Chromatin remodeling in bovine embryos indicates species-specific regulation of genome activation

Michelle M Halstead[1], Xin Ma[1,2], Richard M Schultz[3,4,†], and Pablo J Ross[1. †] *

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

[2] College of Animal Science and Technology, Jilin Agricultural University, Changchun, China

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

[4] Department of Biology, University of Pennsylvania, Philadelphia, PA, USA

* Correspondence: Department of Animal Science, University of California, 2239 Meyer Hall, Davis, CA 95616, USA. Tel: +530-771-7225; E-mail: pross@ucdavis.edu

† Grant support: NIH-NICHD R01HD070044

Running Title: Open chromatin in mammalian embryos

Key words: chromatin accessibility / cow / genome activation / maternal-to-zygotic transition / preimplantation embryo

Abstract

The maternal-to-zygotic transition (MZT) is underpinned by wide-spread transcriptomic and epigenomic remodeling that facilitates totipotency acquisition. Factors regulating MZT vary across species and differences in timing of developmental transitions and motif enrichment at accessible chromatin between human and mouse embryos suggest a distinct regulatory circuitry. Profiling accessible chromatin in bovine preimplantation embryos—timing of developmental transitions in bovine closely resembles that in human—indicated that prior to embryonic genome activation (EGA) accessible chromatin is enriched in maternal transcription factor recognition sites, e.g., CTCF, KLFs, NFY, and SP1, echoing observations in humans and mice, and suggesting that a conserved set of maternal factors regulate chromatin remodeling prior to EGA. In contrast, open chromatin established during EGA was primarily enriched for homeobox motifs and showed remarkable similarities between cattle and humans, indicating that cattle could be a more relevant model for human preimplantation development than mice.

Introduction

Preimplantation development encompasses several critical milestones as embryos progress from fertilization to blastocyst formation. Fusion of the transcriptionally quiescent oocyte and sperm results in a zygote with two haploid pronuclei, which combine during the first round of replication when pronuclear membranes dissolve, allowing maternal and paternal chromosomes to intermingle on the metaphase plate. Subsequent rounds of cleavage ultimately form a blastocyst. However, the cleavage-stage embryo must first complete the maternal-to-zygotic transition (MZT), wherein the embryo assumes control over its own continued development by degrading oocyte-derived products and initiating its own transcriptional program. This dramatic change in gene expression proceeds gradually; minor embryonic genome activation (EGA) results in low levels of transcription in early cleavage-stage embryos[1], and leads to major EGA, which involves wide-spread activation of embryonic transcription[2]. This shift from maternal dependence to self-sufficiency serves at least three functions: elimination of oocyte-specific messages,

replenishment of transcripts that are common to both the oocyte and the embryo, and generation of novel embryonic-specific transcripts.

The MZT also involves wide-spread changes in chromatin structure[3] and other epigenetic marks, which completely restructures the embryonic epigenome. This chromatin remodeling is necessary to eradicate gamete-specific signatures and establish an open chromatin landscape that supports embryonic transcriptional programs. Specifically, chromatin structure defines the genomic context within which transcriptional machinery can operate, thereby determining the cell-specific gene expression patterns that confer cell identity and function. Following fertilization, the zygotic genome is globally demethylated[4], and this loss of DNA methylation coincides with global decreases in repressive histone modifications[3,5]. Generally, epigenetic factors linked to relaxed chromatin are more abundant in mouse zygotes, whereas factors implicated in chromatin compaction become more prevalent during EGA[6], pointing to a highly permissive chromatin landscape in pre-EGA embryos. Indeed, mouse zygotes demonstrate elevated histone mobility[7], highly dispersed chromatin[8], and lack chromocenters (congregations of pericentromeric heterochromatin)[9,10]. Moreover, in mouse, 3-D chromatin architecture is largely absent after fertilization, and is then gradually established throughout preimplantation development[11]. This increasing chromatin compaction and organization facilitates long-distance chromatin interactions in later stage embryos[11]. The co-occurrence of genome activation and chromatin remodeling raises an interesting causality dilemma, namely, whether chromatin remodeling is necessary for transcription activation or whether transcription activation leads to chromatin accessibility. Several maternal products prompt transcription initiation by altering chromatin structure[12–14], and some chromatin compaction occurs in the absence of embryonic transcription[11]. However, inhibiting embryonic transcription pervasively disrupts the establishment of open chromatin during EGA[15,16], suggesting that EGA and chromatin remodeling are likely interdependent.

In mouse and human preimplantation embryos, accessible sites are gradually established as development progresses[15–20]; however, these open regions demonstrate different motif enrichment patterns, implicating distinct sets of transcription factors (TFs) in either murine (RARG, NR5A2, ESRRB),[15] or human EGA (OTX2, GSC, POU5F1)[16,18]. Although some TFs appear to regulate EGA in multiple species, i.e. KLFs[15,16,21], DUX[22–25], ZSCAN4[26,27], and CTCF[15,16,28], it remains unclear if there is any mechanistic conservation across mammals. In fact, the timing of genome activation is highly species-specific: major EGA in mice occurs during the 2-cell stage[1], in humans[29] and pigs[30] during the 4- to-8-cell stage, and in sheep[31] and cattle[32] between the 8- and 16-cell stages. The relative timing with which mice activate their genomes could indicate that the mechanism behind murine EGA differs significantly from other species', which would have significant implications for modeling human preimplantation development. In particular, the timing of bovine EGA more closely resembles that of human EGA, as do global changes to histone PTMs: the active mark trimethylation of lysine 4 on histone 3 decreases in global abundance during human[33] and bovine EGA[34], but increases during murine EGA[35,36]. However, the chromatin remodeling events that underscore bovine preimplantation development have yet to be catalogued, and it remains unclear whether the regulation and execution of the MZT in cattle resembles that which has been observed in humans.

To this end, we here describe the chromatin accessibility landscapes of bovine oocytes and preimplantation embryos using the Assay for Transposase Accessible Chromatin (ATAC-seq)[37]. We find that open chromatin is gained progressively throughout development, with promoters enriched for CTCF and NFY motifs gaining accessibility in earlier stages, and distal regions becoming accessible at later stages. Moreover, embryonic transcription was not required for the appearance of promoter open chromatin in 2- and 4-cell embryos, but was absolutely necessary to establish stage-specific and distal open chromatin, especially in 8-cell embryos, indicating that maternal and embryonic products both participate in chromatin remodeling in a complementary fashion. Sequence enrichment in open chromatin revealed that several TFs likely play roles in both bovine and human EGA (OTX2, SP1), whereas

regulators specific to murine EGA[15] demonstrate no enrichment in bovine (RARG, NR5A2, ESRRB), suggesting that cattle may be a more informative model for human preimplantation development. Nevertheless, several TFs (DUX, KLFs, CTCF, NFY) seem to play a role in preimplantation development in all three species, raising the possibility that events leading to EGA may be mechanistically conserved across mammals, whereas the specific transcriptional programs that are activated may differ between species.

Results and Discussion

Global dynamics of open chromatin in bovine preimplantation embryos

For each developmental stage, ATAC-seq libraries were prepared from cells derived from three separate oocyte collections. A subset of embryos from each collection was also cultured in the presence of a transcriptional inhibitor (α-amanitin), to interrogate the causal relationship between embryonic transcription and chromatin remodeling (Figure 1a). Between 30 and 87 million non-mitochondrial monoclonal uniquely mapping reads were collectively obtained for each developmental stage, and at least 20 million non-mitochondrial monoclonal reads were collectively obtained for transcription blocked embryos (TBEs) at each stage (Table S1). Genome-wide normalized ATAC-seq coverage demonstrated a striking shift in the open chromatin landscape between 4- and 8-cell embryos (Figure 1b), suggesting that large-scale chromatin remodeling coincided with the main transcriptomic shift observed during major EGA (Figure S1). Similarity between replicates (Table S2) indicated that both the technique and embryo production were robust, generating comparable chromatin accessibility profiles across different rounds of oocyte collection and embryo production (Figure 1c). Reads from replicates were pooled together to obtain greater sequencing depth and maximize power for identifying regions of open chromatin. To gauge changes in chromatin accessibility throughout development, regions of open chromatin, or peaks, were called for each stage of development. To minimize bias from sequencing depth, peaks were called from either 30 million monoclonal uniquely mapped reads when comparing different developmental stages, or 20 million reads when comparing TBE with control embryos.

As observed in humans and mice[15,16,18–20], regions of open chromatin were gradually established throughout bovine preimplantation development, with the lowest enrichment for accessible sites in 2-cell embryos (Figure 1d). Rather than reflecting chromatin inaccessibility in 2-cell embryos, this dearth of canonical "open chromatin" probably results from a highly permissive chromatin structure. Indeed, chromocenters are absent from bovine embryos until the early 8-cell stage[9], indicating a relaxed chromatin configuration in early cleavage-stage embryos. Because assays that employ endonucleases, e.g., ATAC-seq and DNase-seq, depend on increased cutting events at consistently accessible loci, we speculate that genome-wide chromatin relaxation would lead to random cutting events genome-wide, resulting in the observed high background and low enrichment in 2-cell embryos. Attempts to use endonuclease-based methods to profile open chromatin in mouse zygotes[15,20] and human 2-cell embryos[18,19] have encountered similar difficulties with low enrichment. As in bovine embryos, chromocenters are also conspicuously absent in pre-EGA mouse embryos[9,10], and electron microscopy[8] and fluorescence recovery after photobleaching analysis[7] also indicate highly dispersed chromatin. Furthermore, a recent study that detected open chromatin based on methylation of accessible GpC sites, rather than endonucleases, found that genome-wide accessibility in human embryos actually decreased from the zygote stage onward[17]. Thus, global chromatin relaxation appears to be a shared characteristic of human, bovine, and murine pre-EGA embryos.

It is tempting to speculate that this "naïve" chromatin state acts as a blank epigenetic slate, which is then gradually compacted and structured to meet the needs of the growing embryo. Indeed, accessible sites were progressively established in 4-cell, 8-cell, and morula stage embryos (Figure 1d). Interestingly, many of these regions were only open transiently at a specific stage, whereas others maintained their accessibility throughout later stages (Figure 1e), suggesting that chromatin remodeling serves two functions: progressive establishment of a "totipotent" chromatin landscape, and transient stage-specific regulation.

Stepwise remodeling yields stage-specific open chromatin with distinct functionality

To delve further into the potential function of regions that lost or gained accessibility between consecutive stages (Figure 2a), intergenic regions of open chromatin were evaluated for sequence enrichment. Although regulatory regions have not been annotated in cattle, intergenic open chromatin could correspond to enhancers, the activity of which is generally highly tissue-specific. Indeed, distal regions that became accessible at each developmental stage were enriched for different recognition motifs (Table S3). These enriched sequences corresponded to the known binding motifs of several TFs that were either maternally provided or expressed in embryos (Figure S2), demonstrating that the changing chromatin structure subjects each stage of development to distinct regulatory circuitry (Figure 2b).

The first major transition in chromatin structure mostly involved loss of hyperaccessible sites in 2-cell embryos following fertilization (Figure 2a). Intergenic loci that lost accessibility during this transition (n=54,264) were enriched for the binding motifs of CTCF (an insulator protein implicated in 3D chromatin organization[38]), FIGLA (an oocyte-specific TF), and RFX factors (a highly conserved family of transcriptional repressors[39]; Figure 2b). Interestingly, one quarter of the intergenic regions that closed in 2-cell embryos regained accessibility at the 4-cell stage (n=14,859), and many of these remained open all the way through to the morula stage (n=6,186). Most intergenic regions that closed after fertilization remained inaccessible in embryos, suggesting that they could contribute to oocyte-specific regulation. On the other hand, regions that re-opened in 4-cell embryos seem more likely to participate in house-keeping functions.

The first major gains in accessibility occurred in 4-cell embryos (Figure 2a). About a quarter of the regions that opened during the 4-cell stage remained accessible in both 8-cell embryos and morulae (Figure 1e), although 74% of these had been previously open in GV oocytes (n=15,466). Surprisingly, nearly half of the regions that opened during the 4-cell stage were only transiently accessible (Figure 1e). This 4-cell-specific open chromatin was most significantly enriched for binding motifs of NFkB family

171 members (Figure 2c), which was particularly intriguing because NFkB activation is specifically necessary
172 in 1-cell mouse embryos for development to progress past EGA[40,41]. Although NFkB factors are
173 maternally provided (Figure 2c), they are initially sequestered in the cytoplasm until they translocate into
174 the nucleus at the early 1-cell stage in mice[40,41] and the 4-to-8 cell stage in cattle[42]. In particular, one of
175 the NFkB subunits capable of activating gene expression – RELA – binds DNA with increased frequency
176 in bovine embryos compared to oocytes[42], suggesting that NFkB activation of target gene expression
177 could be one of the first events in a cascade leading to major EGA. In fact, one of the few genes with
178 upregulated expression in 4-cell embryos, as compared to MII oocytes, was TRIM8 – a positive regulator
179 of NFkB activity (Figure S3). Additionally, RELA binding sites that were accessible in 4- and 8-cell
180 embryos mark genes that encode key regulators of early preimplantation development, raising the
181 possibility that NFkB binding is involved in their transcription initiation (Figure S4), although the
182 contribution of NFkB signaling to gene expression during minor EGA in cattle has not yet been
183 established.
184
185 In contrast to open chromatin in pre-EGA embryos, regions that gained accessibility at the 8-cell stage
186 (Figure 2a) were considerably enriched for the binding motifs of several homeobox TFs, including OTX2,
187 GSC, CRX, PHOX2A, PAX7, and PITX1 (Figure 2b,e), although only OTX1, OTX2, and PITX2 were
188 appreciably expressed during the 8-cell stage (Figure 2f). Most notably, more than a third of the
189 accessible intergenic loci in 8-cell embryos harbored a sequence that most closely matched DUX binding
190 motifs (Figure 2g). The DUX homeoboxes have been extensively implicated in EGA regulation[22–25], and
191 DUXA is expressed transiently and strongly in bovine embryos during EGA (Figure 2h). The
192 synchronized transcription of DUXA and increased accessibility of its binding sites during bovine
193 development strongly indicates a conserved role for DUX in mammalian preimplantation development,
194 especially considering that DUX family members are highly conserved and specific to placental
195 mammals[43,44]. The factors that regulate DUXA expression in cattle remain unclear. Although maternally
196 provided DPPA2 and DPPA4 induce Dux in mouse embryonic stem cells[45], only DPPA3 is maternally

provided in cattle (Figure S5), and it is unknown if DPPA3 is capable of inducing DUXA expression; however, DPPA3 is maternally provided in mice, cattle, and humans[46,47], and knockdown of DPPA3 decreases the developmental competency of mouse[48] and bovine embryos[47], strongly suggesting that it could activate DUXA expression. Further functional validation will be necessary to determine if DUXA is required for bovine embryogenesis, or if is an important but non-essential regulator of EGA, as in mouse[49–51].

The expression profile of DUXA in bovine preimplantation development closely mirrored that of another TF implicated in EGA: ZSCAN4 (Figure S6a)[52]. ZSCAN4 is a downstream target of DUX factors in humans[23] and mice[22,24], and the coordinated expression of these two factors in cattle certainly suggests a similar mechanism may be at play. Although ZSCAN4 depletion disrupts development past EGA in mice[26] and cattle[27], its known binding motifs were not enriched in open chromatin during bovine preimplantation development (Figure S6b-d), suggesting that the binding motif of bovine ZSCAN4 likely differs from those in mice and humans.

Following major EGA, formation of the morula also incurred extensive chromatin remodeling, with even more sites gaining accessibility than during the 8-cell stage (Figure 2a). Compared to earlier stages, intergenic loci that gained accessibility in morulae were primarily enriched for GATA factor binding motifs (Figure 2a; Table S3) – key regulators of trophectoderm establishment and maintenance[53] – indicating that morulae are already initiating differentiation programs necessary for blastocyst formation.

Progressive establishment of maintained open chromatin sets the stage for genome activation

Although many regions only experienced stage-specific gains in accessibility, a stable open chromatin landscape was also progressively established after fertilization. As early as the 2-cell stage, regions began to gain accessibility that was maintained until at least the morula stage. These regions of maintained accessibility were heavily enriched for CTCF motifs (Figure 3a; Table S4b), especially those that were

first established during the 4-cell stage. CTCF binding delineates chromatin loop boundaries, thus determining the genomic space within which genes can interact with their regulatory elements[38]. Therefore, enrichment of CTCF motifs in maintained peaks could point to a gradual establishment of a stable 3-D chromatin architecture in preparation for major EGA. This proposal is consistent with 3-D chromatin architecture in mouse embryos, which is greatly diminished after fertilization and then gradually re-established throughout preimplantation development, facilitating long-distance chromatin contacts, i.e., promoter-enhancer interactions, in later stage embryos[11]. Indeed, transcription during minor EGA in mouse is primarily driven by proximal promoters, whereas enhancers are dispensable for transcription until major EGA[54]. The global reorganization of 3D chromatin architecture in mouse has also been observed in bovine embryos, wherein gene-rich regions switch from a random distribution to a chromosome-specific distribution during major EGA[55]. Collectively, these results suggest that gradual establishment of 3D chromatin architecture is a conserved feature of pre-EGA embryos, although the mechanisms regulating this restructuring remain unknown.

Other than CTCF, maintained peaks were also more enriched for KLF motifs than transiently open sites (Figure 3b). Although most KLFs were not expressed until the 8-cell stage, several KLFs were maternally provided (Figure 3c), including KLF4, a master regulator of induced pluripotent stem cell (iPSC) reprogramming[56]. Recent evidence suggests that KLF4 contributes to reprogramming iPSC by mediating pluripotency-associated enhancer-promoter contacts[57]; thus, the concurrent establishment of open chromatin corresponding to CTCF and KLF binding motifs suggests that pre-EGA embryos are priming a similar mechanism for use during major EGA, especially considering that almost 50% of genes activated during bovine EGA contain KLF motifs in their promoters[21].

In fact, loci that remained open from the 2- and 4-cell stages onward occurred in genic regions more often than stage-specific open chromatin (Figure 3d) and marked the promoters of genes that were functionally enriched for housekeeping functions (Figure 3e), including 33 of the 51 genes that were upregulated in 4-

cell embryos compared to MII oocytes (Figure S3). Moreover, these regions of maintained accessibility were strongly enriched for NFY and SP1 binding sites (Table S4a,b), which is highly reminiscent of chromatin remodeling in mouse embryos, as proximal promoters enriched for NFY are the first regions to gain accessibility[20]. Of note, NFY enhances binding of the pluripotency factors POU5F1, SOX2, and KL4 to their recognition motifs[58,59], and is clearly involved in murine EGA, as NFY knockdown embryos demonstrated impaired open chromatin establishment and downregulation of gene expression[20]. Similarly, SP1 is essential for early mouse development, with knockout embryos arresting at day 11 of gestation[60]. Intriguingly, human zygotes also demonstrate KLF, SP1, and NFY motif enrichment in open chromatin[19], overall suggesting that a conserved set of maternal regulators (CTCF, KLFs, SP1, NFY) participates in chromatin remodeling and transcription activation in pre-EGA embryos.

Both maternal products and embryonic transcription drive chromatin reorganization

Considering that embryonic transcription is extremely limited before the 8-cell stage in cattle, the appearance of open chromatin in pre-EGA embryos suggests that maternal factors participate in chromatin remodeling. To further dissect the maternal contribution to epigenetic reprogramming and EGA, embryos were cultured in the presence of α-amanitin to inhibit POLR2-dependent transcription elongation. Loss of embryonic transcription had a drastic inhibitory effect on the appearance of open chromatin that intensified as development progressed (Figure 4a); 64% of loci that should have opened during the 4-cell stage failed to become accessible without embryonic transcription, and in 8-cell TBE embryos, 96% of loci that should have opened remained closed (Figure 4b), disrupting the chromatin state of key genes, such as KLF4 (Figure 4c), and coinciding with developmental arrest.

Interestingly, inhibiting embryonic transcription did not uniformly affect chromatin remodeling genome-wide. Stage-specific open chromatin was preferentially disrupted in TBEs, whereas maintained open chromatin established in 2- and 4-cell embryos appeared even in the absence of embryonic transcription (Figure 4d). Transcription-independent maintained open chromatin marked the promoters of nearly 60%

of embryonically-expressed genes (n=1,038/1,784 genes identified from data from Bogliotti *et al* 2019; Figure S7), suggesting that maternal factors actively remodel the local chromatin structure of target genes, possibly priming them for expression later on in development. Nevertheless, the appearance of stage-specific open chromatin almost completely depended on embryonic transcription, indicating that maternal and embryonic factors cooperate in a complementary fashion to establish the appropriate chromatin landscape for activation of embryonic transcriptional programs.

Open chromatin in preimplantation embryos is enriched for repetitive elements

Evaluation of changes in chromatin structure as they relate to gene expression gives an incomplete perspective of the genome-wide changes that occur during preimplantation development, because the MZT is not just characterized by a shift in the transcriptome but also in the repeatome. Similar to other mammalian species, 44% of the bovine genome is comprised of repeats derived from retrotransposons[61] – interspersed repeats that are increasingly thought to play major roles in cellular processes and development. Retrotransposons propagate through a 'copy and paste' mechanism, and their expression is generally suppressed to avoid deleterious integrations[62]. However, retrotransposons are often actively transcribed in early embryos, and although this phenomenon was recently thought to be nothing more than opportunistic expression by repetitive elements due to an unusually permissive chromatin state in the developing embryo, the activity of some retrotransposons is actually crucial for development[63]. Although the specific mechanisms behind this necessity are still being investigated, transposable elements have been implicated at several regulatory levels, as they can provide binding sites for TFs, allowing them to be co-opted as alternative promoters and enhancers[64] and participate in 3-D chromatin architecture[65].

Although repetitive element expression in bovine embryos was reported a decade ago using a cDNA array[66], a complete catalogue of repeat transcription throughout bovine preimplantation development was lacking. To address this gap in knowledge, available RNA-seq data[67] were assessed for expression of repetitive elements. Importantly, these libraries were not subjected to polyA selection. As has been

observed in the mouse and human[68,69], expression and accessibility of repetitive elements throughout bovine preimplantation development was highly stage-specific and dynamic (Figure 5a; Figure S8).

Non-long terminal repeat (non-LTR) retrotransposons – long interspersed elements (LINEs) and short interspersed elements (SINEs) – are increasingly transcribed during human preimplantation development[17,68]. In cattle, LINEs also demonstrated increasing expression (Figure 5a) and accessibility (Figure 5b) starting during the 8-cell stage. Although the function of LINE elements in bovine preimplantation embryos has yet to be established, their activation is crucial for mouse development: perturbing LINE expression in mouse preimplantation embryos causes developmental arrest at the 2-cell stage and perturbs gene expression[70]. Moreover, LINE-1 activation may regulate global chromatin accessibility in mouse embryos[71]. Of the SINE families, mammalian-wide interspersed repeats (MIR) expression and accessibility patterns most echoed those of L2 LINE elements, with increased expression and accessibility starting at the 8-cell stage, suggesting that these elements could be acting as enhancers or promoters.

In particular, LTR activation is a key feature in human[68], mouse[69], and bovine preimplantation development[66]. Of these, mammalian LTRs and endogenous retroviral elements (ERVL) were increasingly enriched in open chromatin starting during the 8-cell stage (Figure 5b), although their transcript abundance dropped throughout development (Figure 5a), indicating that other mechanisms likely regulate repeat expression, e.g., DNA methylation or histone modifications[72,73]. Only endogenous retroviral K elements (ERVK) demonstrated both increasing expression (Figure 5a) and moderately increased prevalence in distal open chromatin during the 8-cell stage (Figure 5b), suggesting that ERVK elements function as regulatory elements, as observed in human preimplantation embryos[74].

Mounting evidence suggests that specific types of LTR retrotransposons, especially intact elements, play pivotal roles in early development. During bovine preimplantation development, the most expressed

retrotransposons were ERV1-1_BT and ERV1-2_BT (Figure 6a), as has been observed previously[66]. Upon further inspection, intact ERV1-1_BT elements often co-occurred with MER41_BT repeats in a specific sequence, which demonstrated a highly reproducible pattern of transcription at ERV1-1_BT elements and chromatin accessibility at MER41_BT elements (Figure S9a). Furthermore, MER41_BT elements that were accessible in 8-cell embryos were enriched for the binding motifs of several pluripotency factors, including POU5F1, NFY, KLF4, OTX2, and TEAD (Figure S9b), suggesting that pluripotency factors are driving transposon expression.

DUX has also been implicated in driving the expression of intact ERVL elements in human and mouse pre-EGA embryos[19,22–24]. Specifically, in human embryos DUX4 appears to bind MLT2A1 elements – primate-specific LTRs that flank human ERVL –activating their expression[19]. Considering that intergenic open chromatin was especially enriched for DUX binding sites in 8-cell bovine embryos, it seemed likely that these sites would also correspond to retrotransposons. Indeed, several MLT elements were enriched in 8-cell open chromatin harboring DUX motifs (Table S5), suggesting that bovine DUX may bind and regulate LTRs with sequence similarity to primate-specific MLT2A1 elements. Nearly all LTRs that were enriched in 8-cell open chromatin with DUX binding sites demonstrated dynamic expression profiles throughout development (Figure S10). In particular, increasing accessibility at MLT1A0 elements harboring DUX recognition sites (Figure 6c) was mirrored by a sharp increase in transcription in 8-cell embryos (Figure 6d).

Whether the transcripts derived from LTRs are required for bovine development or are simply a result of opportunistic expression remains to be established. However, evidence in other species suggests that retrotransposon-derived regulatory elements are often co-opted by the embryo as promoters and enhancers[64]: a phenomenon which appears to extend to bovine embryos, as ATAC-seq and RNA-seq suggest that MLT1A0 elements are co-opted as alternative promoters at several loci, including CD1 (Figure 6e), ZNF41, and LPIN2. As such, an interesting balance appears to exist between repetitive

353 elements and the embryo, wherein the repeatome leverages the embryo's existing regulatory network to
354 drive transposition, while simultaneously providing new regulatory elements and TF binding sites that the
355 embryo co-opts to drive the expression of its own transcriptional program.
356
357 **A model for mammalian genome activation**
358 Several lines of evidence suggest that the regulatory circuitry responsible for the MZT may differ
359 between mammals. First, EGA occurs in a highly species-specific fashion, with the major wave of
360 transcription in mice during the 2-cell stage[1], in humans during the 4-to-8-cell stage[29], and cattle during
361 the 8-to-16-cell stage[32]. Second, the expression patterns of repetitive elements are not only highly-stage
362 specific, but also species-specific; primate-specific and murine ERVL elements are strongly expressed
363 during human and mouse preimplantation development, respectively[19,22–24], whereas ERV1 elements were
364 most prominently activated in bovine embryos. Finally, the maternal programs in human and mice are
365 divergent; human maternal programs conspicuously lack the murine maternal effect transcripts POU5F1,
366 HSF1, and DICER1, and are functionally enriched for translational processing[46], reflecting the need to
367 translate maternal messages during the extended period between fertilization and human EGA.
368
369 The relaxed chromatin structure in early preimplantation embryos provides a unique regulatory context
370 for maternal factors, which are essential to support cleavage stage-embryos prior to genome activation,
371 especially in species where EGA is delayed for several cell divisions, e.g., humans and cattle. Whereas
372 the condensed chromatin structure in somatic cells generally restricts DNA-binding proteins to regions of
373 open chromatin, the dispersed chromatin in 2-cell bovine embryos may allow maternal factors to
374 opportunistically and pervasively bind their recognition motifs. Indeed, several maternal factors appear to
375 participate in chromatin restructuring in pre-EGA embryos, leading to open chromatin establishment at
376 promoters enriched for NFY and SP1 binding sites, as well as CTCF and KLF motifs. Although not
377 maternally provided, the DUX family also appears to play a conserved role in genome activation. In
378 humans and mice, DUX factors have been implicated in chromatin remodeling[23,25] and transcription

activation of cleavage-stage genes[23]. We find a similar pattern of DUX expression in bovine embryos, as well as increased accessibility of DUX binding sites around the 8-cell stage, suggesting that DUX may also modulate gene expression and chromatin accessibility during bovine EGA.

Although the chromatin landscape changes markedly upon major EGA in bovine, human[16,17,19], and mouse[15,20] embryos, the regulatory circuitry that is active during this stage in development appears to differ significantly between humans and mice[16], which suggests that regulation of mammalian EGA is highly species-specific. To identify and compare putative regulators of mammalian EGA, intergenic regions that were accessible during major EGA in bovine, human, and mouse embryos were evaluated for binding motif enrichment of actively expressed TFs. Comparing EGA regulatory circuitry between species reflects a stark divergence in regulatory and transcriptional programs that clearly separates humans and cattle from the mouse (Figure S11). Compared to mouse 2-cell embryos, open chromatin in bovine and human 8-cell embryos demonstrated remarkably similar patterns of sequence enrichment (Figure 7). In cattle and humans, SP1, OTX2, and NFY were particularly implicated during major EGA, whereas NR5A2, RARG, and ESRRB were solely enriched in mouse embryos.

Nevertheless, it is unclear if the regulatory factors that are enriched in open chromatin during major EGA are regulators of EGA or simply products of it. For instance, although NR5A2 is enriched in 2-cell mouse embryos, it is an early regulator of inner cell mass and trophectoderm programs and is not essential for genome activation[15]. Similarly, OTX2 is essential for neuronal lineage specification in mice[75], and has been implicated in the transition from naïve to primed pluripotency[76]. However, OTX2 protein is clearly present in human and marmoset zygotes[46], suggesting that this homeobox TF may play an as-of-yet undetermined role in EGA. Although several TFs appeared to only be important in mouse, or in cattle and humans, KLFs were substantially enriched during EGA in all three species. Considering the well-established role of KLFs in somatic cell reprogramming and establishment of pluripotency in multiple species[56,77], KLFs may play a conserved role in the MZT. Although future research will be necessary to

405 elucidate the function of specific regulatory factors, the high consistency between cattle and humans, both
406 with respect to the timing of EGA and the regulatory circuitry that accompanies it, strongly suggests that
407 cattle are a more appropriate model system for human preimplantation development than mouse.
408
409 **Conclusions**
410 Sweeping changes to chromatin structure during bovine preimplantation development suggest that 2-cell
411 embryos are characterized by globally decondensed chromatin, which is gradually compacted as
412 development progresses, echoing similar observations in humans and mice. In particular, it is tempting to
413 speculate that a conserved set of maternal factors establish basal promoter accessibility and the necessary
414 chromatin architecture for enhancer-promoter interactions that will drive gene expression during major
415 EGA (Figure 8). However, the open chromatin landscape during major EGA clearly distinguished mice
416 from cattle and humans, suggesting that whereas maternal regulation of EGA may be conserved across
417 mammals, the transcriptional programs that are subsequently activated have diverged substantially.
418 Practically, this difference suggests that human development may be better modeled in cattle than in
419 mice. Nevertheless, the factors appear to regulate the MZT in cattle, humans, and mice certainly warrant
420 further investigation and validation, which will provide invaluable insight into the regulatory framework
421 that governs successful preimplantation development.
422
423 **Materials and Methods**
424 *Oocyte collection and maturation*
425 Ovaries were procured from a local abattoir and transported to the laboratory in a warm saline solution.
426 Follicles measuring 2-10 mm were aspirated to obtain cumulus oocyte complexes (COCs). Only COCs
427 with healthy layers of cumulus cells were selected for maturation. These were washed in collection
428 medium (6:4 M199 (Sigma M7653)): SOF-Hepes, supplemented with 2% fetal bovine serum (FBS;
429 Hyclone/Thermo Scientific) and transferred to maturation medium (modified M199 medium (Sigma

M2154)) supplemented with ALA-glutamine (0.1 mM), sodium pyruvate (0.2 mM), gentamicin (5 μg/ml), EGF (50 ng/ml), oFSH (50 ng/ml), bLH (3 μg/ml), cysteamine (0.1 mM), and 10% FBS.

In vitro fertilization and embryo culture

After COCs matured for 24 h, MII oocytes were washed in SOF-IVF medium (107.7 mM NaCl, 7.16 mM KCl, 1.19 mM KH_2PO_4, 0.49 mM, $MgCl_2$, 1.17 mM $CaCl_2$, 5.3 mM sodium lactate, 25.07 mM $NaHCO_3$, 0.20 mM sodium pyruvate, 0.5 mM fructose, 1X non-essential amino acids, 5 μg/ml gentamicin, 10 μg/ml heparin, 6 mg/ml fatty acid-free (FFA) BSA) and transferred to drops of SOF-IVF medium under mineral oil. Frozen semen from a Holstein bull was thawed, and 10_6 sperm/ml were added to drops with MII oocytes, which were incubated at 38.5°C for 12-18 h. Zygotes were then removed from fertilization medium, and cumulus cells were removed by vortexing for 5 min in SOF-Hepes medium. Zygotes were then transferred to culture medium (KSOMaa supplemented with 4mg/mL BSA) under mineral oil, and incubated at 38.5°C in 5% CO_2, 5% O_2, and 90% N_2. If embryos were to be transcriptionally inhibited, the culture medium was supplemented with α-amanitin (50 μg/ml) on day one. On day three, the culture medium was supplemented with 5% stem-cell qualified FBS (Gemini Bio 100-525). Blastocyst development was evaluated at 7 days post-insemination (dpi).

Collection of oocytes and preimplantation embryos for ATAC-seq

Oocytes and embryos were collected for ATAC-seq library preparation from three separate collections per developmental stage. Embryos intended for collection at the 2-cell, 4-cell, or 8-cell stages were divided into two groups, one of which was supplemented with α-amanitin, and cultured simultaneously. Germinal vesicle-stage oocytes were collected for ATAC-seq prior to maturation. Preimplantation embryos were collected at the 2-cell (30-32 h post-insemination), 4-cell (2 dpi), 8-cell (3 dpi), and morula stages (5 dpi).

ATAC-seq library preparation

Oocytes or embryos (a minimum of 500 cells) were treated with pronase (10 mg/ml) to completely remove the *zona pellucida* and washed with SOF-Hepes on a warming plate. Cells were then transferred to 1 ml cold SOF-Hepes, and centrifuged at 500 rcf, 4°C, for 5 min. Morulae were subjected to additional vortexing for 3 min in cold ATAC-seq lysis buffer (10 mM Tris-Cl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$ and 0.1% IGEPAL CA-630). Cell pellets were then resuspended in 1 ml cold ATAC-seq cell lysis buffer and centrifuged at 500 rcf, 4°C, for 10 min. Nuclear pellets were then resuspended in 50 µl transposition reaction mix (25 µl TD buffer (Nextera DNA Library Prep Kit, Illumina), 2.5 µl TDE1 enzyme (Nextera DNA Library Prep Kit, Illumina), 22.5 µl nuclease-free H_2O) and incubated for 60 min at 37°C, shaking at 300 rpm. The transposase, which is loaded with Illumina sequencing adapters, cuts DNA where it is not sterically hindered and simultaneously ligates adapters, effectively producing a library in one incubation step. Transposed DNA was purified with the MinElute PCR purification kit (Qiagen, Hilden, Germany) and eluted in 10 µl buffer EB. Libraries were then PCR amplified: 50 µl reactions (25 µl SsoFast Evagreen supermix with low ROX (Bio-Rad, Hercules, CA), 0.6 µl 25 µM custom Nextera PCR primer 1, 0.6 µl 25 µM custom Nextera PCR primer 2 (for a list of primers, see Buenrostro *et al* (2015)[78]), 13 µl nuclease-free H_2O, and 10 µl eluted DNA) were cycled as follows: 72°C for 5 min, 98°C for 30 s, and then thermocycling at 98°C for 10 s, 63°C for 30 s and 72°C for 1 min. Libraries from GV oocytes, 2-cell, and 4-cell embryos were thermocycled for 13 cycles, and 8-cell and morulae libraries were thermocycled for 11 cycles. PCR-amplified libraries were again purified with the MinElute PCR purification kit and eluted in 10 µl buffer EB. Libraries were then evaluated for DNA concentration and nucleosomal laddering patterns using the Bioanalyzer 2100 DNA High Sensitivity chip (Agilent, Palo Alto, CA). Expected nucleosomal laddering was evidenced by the presence of both small fragments, corresponding to hyper-accessible DNA that was frequently transposed, and larger fragments, corresponding to DNA that was wrapped around one or more nucleosomes. This study focused on mapping open chromatin; therefore, the sub-nucleosomal length fraction of each library (150-250 bp) was size selected using the PippinHT system (Sage Science, Beverly, MA) with 3% agarose cassettes. Size-selected libraries were

run on a Bioanalyzer DNA High Sensitivity chip to confirm size-selection and determine DNA concentration. Final libraries were then pooled for sequencing on the Illumina NextSeq platform to generate 40 bp paired end reads.

ATAC-seq read alignment and peak calling

Raw sequencing reads were trimmed with Trim_Galore, a wrapper around Cutadapt (v0.4.0)[79], to remove residual Illumina adapter sequences and low quality (q<20) ends, keeping unpaired reads and reads 10 bp or longer after trimming. Trimmed reads were then aligned to either the GRCm38 (mouse), GRCh38 (human), or ARS-UCD1.2 (cattle) assemblies using BWA aln (-q 15) and sampe[80]. PCR duplicates were removed with PicardTools (v2.8.1), and mitochondrial and low-quality alignments (q<15) were removed with SAMtools (v1.7)[81]. Alignments from biological replicates from each stage were merged and randomly subsampled to equivalent depth with SAMtools for detection of open chromatin. To determine which regions of the genome demonstrated significant enrichment of ATAC-seq signal, broad peaks were called with MACS2 (v2.1.1)[82], using a q-value cutoff of 0.05, and settings *--nomodel --shift -100 --extsize 200*.

RNA-seq alignment and gene expression quantification

Raw sequencing reads were trimmed with Trimmomatic (v0.33)[83]. Low-quality leading and trailing bases (3 bases) were clipped, and Illumina adapter sequences were removed, allowing 2 seed mismatches, a palindrome clip threshold of 30, and a simple clip threshold of 10. Sliding window trimming was conducted with a window size of 4 bases, and a quality threshold of 15. Reads 36 bases or longer were retained after trimming. Trimmed reads were aligned to either the GRCm38 (mouse), GRCh38 (human), or ARS-UCD1.2 (cattle) assemblies with STAR (v2.7.2a)[84] with options *–outFilterScoreMinOverLread 0.85* and *–seedSearchStartLmax 30*. Low quality alignments (q<5) were removed with SAMtools. Raw counts were calculated for genes in the Ensembl 96 annotations for each species with summarizeOverlaps, from the R package GenomicAlignments (v1.18.1)[85], using ‘Union’ mode and

allowing fragments for paired end data. Gene counts were MLE-normalized using the DESeq2 R package (v1.22.2)[86] and submitted to the variance stabilizing transformation for some analyses. DESeq2 was also used for differential expression analysis, with genes demonstrating a logFC > 2 and an adjusted p-value < 0.05 considered differentially expressed.

Repeat expression quantification

Trimmed RNA-seq reads were aligned to the ARS-UCD1.2 genome assembly with STAR with options *–outFilterMultimapNmax 100*, *–winAnchorMultimapNmax 100*, and *–twopassMode Basic*. Raw expression values for individual repetitive elements were calculated for repeats in the RepeatMasker annotation for the ARS-UCD1.2 assembly (downloaded from the UCSC Genome Browser) using TEtoolkit (v2.0.3)[87] in 'multi' mode, which improves quantification of transposable elements transcripts by including ambiguously mapped reads. Raw expression values were MLE-normalized using DESeq2.

Comparison of replicate libraries and ATAC-seq and RNA-seq signal at regions of interest

For both ATAC-seq and RNA-seq data, alignments were converted to bigwig format using bamCoverage from the DeepTools suite[88], which binned the genome into 50 bp windows and calculated normalized signal (reads per kilobase million; RPKM) in each window. The plotPCA function from DeepTools was then used to generate principal components plots, with options *–transpose* and *–log2*. The plotCorrelation function from DeepTools was used to calculate the Spearman correlation coefficient between replicate libraries, based on genome-wide normalized coverage. To assess average accessibility or expression at genomic intervals of interest, average ATAC-seq or RNA-seq normalized signal from bigwig files was visualized using the Deeptools plotHeatmap function.

Comparison and classification of ATAC-seq peaks

Peak sets from different stages were compared using the BEDtools intersect function[89], requiring a minimum of 1 bp overlap to consider a peak shared by both sets. Similarly, peaks were classified as genic

534 if they overlapped either the 2 kb region upstream of a transcription start site (TSS), exons, or introns by 1
535 bp. Otherwise, peaks were considered intergenic.
536
537 *Motif enrichment*
538 Genomic regions were evaluated for binding motif enrichment using the findMotifsGenome.pl script from
539 HOMER (v4.8)[90], using the exact sizes of the input genomic intervals (*–size given*). The most significant
540 known or *de novo* motifs were reported, based on p-value. Known motifs that matched significantly
541 enriched *de novo* motifs were reported if their match score exceeded 0.6.
542
543 *Genome-wide motif location prediction*
544 Position-weight matrices were downloaded from the JASPAR database for TFs of interest[91]. Using the
545 FIMO tool from the MEME suite (v5.0.4)[92], TF motif locations ($p < 1e\text{-}4$) were predicted genome-wide in
546 the ARS-UCD1.2 genome assembly.
547
548 *Repeat class enrichment in genomic intervals*
549 To determine if repetitive elements (either individual elements, families, or classes) were enriched in open
550 chromatin, the number of ATAC-seq peaks overlapping a set of repetitive elements was compared to
551 randomized intervals (ATAC-seq peak locations shuffled with BEDtools shuffle function) overlapping
552 the same set of repetitive elements, yielding a log ratio of random to observed.
553
554 *Functional annotation enrichment analysis*
555 Gene sets were submitted to DAVID (v6.8)[93,94] to identify enriched biological functions. Gene ontology
556 terms with a false discovery rate (FDR) < 0.05 were reported.
557
558
559

Acknowledgements

This work was supported by a grant from NIH-NICHD (R01HD070044) to P.J.R. and R.M.S. Semen used for *in vitro* embryo production was provided by Semex (Madison, WI). M.M.H. was supported by USDA NIFA National Need Fellowship #20143842021796 and an Austin Eugene Lyons Fellowship. The authors thank the entire Ross laboratory for contributions to embryo production and lively discussion.

Author contributions

M.M.H., R.M.S. and P.J.R conceived and designed the experiments; X.M. and M.M.H performed the experiments; M.M.H., R.M.S. and P.J.R wrote the manuscript.

Competing interests

The authors declare no competing financial interests.

Data availability

The following published data sets were used, and accessed through the NCBI GEO repository: for bovine oocytes and *in vitro* produced embryos, raw RNA-seq data were downloaded from accession number GSE52415[67], and mouse and human preimplantation embryo ATAC-seq and RNA-seq data, raw sequencing files were downloaded from accession numbers GSE66390[15] and GSE101571[16], respectively. The ATAC-seq data produced in this study is available via the NCBI SRA repository under the SRA accession number PRJNA595394.

References

1. Ram, P. T. & Schultz, R. M. Reporter gene expression in G2 of the 1-cell mouse embryo. *Dev. Biol.* **156**, 552–556 (1993).
2. Schulz, K. N. & Harrison, M. M. Mechanisms regulating zygotic genome activation. *Nat. Rev. Genet.* **20**, 221–234 (2019).

3. Eckersley-Maslin, M. A., Alda-Catalinas, C. & Reik, W. Dynamics of the epigenetic landscape during the maternal-to-zygotic transition. *Nat. Rev. Mol. Cell Biol.* **19**, 436–450 (2018).

4. White, C. R., MacDonald, W. A. & Mann, M. R. W. Conservation of DNA Methylation Programming Between Mouse and Human Gametes and Preimplantation Embryos. *Biol. Reprod.* **95**, 61 (2016).

5. Ross, P. J. *et al.* Polycomb gene expression and histone H3 lysine 27 trimethylation changes during bovine preimplantation development. *Reproduction* **136**, 777–785 (2008).

6. Funaya, S. & Aoki, F. Regulation of zygotic gene activation by chromatin structure and epigenetic factors. *J. Reprod. Dev.* **63**, 359–363 (2017).

7. Ooga, M., Fulka, H., Hashimoto, S., Suzuki, M. G. & Aoki, F. Analysis of chromatin structure in mouse preimplantation embryos by fluorescent recovery after photobleaching. *Epigenetics* **11**, 85–94 (2016).

8. Ahmed, K. *et al.* Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. *PLoS One* **5**, e10531–e10531 (2010).

9. Martin, C. *et al.* Genome restructuring in mouse embryos during reprogramming and early development. *Dev. Biol.* **292**, 317–332 (2006).

10. Probst, A. V *et al.* A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Dev Cell* **19**, 625–638 (2010).

11. Du, Z. *et al.* Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* **547**, 232–235 (2017).

12. Bultman, S. J. *et al.* Maternal BRG1 regulates zygotic genome activation in the mouse. *Genes Dev* **20**, 1744–1754 (2006).

13. Jimenez, R. *et al.* Maternal SIN3A regulates reprogramming of gene expression during mouse preimplantation development. *Biol. Reprod.* **93**, 89 (2015).

14. Hernandez, C. *et al.* Dppa2/4 Facilitate Epigenetic Remodeling during Reprogramming to Pluripotency. *Cell Stem Cell* **23**, 396-411.e8 (2018).

15. Wu, J. *et al.* The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**, 652–657 (2016).

16. Wu, J. *et al.* Chromatin analysis in human early development reveals epigenetic transition during ZGA. (2018). doi:10.1038/s41586-018-0080-8

17. Li, L. *et al.* Single-cell multi-omics sequencing of human early embryos. *Nat. Cell Biol.* **20**, 847–858 (2018).

18. Gao, L. *et al.* Chromatin Accessibility Landscape in Human Early Embryos and Its Association with Evolution. *Cell* **173**, 248-259.e15 (2018).

19. Liu, L. *et al.* An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos. *Nat. Commun.* **10**, 364 (2019).

20. Lu, F. *et al.* Establishing Chromatin Regulatory Landscape during Mouse Preimplantation Development. *Cell* **165**, 1375–1388 (2016).

21. Bogliotti, Y. S. *et al.* Transcript profiling of bovine embryos implicates specific transcription factors in the maternal-to-embryo transition. *Biol. Reprod.* (2019). doi:10.1093/biolre/ioz209

22. Whiddon, J. L., Langford, A. T., Wong, C.-J., Zhong, J. W. & Tapscott, S. J. Conservation and innovation in the DUX4-family gene network. *Nat. Genet.* **49**, 935–940 (2017).

23. Hendrickson, P. G. *et al.* Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934 (2017).

24. De Iaco, A. *et al.* DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.* **49**, 941–945 (2017).

25. Kruse, K. *et al.* Transposable elements drive reorganisation of 3D chromatin during early embryogenesis. *bioRxiv* 523712 (2019). doi:10.1101/523712

26. Falco, G. *et al.* Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. *Dev. Biol.* **307**, 539–550 (2007).

27. Takahashi, K., Ross, P. J. & Sawai, K. The necessity of ZSCAN4 for preimplantation development and gene expression of bovine embryos. *J. Reprod. Dev.* **65**, 319–326 (2019).

638 28. Chen, X. *et al.* Key role for CTCF in establishing chromatin structure in human embryos. *Nature*
639 **576**, 306–310 (2019).
640 29. Braude, P., Bolton, V. & Moore, S. Human gene expression first occurs between the four- and
641 eight-cell stages of preimplantation development. *Nature* **332**, 459–461 (1988).
642 30. Davis, D. L. Culture and storage of pig embryos. *J. Reprod. Fertil. Suppl.* **33**, 115–124 (1985).
643 31. Crosby, I. M., Gandolfi, F. & Moor, R. M. Control of protein synthesis during early cleavage of
644 sheep embryos. *J. Reprod. Fertil.* **82**, 769–775 (1988).
645 32. Barnes, F. L. & First, N. L. Embryonic transcription in in vitro cultured bovine embryos. *Mol.*
646 *Reprod. Dev.* **29**, 117–123 (1991).
647 33. Zhang, A. *et al.* Dynamic changes of histone H3 trimethylated at positions K4 and K27 in human
648 oocytes and preimplantation embryos. *Fertil. Steril.* **98**, 1009–1016 (2012).
649 34. Wu, X. *et al.* Multiple histone site epigenetic modifications in nuclear transfer and in vitro
650 fertilized bovine embryos. *Zygote* **19**, 31–45 (2011).
651 35. Liu, X. *et al.* Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-
652 implantation embryos. *Nature* **537**, 558–562 (2016).
653 36. Zhang, B. *et al.* Allelic reprogramming of the histone modification H3K4me3 in early mammalian
654 development. *Nature* **537**, 553–557 (2016).
655 37. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of
656 native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding
657 proteins and nucleosome position. *Nat Methods* **10**, 1213–1218 (2013).
658 38. Ghirlando, R. & Felsenfeld, G. CTCF: making the right connections. *Genes Dev* **30**, 881–891
659 (2016).
660 39. Lubelsky, Y., Reuven, N. & Shaul, Y. Autorepression of Rfx1 Gene Expression: Functional
661 Conservation from Yeast to Humans in Response to DNA Replication Arrest. *Mol. Cell. Biol.* **25**,
662 10665–10673 (2005).
663 40. Nishikimi, A., Mukai, J. & Yamada, M. Nuclear translocation of nuclear factor kappa B in early 1-

cell mouse embryos. *Biol. Reprod.* **60**, 1536–1541 (1999).

41. Yang, Y. *et al.* The E3 ubiquitin ligase RNF114 and TAB1 degradation are required for maternal-to-zygotic transition. *EMBO Rep.* **18**, 205–216 (2017).

42. Paciolla, M. *et al.* Nuclear factor-kappa-B-inhibitor alpha (NFKBIA) is a developmental marker of NF-kappaB/p65 activation during in vitro oocyte maturation and early embryogenesis. *Hum. Reprod.* **26**, 1191–1201 (2011).

43. Young, J. M. *et al.* DUX4 binding to retroelements creates promoters that are active in FSHD muscle and testis. *PLoS Genet.* **9**, e1003947–e1003947 (2013).

44. Leidenroth, A. & Hewitt, J. E. A family history of DUX4: phylogenetic analysis of DUXA, B, C and Duxbl reveals the ancestral DUX gene. *BMC Evol. Biol.* **10**, 364 (2010).

45. Eckersley-Maslin, M. *et al.* Dppa2 and Dppa4 directly regulate the Dux-driven zygotic transcriptional program. *Genes Dev.* **33**, 194–208 (2019).

46. Boroviak, T. *et al.* Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. *Development* **145**, dev167833 (2018).

47. Bakhtari, A. & Ross, P. J. DPPA3 prevents cytosine hydroxymethylation of the maternal pronucleus and is required for normal development in bovine embryos. *Epigenetics* **9**, 1271–1279 (2014).

48. Payer, B. *et al.* Stella is a maternal effect gene required for normal early development in mice. *Curr. Biol.* **13**, 2110–2117 (2003).

49. Chen, Z. & Zhang, Y. Loss of DUX causes minor defects in zygotic genome activation and is compatible with mouse development. *Nat. Genet.* **51**, 1 (2019).

50. Guo, M. *et al.* Precise temporal regulation of Dux is important for embryo development. *Cell Res.* (2019). doi:10.1038/s41422-019-0238-4

51. Iaco, A. De, Verp, S., Offner, S. & Trono, D. DUX is a non-essential synchronizer of zygotic genome activation. *bioRxiv* 569434 (2019). doi:10.1101/569434

690 52. Ko, M. S. H. Zygotic Genome Activation Revisited: Looking Through the Expression and
691 Function of Zscan4. *Curr. Top. Dev. Biol.* **120**, 103–124 (2016).
692 53. Bai, H., Sakurai, T., Godkin, J. D. & Imakawa, K. Expression and potential role of GATA factors
693 in trophoblast development. *J. Reprod. Dev.* **59**, 1–6 (2013).
694 54. Wiekowski, M., Miranda, M. & DePamphilis, M. L. Requirements for promoter activity in mouse
695 oocytes and embryos distinguish paternal pronuclei from maternal and zygotic nuclei. *Dev. Biol.*
696 **159**, 366–378 (1993).
697 55. Koehler, D. *et al.* Changes of higher order chromatin arrangements during major genome
698 activation in bovine preimplantation embryos. *Exp. Cell Res.* **315**, 2053–2063 (2009).
699 56. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and
700 adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
701 57. Di Giammartino, D. C. *et al.* KLF4 is involved in the organization and regulation of pluripotency-
702 associated three-dimensional enhancer networks. *Nat. Cell Biol.* **21**, 1179–1190 (2019).
703 58. Oldfield, A. J. *et al.* Histone-fold domain protein NF-Y promotes chromatin accessibility for cell
704 type-specific master transcription factors. *Mol. Cell* **55**, 708–722 (2014).
705 59. Soufi, A. *et al.* Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate
706 reprogramming. *Cell* **161**, 555–568 (2015).
707 60. Marin, M., Karis, A., Visser, P., Grosveld, F. & Philipsen, S. Transcription Factor Sp1 Is Essential
708 for Early Embryonic Development but Dispensable for Cell Growth and Differentiation. *Cell* **89**,
709 619–628 (1997).
710 61. Adelson, D. L., Raison, J. M. & Edgar, R. C. *Characterization and distribution of*
711 *retrotransposons and simple sequence repeats in the bovine genome*. *PNAS August* **4**, (2009).
712 62. Hancks, D. C. & Kazazian Jr, H. H. Roles for retrotransposon insertions in human disease. *Mob.*
713 *DNA* **7**, 9 (2016).
714 63. Spadafora, C. Endogenous reverse transcriptase: a mediator of cell proliferation and
715 differentiation. *Cytogenet. Genome Res.* **105**, 346–350 (2004).

64. Chuong, E. B., Elde, N. C. & Feschotte, C. Regulatory activities of transposable elements: from conflicts to benefits. *Nat. Rev. Genet.* **18**, 71–86 (2017).

65. Raviram, R. *et al.* Analysis of 3D genomic interactions identifies candidate host genes that transposable elements potentially regulate. *Genome Biol.* **19**, 216 (2018).

66. Bui, L. C. *et al.* Retrotransposon expression as a defining event of genome reprograming in fertilized and cloned bovine embryos. *Reproduction* **138**, 289–299 (2009).

67. Graf, A. *et al.* Genome activation in bovine embryos: review of the literature and new insights from RNA sequencing experiments. *Anim Reprod Sci* **149**, 46–58 (2014).

68. Yandım, C. & Karakülah, G. Expression dynamics of repetitive DNA in early human embryonic development. *BMC Genomics* **20**, 439 (2019).

69. Franke, V. *et al.* Long terminal repeats power evolution of genes and gene expression programs in mammalian oocytes and zygotes. *Genome Res.* **27**, 1384–1394 (2017).

70. Beraldi, R., Pittoggi, C., Sciamanna, I., Mattei, E. & Spadafora, C. Expression of LINE-1 retroposons is essential for murine preimplantation development. *Mol Reprod Dev* **73**, 279–287 (2006).

71. Jachowicz, J. W. *et al.* LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat. Genet.* **49**, 1502 (2017).

72. Bestor, T. H. & Bourc'his, D. Transposon silencing and imprint establishment in mammalian germ cells. *Cold Spring Harb. Symp. Quant. Biol.* **69**, 381–387 (2004).

73. Martens, J. H. *et al.* The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *Embo j* **24**, 800–812 (2005).

74. Pontis, J. *et al.* Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic Genome Activation and Control Transcription in Naive Human ESCs. *Cell Stem Cell* **24**, 724-735.e5 (2019).

75. Perea-Gomez, A. *et al.* Otx2 is required for visceral endoderm movement and for the restriction of posterior signals in the epiblast of the mouse embryo. *Development* **128**, 753–765 (2001).

76. Acampora, D., Di Giovannantonio, L. G. & Simeone, A. Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development* **140**, 43–55 (2013).

77. Graf, T. Historical origins of transdifferentiation and reprogramming. *Cell Stem Cell* **9**, 504–516 (2011).

78. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**, 21 29 1–9 (2015).

79. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**, (2011).

80. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).

81. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

82. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

83. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

84. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

85. Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118–e1003118 (2013).

86. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

87. Lerat, E., Fablet, M., Modolo, L., Lopez-Maestre, H. & Vieira, C. TEtools 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–e17 (2017).

88. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).

768 89. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic
769 features. *Bioinformatics* **26**, 841–842 (2010).
770 90. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-
771 regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576–589 (2010).
772 91. Fornes, O. *et al.* JASPAR 2020: update of the open-access database of transcription factor binding
773 profiles. *Nucleic Acids Res.* (2019). doi:10.1093/nar/gkz1001
774 92. Bailey, T. L. *et al.* MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res.* **37**,
775 W202–W208 (2009).
776 93. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large
777 gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
778 94. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward
779 the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
780

781 **Figure legends**

782 Figure 1. Chromatin accessibility in bovine oocytes and *in vitro* preimplantation embryos. A) Schematic
783 of *in vitro* embryo production and ATAC-seq library preparation. B) Principal components analysis
784 (PCA) of ATAC-seq read depth normalized by reads per kilobase million (RPKM) in 50 bp windows
785 covering the whole genome. C) Normalized coverage (RPKM) of replicate ATAC-seq libraries for each
786 stage of development. D) Proportion of genome covered by genic and intergenic ATAC-seq peaks, called
787 from 30 million reads at each developmental stage. E) Categorization of accessible regions in 2-, 4-, and
788 8-cell embryos into stage-specific and maintained peaks (accessibility maintained up until the morula
789 stage). Maintained peaks are carried over from latter stages to show cumulative maintained peaks.
790

791 Figure 2. Gradual establishment of open chromatin enriched for regulatory motifs. A) Regions that
792 significantly lost accessibility from the GV oocyte to 2-cell stage, and regions that significantly gained
793 accessibility from the 2- to 4-cell, 4- to 8-cell, and 8-cell to morula stages, according to 30 million reads

per stage. Accessibility at each region, scaled to average width ± 1 kb, was determined by normalized ATAC-seq read depth (reads per kilobase million; RPKM), based on 20 million reads per stage. B) Top five enriched *de novo* motifs enriched in intergenic peaks with matching known motifs (match score > 0.6) of TFs that were expressed at the given stage. C) Top seven known motifs enriched in 4-cell specific peaks. D) DESeq2 normalized expression of TFs corresponding to enriched motifs in 4-cell specific open chromatin. E) Top seven known motifs enriched in 8-cell specific peaks. F) DESeq2 normalized expression of TFs corresponding to enriched motifs in 8-cell specific open chromatin. G) Proportion of intergenic 8-cell peaks overlapping the *de novo* motif most closely matching the DUXA motif, relative to background. H) DESeq2 normalized expression of DUXA throughout development. RNA-seq data from Graf *et al* (2014).

Figure 3. Binding motif enrichment in regions that opened at the 2-, 4-, or 8-cell stages and remained open until at least the morula stage. Proportion of peaks with A) CTCF or B) KLF5 binding motifs. C) DESeq2 normalized expression of KLFs. RNA-seq data from Graf *et al* (2014). D) Proportion of maintained and stage-specific peaks that were genic or intergenic. E) Gene ontology term enrichment of the 9,456 genes that were marked by maintained genic open chromatin that was first established in 2- or 4-cell embryos.

Figure 4. Effect of transcription inhibition on chromatin remodeling. A) Accessibility status of loci that should have opened or closed between consecutive stages in TBEs. ATAC-seq peaks were called based on 20 million reads per stage and categorized as either genic or intergenic. B) Normalized read depth (RPKM) at loci which opened during the 4- to 8-cell transition in 8-cell control and transcriptionally inhibited embryos. Regions scaled to average width ± 1 kb. C) Normalized ATAC-seq and RNA-seq signal (RPKM) in 8-cell and 8-cell TBEs at the KLF4 locus. RNA-seq data from Bogliotti *et al* (2019). D) Proportion of stage-specific and maintained peaks that should appear at each stage, but which fail to open in TBEs.

Figure 5. Expression and accessibility dynamics of repeat families during bovine preimplantation development. A) DESeq2 normalized expression profiles of select LTR, LINE, and SINE repeat families. RNA-seq data from Graf *et al* (2014). B) Enrichment of several transposable element families in ATAC-seq peaks, called from 30 million reads. Promoter peaks fell within the 2 kb region upstream of transcription start sites (TSS). Intergenic peaks did not overlap the 2 kb regions upstream of TSS, exons, or introns.

Figure 6. Activity of LTR elements in bovine preimplantation embryos. A) Enrichment of LTR elements in open chromatin in 4-cell, 8-cell, and morula-stage embryos, and expression of the same LTR elements. Variance-stabilized normalized expression shown for three replicates per stage. B) ATAC-seq and RNA-seq read coverage at a highly expressed ERV1-2_BT locus. ATAC-seq tracks show coverage from 30 million reads per library; RNA-seq tracks show combined coverage from three replicates. TF motifs predicted from JASPAR binding motifs MA0039.3 (KLF4) and MA0712.1 (OTX2), respectively. C) Average normalized ATAC-seq and D) RNA-seq signal (RPKM) at MLT1A0 repeats overlapped by 8-cell intergenic open chromatin harboring DUXA motifs, predicted from JASPAR motif MA0468.1. E) Co-option of an accessible MLT1A0 element with a DUXA binding motif as an alternative promoter upstream of the C1D locus. RNA-seq data from Graf *et al* (2014).

Figure 7. Inference of key regulatory factors during EGA in cattle, human, and mouse, based on enrichment of TF binding motifs in open chromatin and expression of the corresponding TFs. Bovine RNA-seq data from Graf *et al* (2014); human ATAC-seq and RNA-seq data from Wu *et al* (2018); mouse ATAC-seq and RNA-seq from Wu *et al* (2016).

Figure 8. Potential mechanistic model depicting events leading to major EGA. We postulate that chromatin structure is globally decondensed following fertilization, allowing opportunistic binding of maternal factors which initiate a minor wave of transcription and begin to establish 3D chromatin

846 architecture. This sets the stage for major EGA, wherein maternal products, minor EGA products, and
847 promoter-enhancer contacts collectively regulate the first major wave of gene expression and continue to
848 refine 3D chromatin structure.

Figure 1.

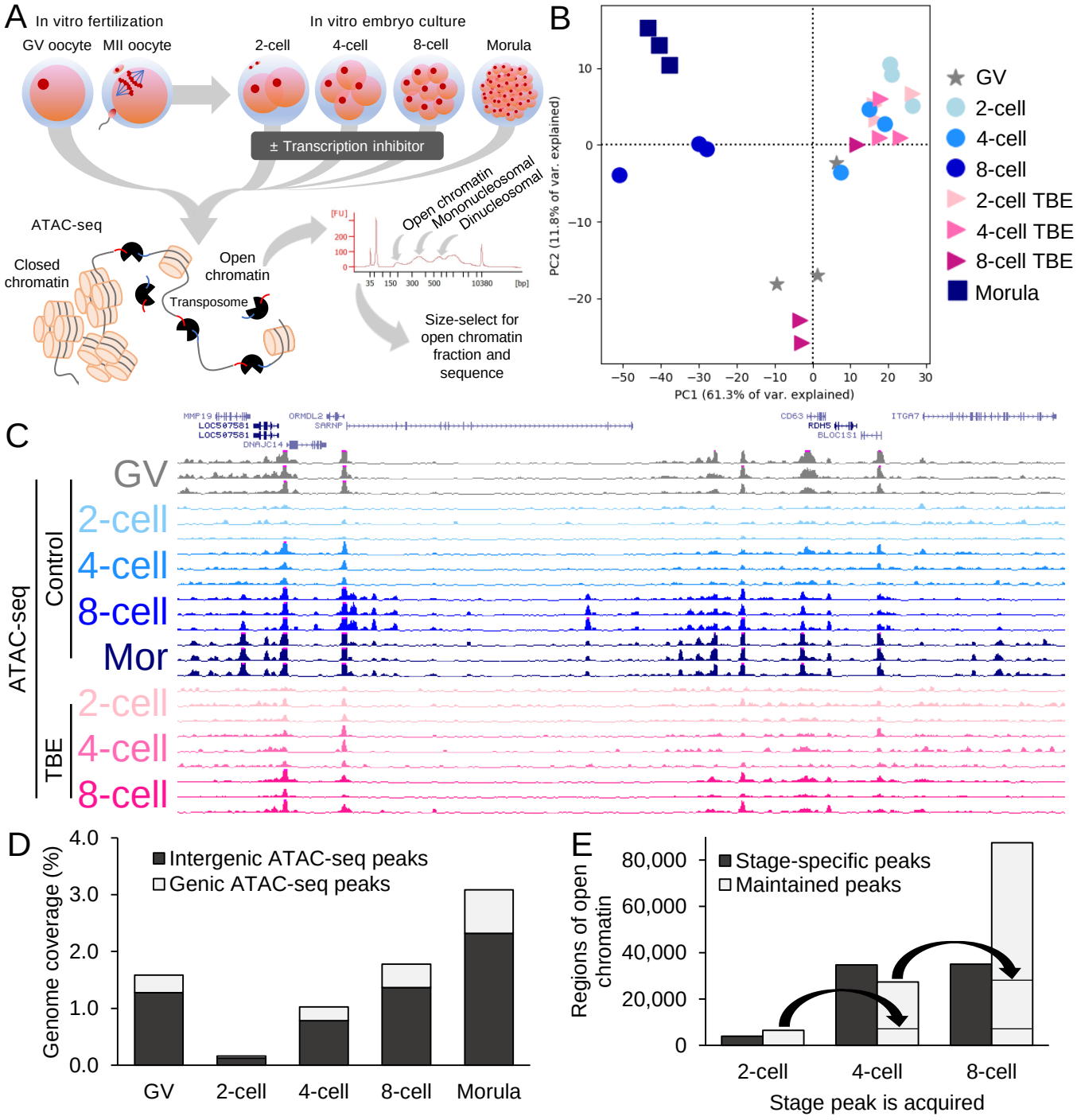

A
In vitro fertilization
GV oocyte
MII oocyte
In vitro embryo culture
2-cell
4-cell
8-cell
Morula
± Transcription inhibitor
ATAC-seq
Closed chromatin
Open chromatin
Transposome
Open chromatin
Mononucleosomal
Dinucleosomal
Size-select for open chromatin fraction and sequence
B
PC2 (11.8% of var. explained)
PC1 (61.3% of var. explained)
GV
2-cell
4-cell
8-cell
2-cell TBE
4-cell TBE
8-cell TBE
Morula
C
ATAC-seq
Control
TBE
GV
2-cell
4-cell
8-cell
Mor
2-cell
4-cell
8-cell
D
Genome coverage (%)
Intergenic ATAC-seq peaks
Genic ATAC-seq peaks
GV
2-cell
4-cell
8-cell
Morula
E
Regions of open chromatin
Stage-specific peaks
Maintained peaks
2-cell
4-cell
8-cell
Stage peak is acquired

Figure 2.

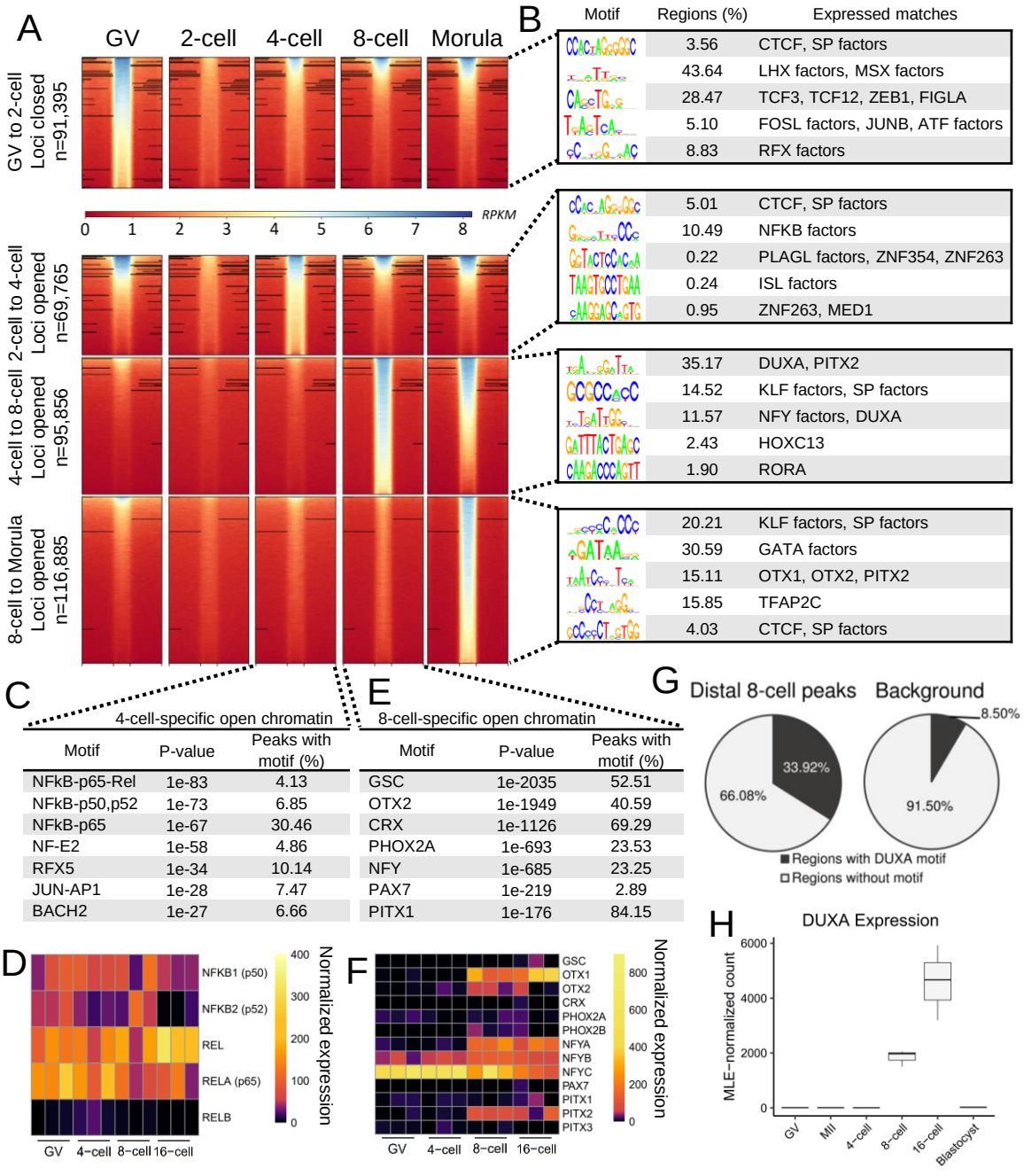

A
GV
2-cell
4-cell
8-cell
Morula
GV to 2-cell Loci closed n=91,395
2-cell to 4-cell Loci opened n=69,765
4-cell to 8-cell Loci opened n=95,856
8-cell to Morula Loci opened n=116,885
0 1 2 3 4 5 6 7 8
RPKM
B
Motif
Regions (%)
Expressed matches
3.56 CTCF, SP factors
43.64 LHX factors, MSX factors
28.47 TCF3, TCF12, ZEB1, FIGLA
5.10 FOSL factors, JUNB, ATF factors
8.83 RFX factors
5.01 CTCF, SP factors
10.49 NFKB factors
0.22 PLAGL factors, ZNF354, ZNF263
0.24 ISL factors
0.95 ZNF263, MED1
35.17 DUXA, PITX2
14.52 KLF factors, SP factors
11.57 NFY factors, DUXA
2.43 HOXC13
1.90 RORA
20.21 KLF factors, SP factors
30.59 GATA factors
15.11 OTX1, OTX2, PITX2
15.85 TFAP2C
4.03 CTCF, SP factors
C
4-cell-specific open chromatin
Motif | P-value | Peaks with motif (%)
NFkB-p65-Rel | 1e-83 | 4.13
NFkB-p50,p52 | 1e-73 | 6.85
NFkB-p65 | 1e-67 | 30.46
NF-E2 | 1e-58 | 4.86
RFX5 | 1e-34 | 10.14
JUN-AP1 | 1e-28 | 7.47
BACH2 | 1e-27 | 6.66
E
8-cell-specific open chromatin
Motif | P-value | Peaks with motif (%)
GSC | 1e-2035 | 52.51
OTX2 | 1e-1949 | 40.59
CRX | 1e-1126 | 69.29
PHOX2A | 1e-693 | 23.53
NFY | 1e-685 | 23.25
PAX7 | 1e-219 | 2.89
PITX1 | 1e-176 | 84.15
G
Distal 8-cell peaks
Background
33.92%
66.08%
8.50%
91.50%
Regions with DUXA motif
Regions without motif
D
NFKB1 (p50)
NFKB2 (p52)
REL
RELA (p65)
RELB
400 300 200 100 0
Normalized expression
GV 4−cell 8−cell 16−cell
F
GSC
OTX1
OTX2
CRX
PHOX2A
PHOX2B
NFYA
NFYB
NFYC
PAX7
PITX1
PITX2
PITX3
800 600 400 200 0
Normalized expression
GV 4−cell 8−cell 16−cell
H
DUXA Expression
MLE−normalized count
6000 4000 2000 0
GV MII 4−cell 8−cell 16−cell Blastocyst

Figure 3.

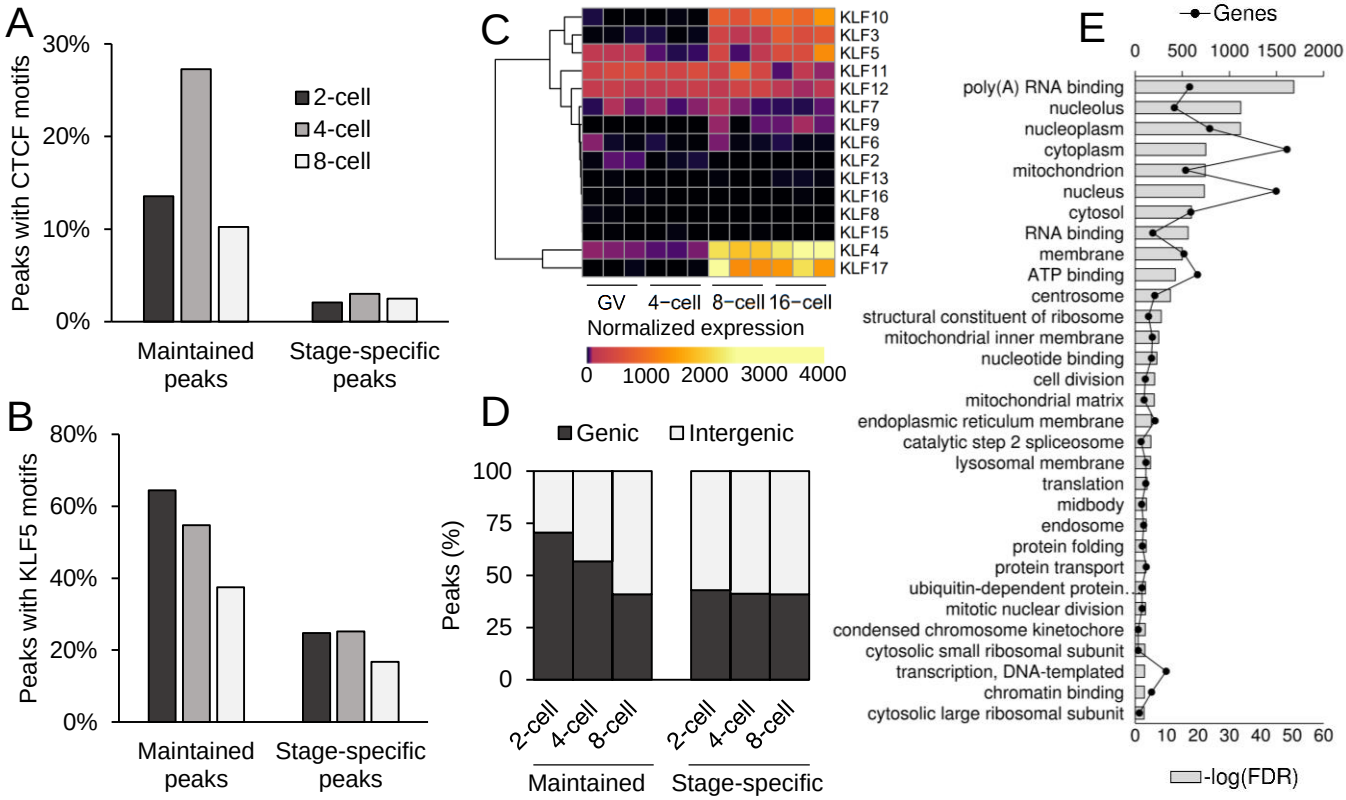

A
Peaks with CTCF motifs
30%
20%
10%
0%
2-cell
4-cell
8-cell
Maintained peaks
Stage-specific peaks
B
Peaks with KLF5 motifs
80%
60%
40%
20%
0%
Maintained peaks
Stage-specific peaks
C
KLF10
KLF3
KLF5
KLF11
KLF12
KLF7
KLF9
KLF6
KLF2
KLF13
KLF16
KLF8
KLF15
KLF4
KLF17
GV
4−cell
8−cell
16−cell
Normalized expression
0
1000
2000
3000
4000
D
Genic
Intergenic
Peaks (%)
100
75
50
25
0
2-cell
4-cell
8-cell
2-cell
4-cell
8-cell
Maintained
Stage-specific
E
Genes
0
500
1000
1500
2000
poly(A) RNA binding
nucleolus
nucleoplasm
cytoplasm
mitochondrion
nucleus
cytosol
RNA binding
membrane
ATP binding
centrosome
structural constituent of ribosome
mitochondrial inner membrane
nucleotide binding
cell division
mitochondrial matrix
endoplasmic reticulum membrane
catalytic step 2 spliceosome
lysosomal membrane
translation
midbody
endosome
protein folding
protein transport
ubiquitin-dependent protein...
mitotic nuclear division
condensed chromosome kinetochore
cytosolic small ribosomal subunit
transcription, DNA-templated
chromatin binding
cytosolic large ribosomal subunit
0
10
20
30
40
50
60
-log(FDR)

Figure 4.

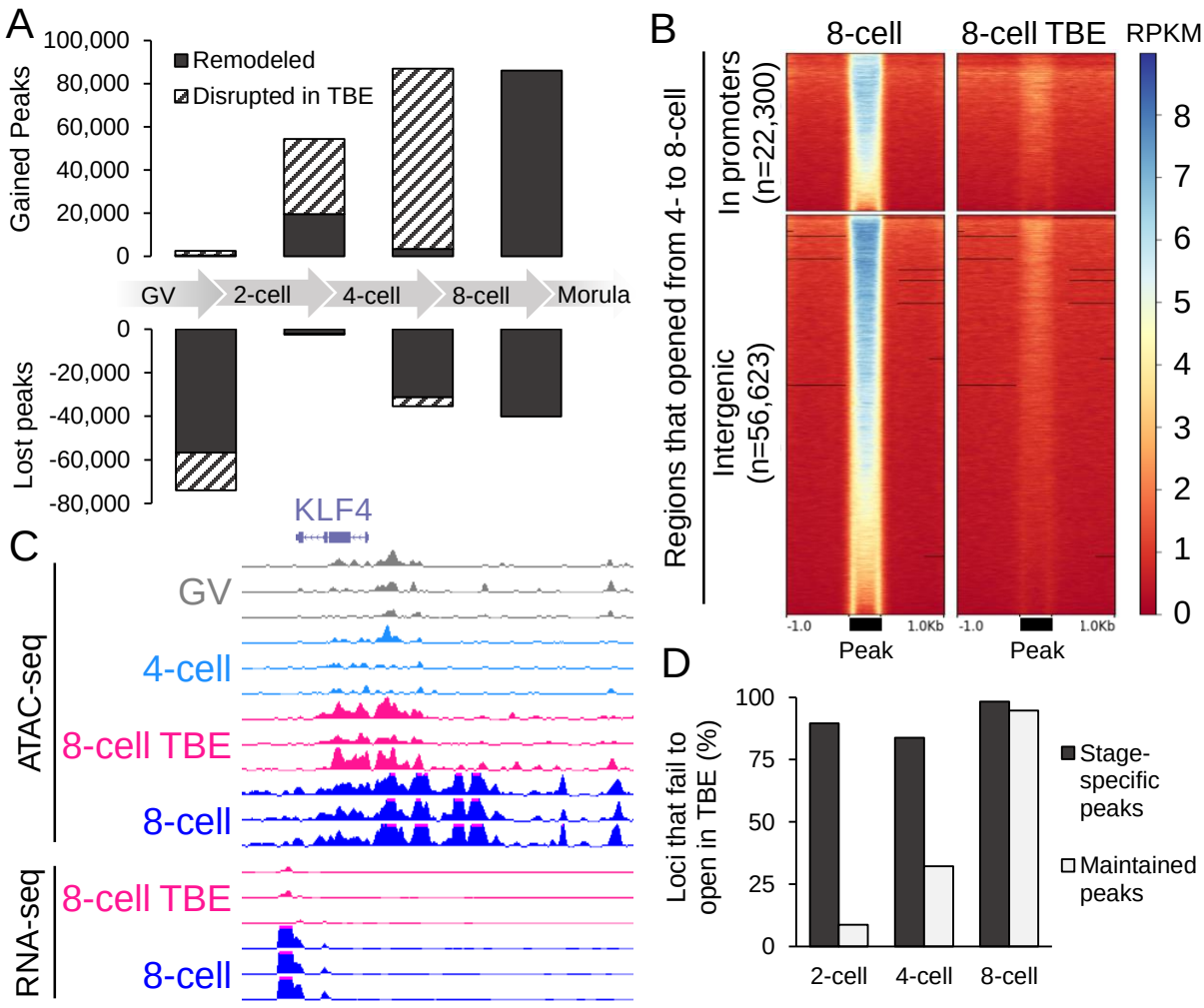

A
Gained Peaks
100,000
80,000
60,000
40,000
20,000
0
Remodeled
Disrupted in TBE
GV
2-cell
4-cell
8-cell
Morula
0
-20,000
-40,000
-60,000
-80,000
Lost peaks
B
8-cell
8-cell TBE
RPKM
Regions that opened from 4- to 8-cell
In promoters (n=22,300)
Intergenic (n=56,623)
8
7
6
5
4
3
2
1
0
-1.0
1.0Kb
Peak
C
KLF4
ATAC-seq
GV
4-cell
8-cell TBE
8-cell
RNA-seq
8-cell TBE
8-cell
D
Loci that fail to open in TBE (%)
100
75
50
25
0
2-cell
4-cell
8-cell
Stage-specific peaks
Maintained peaks

Figure 5.

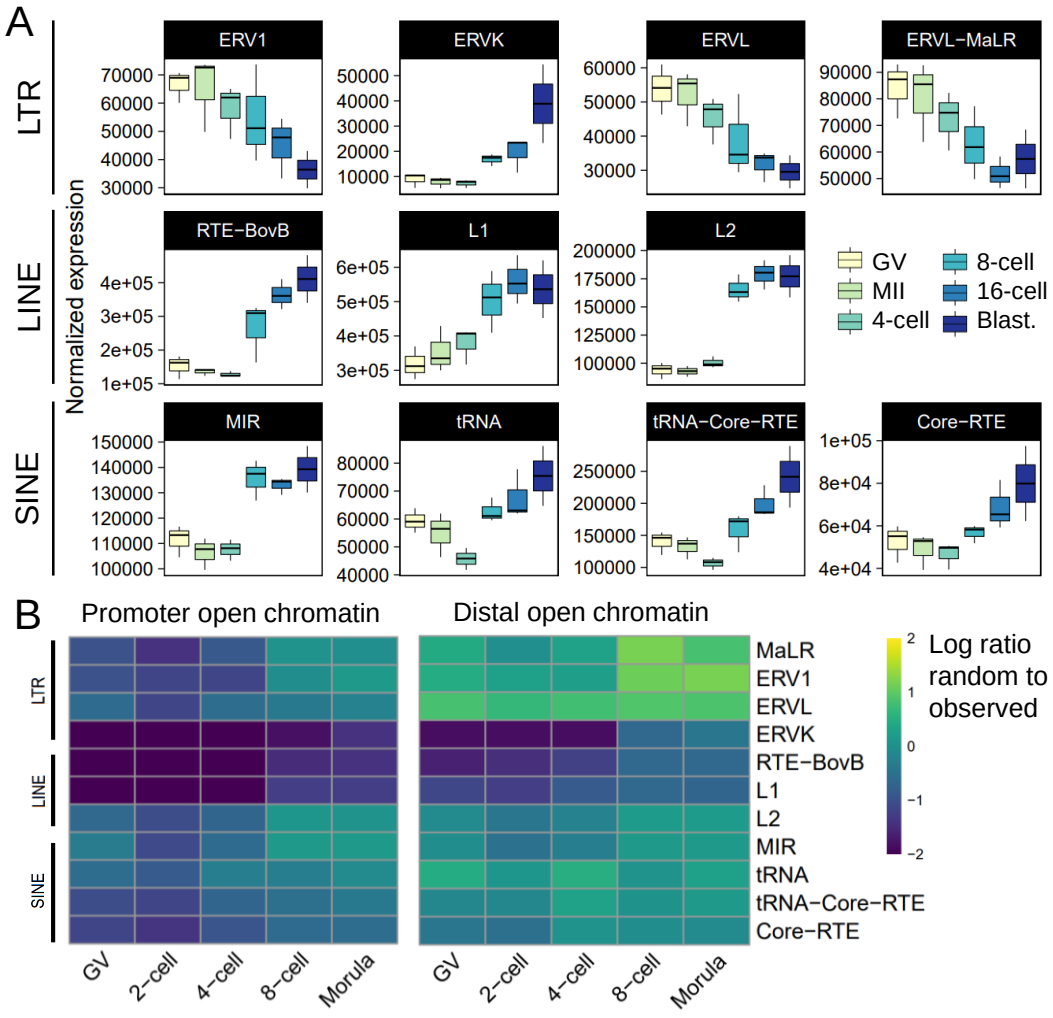
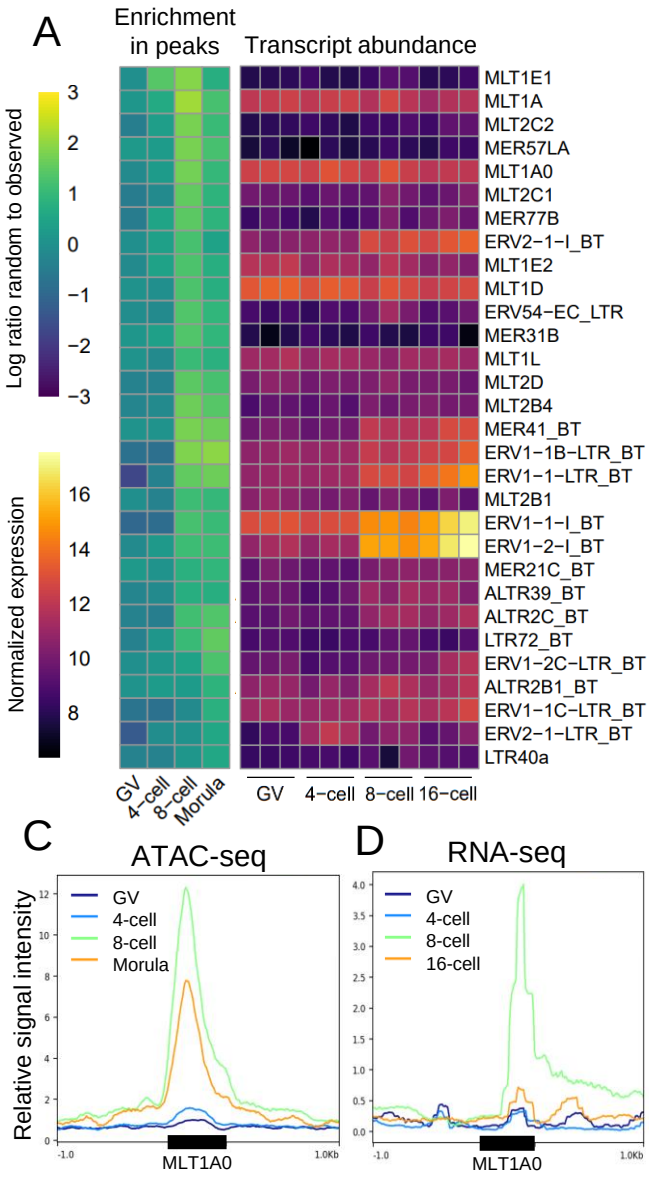
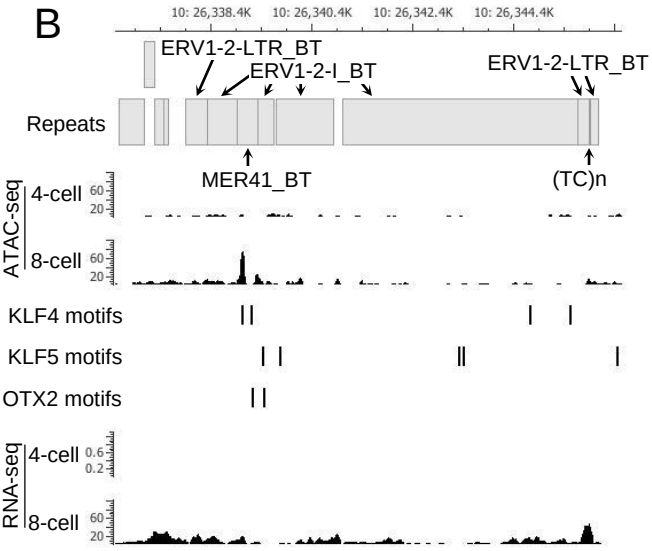

A
LTR
LINE
SINE
Normalized expression
ERV1
ERVK
ERVL
ERVL−MaLR
RTE−BovB
L1
L2
MIR
tRNA
tRNA−Core−RTE
Core−RTE
GV
MII
4-cell
8-cell
16-cell
Blast.
B
Promoter open chromatin
Distal open chromatin
MaLR
ERV1
ERVL
ERVK
RTE−BovB
L1
L2
MIR
tRNA
tRNA−Core−RTE
Core−RTE
Log ratio random to observed
GV
2−cell
4−cell
8−cell
Morula

Figure 6.


A
Enrichment in peaks
Transcript abundance
Log ratio random to observed
Normalized expression
MLT1E1
MLT1A
MLT2C2
MER57LA
MLT1A0
MLT2C1
MER77B
ERV2−1−I_BT
MLT1E2
MLT1D
ERV54−EC_LTR
MER31B
MLT1L
MLT2D
MLT2B4
MER41_BT
ERV1−1B−LTR_BT
ERV1−1−LTR_BT
MLT2B1
ERV1−1−I_BT
ERV1−2−I_BT
MER21C_BT
ALTR39_BT
ALTR2C_BT
LTR72_BT
ERV1−2C−LTR_BT
ALTR2B1_BT
ERV1−1C−LTR_BT
ERV2−1−LTR_BT
LTR40a
GV
4−cell
8−cell
Morula
16−cell


B
10: 26,338.4K
10: 26,340.4K
10: 26,342.4K
10: 26,344.4K
ERV1-2-LTR_BT
ERV1-2-I_BT
ERV1-2-LTR_BT
Repeats
MER41_BT
(TC)n
ATAC-seq
4-cell
8-cell
KLF4 motifs
KLF5 motifs
OTX2 motifs
RNA-seq
4-cell
8-cell

C
ATAC-seq
Relative signal intensity
GV
4-cell
8-cell
Morula
MLT1A0
D
RNA-seq
GV
4-cell
8-cell
16-cell
MLT1A0

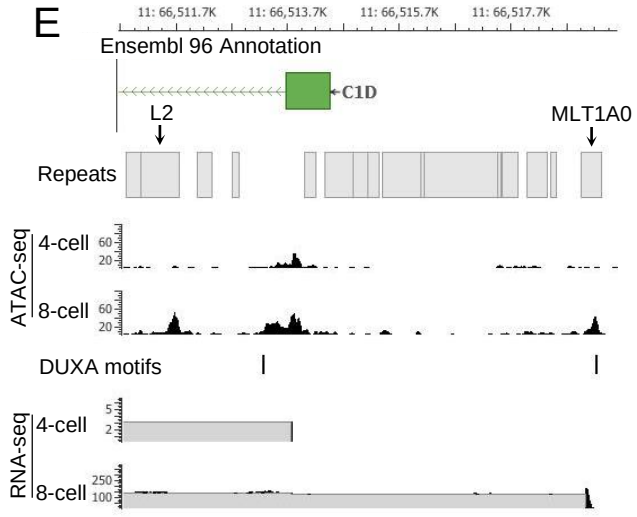

E
11: 66,511.7K
11: 66,513.7K
11: 66,515.7K
11: 66,517.7K
Ensembl 96 Annotation
C1D
L2
MLT1A0
Repeats
ATAC-seq
4-cell
8-cell
DUXA motifs
RNA-seq
4-cell
8-cell

Figure 7.

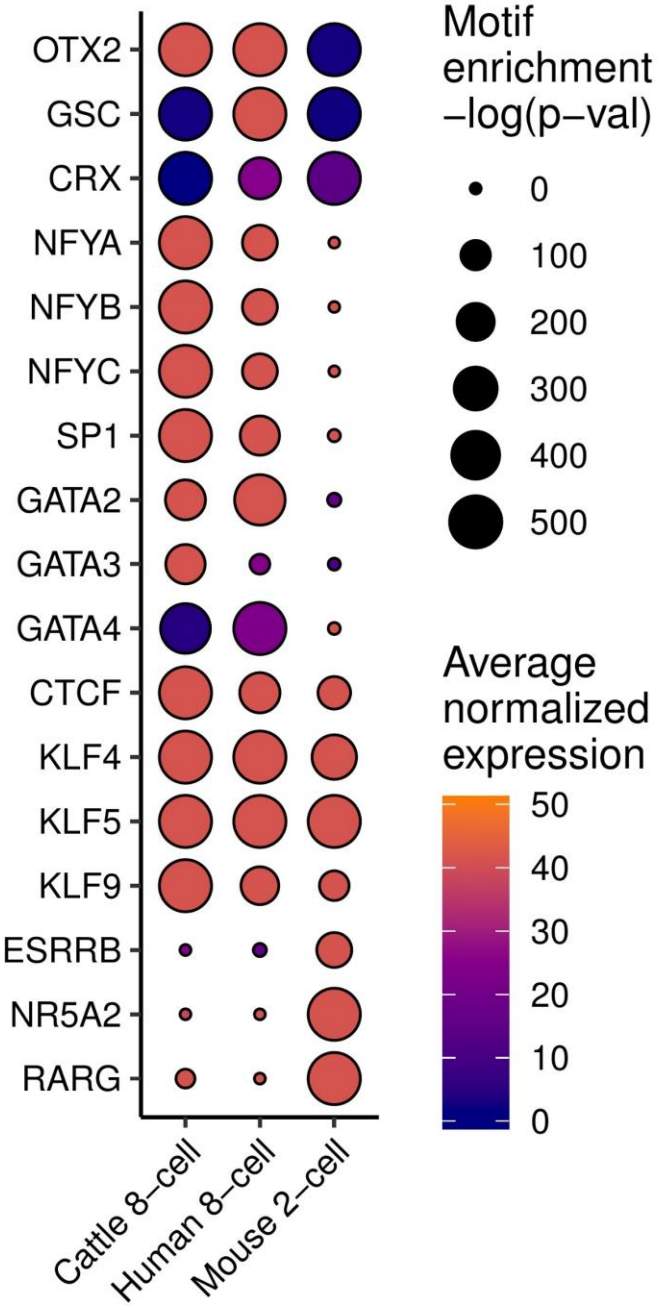
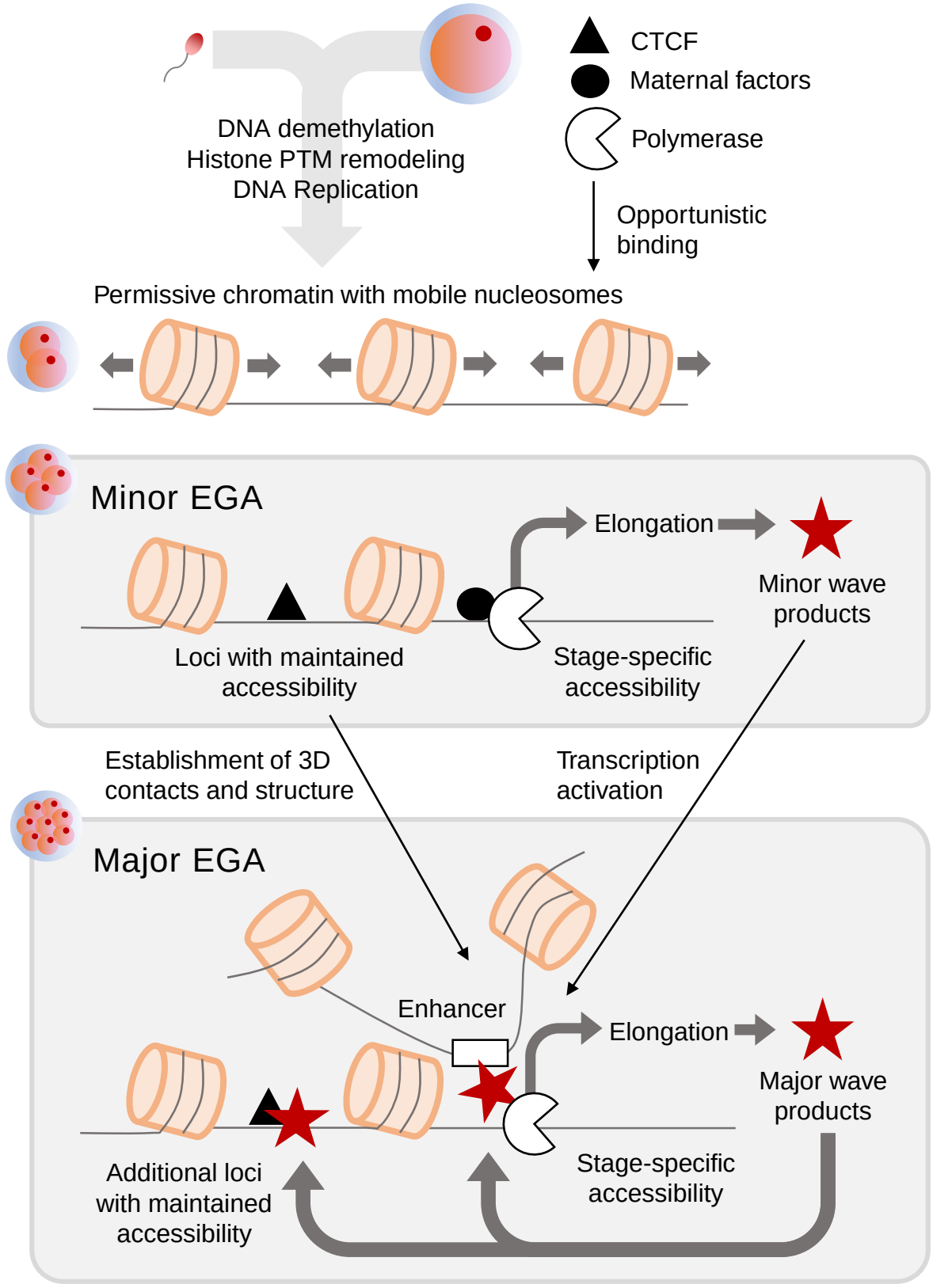

OTX2
GSC
CRX
NFYA
NFYB
NFYC
SP1
GATA2
GATA3
GATA4
CTCF
KLF4
KLF5
KLF9
ESRRB
NR5A2
RARG
Cattle 8–cell
Human 8–cell
Mouse 2–cell
Motif enrichment –log(p–val)
0
100
200
300
400
500
Average normalized expression
50
40
30
20
10
0

Figure 8.


CTCF
Maternal factors
Polymerase
Opportunistic binding
DNA demethylation
Histone PTM remodeling
DNA Replication
Permissive chromatin with mobile nucleosomes
Minor EGA
Elongation
Minor wave products
Loci with maintained accessibility
Stage-specific accessibility
Establishment of 3D contacts and structure
Transcription activation
Major EGA
Enhancer
Elongation
Major wave products
Additional loci with maintained accessibility
Stage-specific accessibility