1.2) with STAR (v2.7.2a; Dobin *et al*, 2013) with default parameters, except “—seedSearchStartLmax 30 —outFilterScoreMinOverLread 0.85.” Low-quality alignments ($q < 5$) were filtered using SAMtools. HTseq-count (v0.10.0; Love *et al*, 2014) was used to calculate the raw count for genes in the Ensembl (v104) annotation with parameters “—mode = intersection-nonempty —type = exon.” Differentially expressed genes (DEG) were identified using the DESeq2 R package (v1.34.0; Anders & Huber, 2010). To identify maternal and embryonic gene sets, DEG ($\log_2FC > 1$ and $FDR < 0.01$) were determined from pair-wise comparisons of MII oocytes, 8- to 16-cell embryos, and 8- to 16-cell embryos cultured in the presence of α -amanitin, as described previously (Bogliotti *et al*, 2020). To identify minor EGA genes, DEG ($\log_2FC > 1$ and $FDR < 0.05$) were determined between MII oocytes and 4-cell embryos (Graf *et al*, 2014b). A looser FDR cutoff was used to capture subtler differences in gene expression. TEtoolkit in multimode (Lerat *et al*, 2017) was used to calculate raw counts for repetitive elements in the RepeatMasker annotation (downloaded from the UCSC Genome Browser). DESeq2 (v1.34.0) was used to identify differentially expressed repeats between the 4- and 16-cell stage ($FDR < 0.01$). Raw counts were normalized by counts per million (CPM) and z-score transformed for visualization.

DNA methylation data processing and identification of PMDs

Post-bisulfite adaptor tagging (PBAT) DNA library data for bovine oocytes and embryos were accessed through the NCBI GEO repository (accession GSE143850 (Data ref: Ivanova *et al*, 2020b)). Raw reads were aligned to the cattle reference genome (ARS-UCD1.2) using Bismark (v0.14.3; Krueger & Andrews, 2011), with options “—bowtie2 —pbat.” Context-dependent CpG methylation was then extracted using the “bismark_methylation_extractor” command with options “—s —bedGraph —cytosine_report.” To identify PMDs, the genome was first binned into 1 Kb windows with the BEDtools (v2.27.1; Quinlan & Hall, 2010) command makewindows. Bins that contained 5 or more CpGs, and which had an average methylation level less than or equal to 0.5 in GV oocytes, were considered hypomethylated windows. These were merged, allowing for a maximum gap between windows of 10 Kb. Regions that fell within 2.5 Kb of TSS were excluded from the final set of PMDs using BEDtools subtract. Methylation status of genomic features (e.g., gene bodies, peaks, chromatin states, etc.) was determined by taking the average methylation level of all CpGs that overlapped the given feature.

Reproducibility of CUT&tag and ATAC-seq data

Informative reads were converted to bigwig format, and signal was normalized by counts per million (CPM) with bamCoverage from the DeepTools suite (v3.4.3; Ramírez *et al*, 2014) with default parameters. Resulting bigwig files were uploaded to the UCSC genome browser (Kent *et al*, 2002) for track visualization. The viewing range was set from 0 to 1.5 CPM, and values exceeding the maximum range were indicated by pink bars. The smoothing window was set to 4 pixels, and windowing function was set to “mean + whiskers,” which shows the average signal in the darkest shade, one standard deviation away from the mean in a medium shade, and the maximum and minimum in the lightest shade. To

calculate the correlation between libraries and conduct principal components analyses, bigwig files were first consolidated using the multiBigWigSummary function from Deeptools, using a bin size of 500 bp for activating marks (ATAC-seq, H3K4me3, H3K27ac) and 10 Kb for repressive marks (H3K9me3, H3K27me3). The output was then piped to plotPCA with parameters “—transpose —log2 —ntop 100000” (active marks) or “—transpose —log2 —ntop 50000” (repressive marks), and plotCorrelation with parameters “—corMethod pearson —skipZeros.” Finally, files containing normalized signal (RPKM; reads per kilobase million) of published CUT&RUN H3K27me3 data for bovine GV oocytes and blastocysts were downloaded from GSE163620 (Lu *et al*, 2021). Using the liftOver tool from the UCSC tool suite, these signal files were converted to the most recent genome assembly. To compare these data to CUT&Tag, the H3K27me3 libraries from GV oocytes and blastocysts were also normalized by RPKM using bamCoverage from DeepTools, and the multiBigWigSummary and plotCorrelation functions were implemented as previously described to determine correlation between libraries.

Peak calling

To identify regions with signal enrichment, or “peaks,” for each library, we first compared two different methods for peak calling: Epic2 (Stovner & Sætrom, 2019) and Macs2 (Zhang *et al*, 2008), using different parameters for narrow and broad epigenetic marks (Appendix Table S14). Peaks from biological replicates were compared with the BEDtools jaccard command. We found that Epic2 generated more consistent peak calls between biological replicates, was less sensitive to differences in read depth, and more logically captured regions of enrichment, especially for repressive marks (Appendix Fig S10). Therefore, Epic2 was selected as the optimal peak caller, using different parameters for narrow and broad peaks (Appendix Table S14). Unless otherwise stated, only the overlapping regions, identified by BEDtools intersect, of peaks called in both biological replicates were identified as “true” peaks and used for further analyses. Peaks were classified as genic if they overlapped promoters (2 Kb regions upstream of TSS) or gene bodies. Otherwise, peaks were considered intergenic. Genomic coverage of peaks was calculated using BEDtools genomecov.

Enrichment of CUT&Tag and ATAC-seq signal

To visualize the signal intensity of histone modifications and chromatin accessibility at specific genomic regions (e.g., TSS, gene bodies, peaks, PMDs), normalized signal from individual libraries or pooled replicates was analyzed using the computeMatrix function of DeepTools with option “—skipZeros.” Signal at PMDs was visualized using a bin size of 100 bp. Signal at genes and peaks was visualized using a bin size of 10 bp. The output was then piped to the plotProfile and plotHeatmap functions from DeepTools. In cases where regions were clustered, the option “—kmeans” was used with plotHeatmap and cluster members were extracted using “—outFile-SortedRegions.” Signal at PMDs was summarized by taking the average of all 100 bp bins for a given PMD, and then calculating the mean signal and standard error of all PMDs in a given cluster for a given developmental stage.

Comparison of DNA methylation and intragenic H3K27me3 signal

For H3K27me3 libraries, signal in gene bodies was quantified using HTseq-count with parameters “—mode = intersection-nonempty—type = gene.” Raw counts were processed using DESeq2 to obtain log₂FC values between GV oocytes and 2-cell embryos. Correlations between gene body H3K27me3 (log₂FC) and change in CpG methylation (2-cell versus GV oocytes) were calculated by linear regression.

Functional enrichment analysis

Gene sets were identified based on overlap with peaks or chromatin states using the intersect function from BEDtools. Functional enrichment analysis of gene sets was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID, v2021q4; Huang *et al*, 2009b, 2009c). Terms with an FDR less than 0.05 were considered significant.

Identification of differentially activated and repressed genes in ICM and TE

For H3K27me3 and H3K27ac libraries, signal in promoters (2 Kb regions upstream of genes) was quantified using HTseq-count in mode “intersection-nonempty.” Genes with differential promoter signal (log₂FC > 1 and FDR < 0.05) between ICM and TE were identified using DESeq2, and counts were z-score transformed for visualization.

Transcription factor footprinting analysis

Transcription factor footprints were identified from ATAC-seq data using HINT from the Regulatory Genomic Toolbox (Li *et al*, 2019) as previously described (Halstead *et al*, 2020a). For footprint detection, alignments from biological replicates were pooled using the merge function from SAMtools, and peaks were called from pooled alignments using Epic2. TFs with a significant difference in binding activity between two cells types were reported (two-tailed *t* test; *P* < 0.05).

Chromatin state annotation

ChromHMM (v1.22; Ernst & Kellis, 2012) was used to train a chromatin state model incorporating CUT&Tag (H3K4me3, H3K27ac, H3K27me3, and H3K9me3) and ATAC-seq from all developmental stages except for blastocyst. The biological replicates of the same developmental stage were merged into one file using the merge function from SAMtools. Specifically, the alignment files from all developmental stages were first binarized using the “BinarizeBam” command with the default parameters. The output was then piped to generate the segmentation model using the “LearnModel” command with the default parameters. Multistate models were trained and the 13-state model was finally selected as it exhibited the maximum number of states with distinct chromatin mark combinations. Fold enrichment of chromatin states at genomic features was calculated using the “OverlapEnrichment” command. Fold enrichment was calculated as (C/A)/(B/D), where A represented the bases in a chromatin state, B the bases in a genomic feature, C

the number of bases overlapped between a chromatin state and genomic feature, and D the bases in the genome. Chromatin state-fold enrichment was calculated for gene elements (Ensembl v104 annotation), repetitive element families (RepeatMasker downloaded from UCSC genome annotation database), and mammalian conserved elements identified from multiple sequence alignments using the Genomic Evolutionary Rate Profile (GERP) software based on 103 mammals (ftp://ftp.ensembl.org/pub/release-100/bed/ensembl-compara/103_mammals.gerp_constrained_element/). To determine if a specific type of repetitive element was enriched for a given chromatin state at a given developmental stage, the overlap of repetitive elements with a given chromatin state (“observed” overlap) was compared to the overlap of randomized intervals (repetitive element locations randomized using BEDtools shuffle) with the same chromatin state (“expected” overlap). The ratio of observed to expected was log-transformed to determine the log ratio (LR) of random to observed.

Data availability

CUT&Tag and ATAC-seq data produced in this study have been deposited in the NCBI GEO database (GSE193640). A UCSC track hub is available to view predicted chromatin states, ATAC-seq, and CUT&Tag read depth (https://genome-euro.ucsc.edu/s/mmhalstead/Bovine_Embryo_Epigenome).

Expanded View for this article is available [online](#).

Acknowledgements

The authors thank the DNA Technologies Core at the University of California, Davis, for their assistance with sequencing and data processing. Funding for this project was provided by NIH grant HD070044 to PJR and RMS. MMH and AB-G were funded by the REVIVE Labex (Investissement d’Avenir, ANR-10-LABX-73) and supported by the PHASE Department of the French National Research Institute for Agriculture, Food and Environment (INRAE).

Author contributions

Chuan Zhou: Conceptualization; formal analysis; investigation; methodology; writing – original draft. **Michelle M Halstead:** Formal analysis; methodology; writing – original draft. **Amèlie Bonnet-Garnier:** Supervision; writing – original draft. **Richard M Schultz:** Conceptualization; supervision; writing – original draft; project administration. **Pablo J Ross:** Conceptualization; supervision; writing – original draft; project administration.

Disclosure and competing interests statement

The authors declare that they have no conflict of interest.

References

- Ahmed K, Dehghani H, Rugg-Gunn P, Fussner E, Rossant J, Bazett-Jones DP (2010) Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. *PLoS ONE* 5: e10531
- Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11: R106
- Becker JS, Nicetto D, Zaret KS (2016) H3K9me3-dependent heterochromatin: barrier to cell fate changes. *Trends Genet* 32: 29–41

- Berg DK, Smith CS, Pearton DJ, Wells DN, Broadhurst R, Donnison M, Pfeffer PL (2011) Trophectoderm lineage determination in cattle. *Dev Cell* 20: 244–255
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125: 315–326
- Blakeley P, Fogarty NME, Del Valle I, Wamaitha SE, Hu TX, Elder K, Snell P, Christie L, Robson P, Niakan KK (2015) Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* 142: 3151–3165
- Bogliotti YS, Chung N, Paulson EE, Chitwood J, Halstead M, Kern C, Schultz RM, Ross PJ (2018) Gene Expression Omnibus GSE110040 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE110040>). [DATASET]
- Bogliotti YS, Chung N, Paulson EE, Chitwood J, Halstead M, Kern C, Schultz RM, Ross PJ (2020) Transcript profiling of bovine embryos implicates specific transcription factors in the maternal-to-embryo transition. *Biol Reprod* 102: 671–679
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120
- Bosnakovski D, Gearhart MD, Ho Choi S, Kyba M (2021) Dux facilitates post-implantation development, but is not essential for zygotic genome activation. *Biol Reprod* 104: 83–93
- Buenostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10: 1213–1218
- Bushnell DA (2002) Structural basis of transcription: a-amanitin-RNA polymerase. 2. Cocrystal at 2.8 inches resolution. *Proc Natl Acad Sci USA* 99: 1218–1222
- Canovas S, Cibelli JB, Ross PJ (2012) Jumonji domain-containing protein 3 regulates histone 3 lysine 27 methylation during bovine preimplantation development. *Proc Natl Acad Sci USA* 109: 2400–2405
- Cauffman G, van de Velde H, Liebaers I, van Steirteghem A (2004) Oct-4 mRNA and protein expression during human preimplantation development. *Mol Hum Reprod* 11: 173–181
- Chazaud C, Yamanaka Y, Pawson T, Rossant J (2006) Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev Cell* 10: 615–624
- Chen Z, Zhang Y (2019) Loss of DUX causes minor defects in zygotic genome activation and is compatible with mouse development. *Nat Genet* 51: 947–951
- Chen Z, Zhang Y (2020) Maternal H3K27me3-dependent autosomal and X chromosome imprinting. *Nat Rev Genet* 21: 555–571
- Chen AE, Egli D, Niakan K, Deng J, Akutsu H, Yamaki M, Cowan C, Fitz-Gerald C, Zhang K, Melton DA (2009) Optimal timing of inner cell mass isolation increases the efficiency of human embryonic stem cell derivation and allows generation of sibling cell lines. *Cell Stem Cell* 4: 103–106
- Choi SH, Gearhart MD, Cui Z, Bosnakovski D, Kim M, Schennum N, Kyba M (2016) DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes. *Nucleic Acids Res* 44: 5161–5173
- Dahl JA, Jung I, Aanes H, Greggains GD, Manaf A, Lerdrup M, Li G, Kuan S, Li B, Lee AY et al (2016) Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* 537: 548–552
- de Iaco A, Planet E, Coluccio A, Verp S, Duc J, Trono D (2017) DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat Genet* 49: 941–945
- de Iaco A, Verp S, Offner S, Grun D, Trono D (2020) DUX is a non-essential synchronizer of zygotic genome activation. *Development* 147: dev177725
- de Paepe C, Cauffman G, Verloes A, Sterckx J, Devroey P, Tournaye H, Liebaers I, van de Velde H (2013) Human trophectoderm cells are not yet committed. *Hum Reprod* 28: 740–749
- Degliincerti A, Croft GF, Pietila LN, Zernicka-Goetz M, Siggia ED, Brivanlou AH (2016) Self-organization of the in vitro attached human embryo. *Nature* 533: 251–254
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15–21
- 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
- Eckersley-Maslin MA, Alda-Catalinas C, Reik W (2018) Dynamics of the epigenetic landscape during the maternal-to-zygotic transition. *Nat Rev Mol Cell Biol* 19: 436–450
- Ernst J, Kellis M (2012) ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 9: 215–216
- Frum T, Ralston A (2015) Cell signaling and transcription factors regulating cell fate during formation of the mouse blastocyst. *Trends Genet* 31: 402–410
- Gao L, Wu K, Liu Z, Yao X, Yuan S, Tao W, Yi L, Yu G, Hou Z, Fan D et al (2018) Chromatin accessibility landscape in human early embryos and its association with evolution. *Cell* 173: 248–259
- Gerri C, McCarthy A, Alanis-Lobato G, Demtschenko A, Bruneau A, Loubersac S, Fogarty NME, Hampshire D, Elder K, Snell P (2020a) Initiation of a conserved trophectoderm program in human, cow and mouse embryos. *Nature* 587: 443–447
- Gerri C, Mencheri S, Mahadevaiah SK, Turner JMA, Niakan KK (2020b) Human embryogenesis: a comparative perspective. *Annu Rev Cell Dev Biol* 36: 411–440
- Graf A, Krebs S, Zakhartchenko V, Schwab B, Blum H, Wolf E (2014a) Gene Expression Omnibus GSE52415 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE52415>). [DATASET]
- Graf A, Krebs S, Zakhartchenko V, Schwab B, Blum H, Wolf E (2014b) Fine mapping of genome activation in bovine embryos by RNA sequencing. *Proc Natl Acad Sci USA* 111: 4139–4144
- Green JA, Xie S, Quan X, Bao B, Gan X, Mathialagan N, Beckers J-F, Roberts RM (2000) Pregnancy-associated bovine and ovine glycoproteins exhibit spatially and temporally distinct expression patterns during pregnancy. *Biol Reprod* 62: 1624–1631
- Greenberg MVC, Bourc'his D (2019) The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol* 20: 590–607
- Guo G, Stirparo GG, Strawbridge SE, Spindlow D, Yang J, Clarke J, Dattani A, Yanagida A, Li MA, Myers S (2021) Human naive epiblast cells possess unrestricted lineage potential. *Cell Stem Cell* 28: 1040–1056
- Hagarman JA, Motley MP, Kristjansdottir K, Soloway PD (2013) Coordinate regulation of DNA methylation and H3K27me3 in mouse embryonic stem cells. *PLoS ONE* 8: e53880
- Hall VJ, Christensen J, Gao Y, Schmidt MH, Hyttel P (2009) Porcine pluripotency cell signaling develops from the inner cell mass to the epiblast during early development. *Dev Dyn* 238: 2014–2024
- Halstead MM, Ross PJ (2020) Gene Expression Omnibus GSE143658 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE143658>). [DATASET]
- Halstead MM, Ma X, Zhou C, Schultz RM, Ross PJ (2020a) Chromatin remodeling in bovine embryos indicates species-specific regulation of genome activation. *Nat Commun* 11: 1–16

- Halstead MM, Navarro M, Ross PJ (2020b) Gene Expression Omnibus GSE171104 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE171104>). [DATASET]
- Hendrickson PG, Doráis JA, Grow EJ, Whiddon JL, Lim J-W, Wike CL, Weaver BD, Pflueger C, Emery BR, Wilcox AL *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
- Home P, Saha B, Ray S, Dutta D, Gunewardena S, Yoo B, Pal A, Vivian JL, Larson M, Petroff M (2012) Altered subcellular localization of transcription factor TEAD4 regulates first mammalian cell lineage commitment. *Proc Natl Acad Sci USA* 109: 7362–7367
- Huang B, Yang X-D, Zhou M-M, Ozato K, Chen L-F (2009a) Brd4 coactivates transcriptional activation of NF- κ B via specific binding to acetylated RelA. *Mol Cell Biol* 29: 1375–1387
- Huang DW, Sherman BT, Lempicki RA (2009b) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44–57
- Huang DW, Sherman BT, Lempicki RA (2009c) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37: 1–13
- Ivanova E, Canovas S, Garcia-Martínez S, Romar R, Lopes JS, Rizos D, Sanchez-Calabuig MJ, Krueger F, Andrews S, Perez-Sanz F (2020a) DNA methylation changes during preimplantation development reveal interspecies differences and reprogramming events at imprinted genes. *Clin Epigenetics* 12: 1–18
- Ivanova E, Canovas S, Garcia-Martínez S, Romar R, Lopes JS, Rizos D, Sanchez-Calabuig MJ, Krueger F, Andrews S, Perez-Sanz F (2020b) Gene Expression Omnibus GSE143850 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE143850>). [DATASET]
- Jambhekar A, Dhall A, Shi Y (2019) Roles and regulation of histone methylation in animal development. *Nat Rev Mol Cell Biol* 20: 625–641
- Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S (2019) CUT&tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* 10: 1–10
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D (2002) The human genome browser at UCSC. *Genome Res* 12: 996–1006
- Kigami D, Minami N, Takayama H, Imai H (2003) MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biol Reprod* 68: 651–654
- Kirchhof N, Carnwath JW, Lemme E, Anastassiadis K, Scholer H, Niemann H (2000) Expression pattern of Oct-4 in preimplantation embryos of different species. *Biol Reprod* 63: 1698–1705
- Kobolak J, Kiss K, Polgar Z, Mamo S, Rogel-Gaillard C, Tancos Z, Bock I, Baji AG, Tar K, Purity MK (2009) Promoter analysis of the rabbit POU5F1 gene and its expression in preimplantation stage embryos. *BMC Mol Biol* 10: 1–12
- Krueger F, Andrews SR (2011) Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. *Bioinformatics* 27: 1571–1572
- Lea RA, McCarthy A, Boeing S, Fallesen T, Elder K, Snell P, Christie L, Adkins S, Shaikly V, Taranissi M (2021) KLF17 promotes human naïve pluripotency but is not required for its establishment. *Development* 148: dev199378
- Lerat E, Fablet M, Modolo L, Lopez-Maestre H, Vieira C (2017) TTools facilitates big data expression analysis of transposable elements and reveals an antagonism between their activity and that of piRNA genes. *Nucleic Acids Res* 45: e17
- Li H, Durbin R (2010) Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics* 26: 589–595
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup 1000 Genome Project Data Processing (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078–2079
- Li L, Guo F, Gao Y, Ren Y, Yuan P, Yan L, Li R, Lian Y, Li J, Hu B *et al* (2018) Single-cell multi-omics sequencing of human early embryos. *Nat Cell Biol* 20: 847–858
- Li Z, Schulz MH, Look T, Begemann M, Zenke M, Costa IG (2019) Identification of transcription factor binding sites using ATAC-seq. *Genome Biol* 20: 45
- Liu X, Wang C, Liu W, Li J, Li C, Kou X, Chen J, Zhao Y, Gao H, Wang H *et al* (2016) Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* 537: 558–562
- 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
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 1–21
- Lu F, Liu Y, Inoue A, Suzuki T, Zhao K, Zhang Y (2016) Establishing chromatin regulatory landscape during mouse preimplantation development. *Cell* 165: 1375–1388
- Lu X, Zhang Y, Xia W (2020) Gene Expression Omnibus GSE163620 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE163620>). [DATASET]
- Lu X, Zhang Y, Wang L, Wang L, Wang H, Xu Q, Xiang Y, Chen C, Kong F, Xia W (2021) Evolutionary epigenomic analyses in mammalian early embryos reveal species-specific innovations and conserved principles of imprinting. *Sci Adv* 7: eabi6178
- Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL (2012) Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* 487: 57–63
- Manes C (1973) The participation of the embryonic genome during early cleavage in the rabbit. *Dev Biol* 32: 453–459
- Markov GJ, Mai T, Nair S, Shcherbina A, Wang YX, Burns DM, Kundaje A, Blau HM (2021) AP-1 is a temporally regulated dual gatekeeper of reprogramming to pluripotency. *Proc Natl Acad Sci USA* 118: e2104841118
- Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17: 10–12
- Matsumura Y, Nakaki R, Inagaki T, Yoshida A, Kano Y, Kimura H, Tanaka T, Tsutsumi S, Nakao M, Doi T (2015) H3K4/H3K9me3 bivalent chromatin domains targeted by lineage-specific DNA methylation pauses adipocyte differentiation. *Mol Cell* 60: 584–596
- Memili E, First NL (1998) Developmental changes in RNA polymerase II in bovine oocytes, early embryos, and effect of alpha-amanitin on embryo development. *Mol Reprod Dev* 51: 381–389
- Ming H, Sun J, Pasquariello R, Gatenby L, Herrick JR, Yuan Y, Pinto CR, Bondioli KR, Krisher RL, Jiang Z (2021) The landscape of accessible chromatin in bovine oocytes and early embryos. *Epigenetics* 16: 300–312
- Navarro M, Halstead MM, Rincon G, Mutto AA, Ross PJ (2022) bESC from cloned embryos do not retain transcriptomic or epigenetic memory from somatic donor cells. *Reproduction* 164: 243–257
- Niakan KK, Han J, Pedersen RA, Simon C, Pera RAR (2012) Human pre-implantation embryo development. *Development* 139: 829–841
- Nishikimi A, Mukai J, Yamada M (1999) Nuclear translocation of nuclear factor kappa B in early 1-cell mouse embryos. *Biol Reprod* 60: 1536–1541
- Ooga M, Fulka H, Hashimoto S, Suzuki MG, Aoki F (2016) Analysis of chromatin structure in mouse preimplantation embryos by fluorescent recovery after photobleaching. *Epigenetics* 11: 85–94

- Orsini MW (1962) Study of ovo-implantation in the hamster, rat, mouse, Guinea-pig and rabbit in cleared uterine tracts. *Reproduction* 3: 288–293
- Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D, Knowles BB (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell* 7: 597–606
- Pérez-Gómez A, González-Brusi L, Bermejo-Álvarez P, Ramos-Ibeas P (2021) Lineage differentiation markers as a proxy for embryo viability in farm ungulates. *Front Vet Sci* 8: 680539
- Plusa B, Piliszek A, Frankenberg S, Artus J, Hadjantonakis A-K (2008) Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development* 135: 3081–3091
- Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841–842
- Ralston A, Cox BJ, Nishioka N, Sasaki H, Chea E, Rugg-Gunn P, Guo G, Robson P, Draper JS, Rossant J (2010) Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development* 137: 395–403
- Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* 42: W187–W191
- Roode M, Blair K, Snell P, Elder K, Marchant S, Smith A, Nichols J (2012) Human hypoblast formation is not dependent on FGF signalling. *Dev Biol* 361: 358–363
- Rose NR, Klose RJ (2014) Understanding the relationship between DNA methylation and histone lysine methylation. *Biochim Biophys Acta* 1839: 1362–1372
- Ross PJ, Sampaio RV (2018) Epigenetic remodeling in preimplantation embryos: cows are not big mice. *Anim Reprod* 15: 204–214
- Ross PJ, Ragina NP, Rodriguez RM, Lager AE, Siripattarapavat K, Lopez-Corralles N, Cibelli JB (2008) Polycomb gene expression and histone H3 lysine 27 trimethylation changes during bovine preimplantation development. *Reproduction* 136: 777–785
- Schultz RM, Stein P, Svoboda P (2018) The oocyte-to-embryo transition in mouse: past, present, and future. *Biol Reprod* 99: 160–174
- Schulz KN, Harrison MM (2019) Mechanisms regulating zygotic genome activation. *Nat Rev Genet* 20: 221–234
- Simmet K, Zakhartchenko V, Wolf E (2018) Comparative aspects of early lineage specification events in mammalian embryos—insights from reverse genetics studies. *Cell Cycle* 17: 1688–1695
- Skene PJ, Henikoff S (2017) CUT&RUN: Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *bioRxiv* <https://doi.org/10.1101/193219> [PREPRINT]
- Stovner EB, Sætrum P (2019) epic2 efficiently finds diffuse domains in ChIP-seq data. *Bioinformatics* 35: 4392–4393
- Tardat M, Albert M, Kunzmann R, Liu Z, Kaustov L, Thierry R, Duan S, Brykczynska U, Arrowsmith CH, Peters AHFM (2015) Cbx2 targets PRC1 to constitutive heterochromatin in mouse zygotes in a parent-of-origin-dependent manner. *Mol Cell* 58: 157–171
- van Eijk MJT, van Rooijen MA, Modina S, Scesi L, Folkers G, van Tol HTA, Bevers MM, Fisher SR, Lewin HA, Rakacolli D (1999) Molecular cloning, genetic mapping, and developmental expression of bovine POU5F1. *Biol Reprod* 60: 1093–1103
- Vuoristo S, Bhagat S, Hydén-Granskog C, Yoshihara M, Gawrylski L, Jouhilahti E-M, Ranga V, Tamirat M, Huhtala M, Kirjanov I et al (2022) DUX4 is a multifunctional factor priming human embryonic genome activation. *iScience* 25: 104137
- Wang C, Liu X, Gao Y, Yang L, Li C, Liu W, Chen C, Kou X, Zhao Y, Chen J et al (2018) Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat Cell Biol* 20: 620–631
- Wiekowski M, Miranda M, DePamphilis ML (1993) Requirements for promoter activity in mouse oocytes and embryos distinguish paternal pronuclei from maternal and zygotic nuclei. *Dev Biol* 159: 366–378
- Wu J, Huang B, Chen H, Yin Q, Liu Y, Xiang Y, Zhang B, Liu B, Wang Q, Xia W et al (2016) The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* 534: 652–657
- Wu J, Xu J, Liu B, Yao G, Wang P, Lin Z, Huang B, Wang X, Li T, Shi S et al (2018) Chromatin analysis in human early development reveals epigenetic transition during ZGA. *Nature* 557: 256–260
- Xia W, Xu J, Yu G, Yao G, Xu K, Ma X, Zhang N, Liu B, Li T, Lin Z (2019) Resetting histone modifications during human parental-to-zygotic transition. *Science* 365: 353–360
- Xu R, Li S, Wu Q, Li C, Jiang M, Guo L, Chen M, Yang L, Dong X, Wang H (2022) Stage-specific H3K9me3 occupancy ensures retrotransposon silencing in human pre-implantation embryos. *Cell Stem Cell* 29: 1051–1066
- Yeo J-C, Jiang J, Tan Z-Y, Yim G-R, Ng J-H, Göke J, Kraus P, Liang H, Gonzales KAU, Chong H-C (2014) Klf2 is an essential factor that sustains ground state pluripotency. *Cell Stem Cell* 14: 864–872
- Yu H, Chen M, Hu Y, Ou S, Yu X, Liang S, Li N, Yang M, Kong X, Sun C (2022) Dynamic reprogramming of H3K9me3 at hominoid-specific retrotransposons during human preimplantation development. *Cell Stem Cell* 29: 1031–1050
- Zaret KS (2020) Pioneer transcription factors initiating gene network changes. *Annu Rev Genet* 54: 367–385
- Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al (2008) Model-based analysis of ChIP-seq (MACS). *Genome Biol* 9: R137
- Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, Ming J, Wu X, Zhang Y, Xu Q et al (2016) Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* 537: 553–557
- Zhang W, Chen Z, Yin Q, Zhang D, Racowsky C, Zhang Y (2019) Maternal-biased H3K27me3 correlates with paternal-specific gene expression in the human morula. *Genes Dev* 33: 382–387
- Zheng H, Huang B, Zhang B, Xiang Y, Du Z, Xu Q, Li Y, Wang Q, Ma J, Peng X et al (2016) Resetting epigenetic memory by reprogramming of histone modifications in mammals. *Mol Cell* 63: 1066–1079
- Zhong H, May MJ, Jimi E, Ghosh S (2002) The phosphorylation status of nuclear NF- κ B determines its association with CBP/p300 or HDAC-1. *Mol Cell* 9: 625–636
- Zhou C, Wang Y, Zhang J, Su J, An Q, Liu X, Zhang M, Wang Y, Liu J, Zhang Y (2019) H3K27me3 is an epigenetic barrier while KDM6A overexpression improves nuclear reprogramming efficiency. *FASEB J* 33: 4638–4652