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. Although the MZT is a universal feature of mammalian preimplantation development, the factors that regulate this transition remain unclear because the timing of embryonic genome activation (EGA) is highly species-specific, suggesting divergence of regulatory circuitry. Indeed, open chromatin during mouse and human EGA shows differential motif enrichment, implicating distinct regulators. To better assess chromatin remodeling during mammalian development, chromatin accessibility was profiled in bovine oocytes and *in vitro* produced embryos. Accessibility that was established pre-EGA marked maternal factor recognition sites, i.e. 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. Subsequently, open chromatin established during EGA was primarily enriched for homeobox motifs, including that of DUX, a potential pioneer factor in mice and humans. Remarkably, the open chromatin landscape during bovine EGA clearly resembled that of humans', revealing that cattle could be a more relevant model for human development than mouse.

Introduction

Preimplantation development encompasses several critical milestones as embryos progress following fertilization to blastocyst implantation. Fusion of the transcriptionally quiescent oocyte and sperm results in a zygote with two haploid pronuclei. Gametic syngamy is completed 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 lead to formation of 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 (Abe *et al*, 2018; Ram & Schultz, 1993), and leads to major EGA, which

involves wide-spread activation of embryonic transcription (Schulz & Harrison, 2019; Vastenhouw *et al*, 2019). 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 (Latham *et al*, 1991).

While embryos undergo this massive shift in transcriptional program during the MZT, wide-spread changes in chromatin structure (Eckersley-Maslin *et al*, 2018; Xu & Xie, 2018) and other epigenetic marks are simultaneously restructuring the embryonic epigenome. 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 (Iurlaro *et al*, 2017; White *et al*, 2016), and this loss of DNA methylation coincides with global decreases in repressive histone modifications (Eckersley-Maslin *et al*, 2018; Xu & Xie, 2018; Ross *et al*, 2008). Overall, epigenetic factors linked to relaxed chromatin are more abundant in mouse zygotes, whereas factors implicated in chromatin compaction become more prevalent during EGA (Funaya & Aoki, 2017), pointing to a more permissive chromatin landscape in pre-EGA embryos. Indeed, mouse zygotes demonstrate elevated histone mobility (Ooga *et al*, 2016), highly dispersed chromatin(Ahmed *et al*, 2010), and lack chromocenters (congregations of pericentromeric heterochromatin) (Martin *et al*, 2006; Probst *et al*, 2007, 2010). Moreover, in mouse, 3-D chromatin architecture is largely absent after fertilization, and is gradually established throughout preimplantation development, which facilitates long-distance chromatin interactions in later stage embryos (Du *et al*, 2017), and indicates increasing chromatin compaction and organization. The co-occurrence of genome activation and chromatin remodeling raises an interesting causality dilemma, namely, whether chromatin remodeling is necessary for transcription activation, or

69 whether transcription activation leads to chromatin accessibility. Several maternal products prompt
70 transcription initiation by altering chromatin structure (Bultman *et al*, 2006; Jimenez *et al*, 2015;
71 Hernandez *et al*, 2018), and some chromatin compaction occurs in the absence of embryonic transcription
72 (Du *et al*, 2017). However, inhibiting embryonic transcription pervasively disrupts the establishment of
73 open chromatin during EGA(Wu *et al*, 2016, 2018), suggesting that EGA and chromatin remodeling are
74 likely interdependent.
75
76 In mouse and human preimplantation embryos, accessible sites are gradually established as development
77 progresses (Li *et al*, 2018; Wu *et al*, 2018; Gao *et al*, 2018; Liu *et al*, 2019; Lu *et al*, 2016; Wu *et al*,
78 2016); however, these open regions demonstrate different motif enrichment patterns, implicating distinct
79 sets of transcription factors (TFs) in either murine (RARG, NR5A2, ESRRB), (Wu *et al*, 2016) or human
80 EGA (OTX2, GSC, POU5F1) (Wu *et al*, 2018; Gao *et al*, 2018). Although some TFs appear to regulate
81 EGA in multiple species, i.e. Krüppel-like factors (KLFs) (Bogliotti *et al*, 2019; Wu *et al*, 2016, 2018),
82 double homeobox (DUX) (Whiddon *et al*, 2017; Hendrickson *et al*, 2017; De Iaco *et al*, 2017; Young *et*
83 *al*, 2013; Kruse *et al*, 2019; Iturbide & Torres-Padilla, 2017), ZSCAN4 (Falco *et al*, 2007; Takahashi *et*
84 *al*, 2019), and CTCF (Wu *et al*, 2016, 2018), it remains unclear if there is any mechanistic conservation
85 across mammals. In fact, the timing of genome activation is highly species-specific: major EGA in mice
86 occurs during the 2-cell stage (Latham *et al*, 1991; Ram & Schultz, 1993), in humans (Braude *et al*, 1988)
87 and pigs (Davis, 1985) during the 4- to-8-cell stage, and in sheep (Crosby *et al*, 1988) and cattle (Frei *et*
88 *al*, 1989; Barnes & First, 1991) between the 8- and 16-cell stages. The relative timing with which mice
89 activate their genomes could indicate that the mechanism behind murine EGA differs significantly from
90 other species', which would have significant implications for modeling human preimplantation
91 development. In particular, the timing of bovine EGA more closely resembles that of human EGA, as do
92 global changes to histone PTMs. For instance, the active mark trimethylation of lysine 4 on histone 3
93 (H3K4me3) decreases in global abundance during human(Zhang *et al*, 2012) and bovine EGA(Wu *et al*,
94 2011), but increases during murine EGA(Liu *et al*, 2016; Zhang *et al*, 2016). However, the chromatin

95 remodeling events that underscore bovine preimplantation development have yet to be catalogued, and it
96 remains unclear whether the regulation and execution of the MZT in cattle resembles that which has been
97 observed in humans.
98
99 To this end, we here describe the chromatin accessibility landscapes of bovine oocytes and
100 preimplantation embryos. We find that open chromatin is gained progressively throughout development,
101 with promoters enriched for CTCF and NFY motifs gaining accessibility in earlier stages, and distal open
102 chromatin becoming accessible at later stages. Moreover, embryonic transcription was not required for
103 the appearance of promoter open chromatin in 2- and 4-cell embryos, but was absolutely necessary to
104 establish stage-specific and distal open chromatin, especially in 8-cell embryos, indicating that maternal
105 and embryonic products both participate in chromatin remodeling in a complementary fashion. Sequence
106 enrichment in open chromatin revealed that several TFs likely play roles in both bovine and human EGA
107 (OTX2, SP1), whereas regulators specific to murine EGA(Wu *et al*, 2016) demonstrate no enrichment in
108 bovine (RARG, NR5A2, ESRRB), suggesting that cattle may be a more informative model for human
109 preimplantation development. Nevertheless, several TFs (DUX, KLFs, CTCF, NFY) seem to play a role
110 in preimplantation development in all three species, raising the possibility that genome activation may be
111 mechanistically conserved across mammals, whereas the specific transcriptional program that is activated
112 may differ between species.
113
114 **Results and Discussion**
115 **Global dynamics of open chromatin in bovine preimplantation embryos**
116 For each developmental stage, ATAC-seq libraries were prepared from cells derived from three separate
117 oocyte collections. A subset of embryos from each collection was also cultured in the presence of the
118 transcriptional inhibitor, α-amanitin, to interrogate the relationship between embryonic transcription and
119 chromatin remodeling (Figure 1a). Between 30 and 87 million non-mitochondrial monoclonal uniquely
120 mapping reads were collectively obtained for each developmental stage, and at least 20 million non-

121 mitochondrial monoclonal reads were collectively obtained for transcription blocked embryos (TBE) at
122 each stage (Table S1). Genome-wide normalized ATAC-seq coverage demonstrated a striking shift in the
123 open chromatin landscape between 4- and 8-cell embryos (Figure 1b), suggesting that large-scale
124 chromatin remodeling coincided with the main transcriptomic shift observed during major EGA (Figure
125 S1). Similarity between replicates (Table S2) indicated that both the technique and embryo production
126 were robust, generating comparable chromatin accessibility profiles across different rounds of oocyte
127 collection and embryo production (Figure 1c). Reads from replicates were pooled together to obtain
128 greater sequencing depth and maximize power for identifying regions of open chromatin. To gauge
129 changes in chromatin accessibility throughout development, regions of open chromatin, or peaks, were
130 called for each stage of development. To minimize bias from sequencing depth, peaks were called from
131 either 30 million monoclonal uniquely mapped reads when comparing controls, or 20 million reads when
132 comparing TBE with control embryos.
133
134 As observed in humans and mice (Lu *et al*, 2016; Wu *et al*, 2016, 2018; Gao *et al*, 2018; Liu *et al*, 2019),
135 regions of open chromatin were gradually established throughout bovine preimplantation development,
136 with the lowest enrichment for accessible sites in 2-cell embryos (Figure 1d). Rather than reflecting
137 chromatin inaccessibility in 2-cell embryos, this dearth of canonical "open chromatin" probably results
138 from a highly permissive chromatin structure. Indeed, chromocenters are absent from bovine embryos
139 until the early 8-cell stage (Martin *et al*, 2006), indicating a relaxed chromatin configuration in early
140 cleavage-stage embryos. Because assays that employ endonucleases – e.g., ATAC-seq and DNase-seq –
141 depend on increased cutting events at consistently accessible loci, we speculate that genome-wide
142 chromatin relaxation would lead to random cutting events genome-wide, resulting in the observed high
143 background and low enrichment in 2-cell embryos. Attempts to use endonuclease-based methods to
144 profile open chromatin in mouse zygotes (Lu *et al*, 2016; Wu *et al*, 2016) and human 2-cell embryos (Liu
145 *et al*, 2019; Gao *et al*, 2018) have encountered similar difficulties with low enrichment. As in bovine
146 embryos, chromocenters are also conspicuously absent in pre-EGA mouse embryos (Martin *et al*, 2006;

147 Probst *et al*, 2007, 2010), and electron microscopy (Ahmed *et al*, 2010) and fluorescence recovery after
148 photobleaching analysis (Ooga *et al*, 2016) also indicate highly dispersed chromatin. Furthermore, a
149 recent study that detected open chromatin based on methylation of accessible GpC sites, rather than
150 endonucleases, found that genome-wide accessibility in human embryos actually decreased from the
151 zygote stage onward (Li *et al*, 2018). Global chromatin relaxation appears to be a shared characteristic of
152 human, bovine, and murine pre-EGA embryos. Thus, it is tempting to speculate that this "naïve"
153 chromatin state acts as a blank epigenetic slate, which is then gradually compacted and structured to meet
154 the needs of the growing embryo. Indeed, accessible sites were progressively established in 4-cell, 8-cell,
155 and morula stage embryos (Figure 1d). Interestingly, many of these regions were only open transiently at
156 a specific stage, whereas others maintained their accessibility throughout later stages (Figure 1e),
157 suggesting that chromatin remodeling serves two functions: progressive establishment of a "totipotent"
158 chromatin landscape, and transient stage-specific regulation.
159
160 **Stepwise remodeling yields stage-specific open chromatin with distinct functionality**
161 To delve further into the potential function of regions that lost or gained accessibility between
162 consecutive stages (Figure 2a), intergenic regions of open chromatin were evaluated for sequence
163 enrichment, revealing that the binding motifs of distinct groups of TFs were gaining accessibility at each
164 developmental stage (Table S3). These enriched sequences corresponded to the known binding motifs of
165 several TFs that were either maternally provided or expressed in embryos (Figure S2), demonstrating that
166 changing chromatin structure subjects each stage of development to distinct regulatory circuitry (Figure
167 2b).
168
169 The first major transition in chromatin structure mostly involved loss of hyperaccessible sites in 2-cell
170 embryos following fertilization (Figure 2a). Intergenic loci that lost accessibility during this transition
171 (n=54,264) were enriched for the binding motifs of CTCF – an insulator protein implicated in 3D
172 chromatin organization (Ghirlando & Felsenfeld, 2016) – and several RFX factors – a highly conserved

173 family of transcriptional repressors (Lubelsky *et al*, 2005) (Figure 2b). Interestingly, one quarter of the
174 intergenic regions that closed in 2-cell embryos regained accessibility at the 4-cell stage (n=14,859), and
175 many of these remained open all the way through to the morula stage (n=6,186). Most intergenic regions
176 that close after fertilization likely contribute to oocyte-specific regulation, considering that they remain
177 inaccessible in embryos, whereas regions that re-opened in 4-cell embryos could be implicated in general
178 house-keeping functions.
179
180 The first major gains in accessibility occurred in 4-cell embryos (Figure 2a). About a quarter of the
181 regions that opened during the 4-cell stage maintained their accessibility in both 8-cell embryos and
182 morulae (Figure 1e), although 74% of these had been previously open in GV oocytes (n=15,466).
183 Surprisingly, nearly half of the regions that opened during the 4-cell stage were only transiently accessible
184 (Figure 1e). Four-cell-specific open chromatin was most significantly enriched for binding motifs of
185 NFkB family members (Figure 2c), which was particularly intriguing because NFkB activation is
186 specifically necessary in 1-cell mouse embryos for development to progress past EGA (Nishikimi *et al*,
187 1999; Yang *et al*, 2017). Although NFkB factors are maternally provided (Figure 2c), they are initially
188 sequestered in the cytoplasm until they translocate into the nucleus at the early 1-cell stage in mice
189 (Nishikimi *et al*, 1999; Yang *et al*, 2017) and the 4-to-8 cell stage in cattle (Paciolla *et al*, 2011). In
190 particular, one of the NFkB subunits capable of activating gene expression – RELA – binds DNA with
191 increased frequency in cattle embryos compared to oocytes (Paciolla *et al*, 2011), suggesting that NFkB
192 activation of target gene expression could be one of the first events in a cascade leading to major EGA. In
193 fact, one of the few genes with upregulated expression in 4-cell embryos, as compared to MII oocytes,
194 was TRIM8 – a positive regulator of NFkB activity (Figure S3). Additionally, RELA binding sites that
195 were accessible in 4- and 8-cell embryos mark genes that encode key regulators of early preimplantation
196 development, raising the possibility that NFkB binding is involved in their transcription initiation (Figure
197 S4), although the contribution of NFkB signaling to gene expression during minor EGA in cattle has not
198 yet been established.

199 In contrast to open chromatin in pre-EGA embryos, regions that gained accessibility at the 8-cell stage
200 (Figure 2a) were considerably enriched for the binding motifs of several homeobox TFs, including OTX2,
201 GSC, CRX, PHOX2A, PAX7, and PITX1 (Figure 2b,e), although only OTX1, OTX2, and PITX2 were
202 appreciably expressed during the 8-cell stage (Figure 2f). More so than other homeobox recognition
203 motifs, 8-cell open chromatin was most enriched for binding motifs of the double homeobox (DUX)
204 family, which has been extensively implicated in EGA regulation (Whiddon *et al*, 2017; Hendrickson *et*
205 *al*, 2017; De Iaco *et al*, 2017; Young *et al*, 2013; Kruse *et al*, 2019; Iturbide & Torres-Padilla, 2017).
206 More than a third of the accessible intergenic loci in 8-cell embryos harbored a sequence that most closely
207 matched double homeobox (DUX) binding motifs (Figure 2g). DUX factors are increasingly recognized
208 as potential developmental pioneer factors, participating in chromatin remodeling(Kruse *et al*, 2019;
209 Hendrickson *et al*, 2017), and expression of repetitive elements(Whiddon *et al*, 2017; Hendrickson *et al*,
210 2017; De Iaco *et al*, 2017; Young *et al*, 2013) and cleavage-stage genes(Hendrickson *et al*, 2017; Iturbide
211 & Torres-Padilla, 2017). In cattle, DUXA is expressed transiently and strongly during the 8- and 16-cell
212 stages (Figure 2h), similar to DUX4 and Dux expression in human and mouse embryos, respectively, for
213 a specific window prior to EGA(Hendrickson *et al*, 2017; De Iaco *et al*, 2017). Although the maternal
214 factors DPPA2 and DPPA4 induce Dux expression in mouse embryonic stem cells (ESC) (Eckersley-
215 Maslin *et al*, 2019), only DPPA3 is maternally provided in cattle (Figure S5). It remains unclear if
216 DPPA3 is capable of inducing DUX expression; however, DPPA3 is maternally provided in mice, cattle,
217 and humans (Boroviak *et al*, 2018; Bakhtari & Ross, 2014), and knockdown DPPA3 decreases the
218 developmental competency of mouse (Payer *et al*, 2003) and bovine (Bakhtari & Ross, 2014) embryos.
219 Although the maternal factors that induce DUXA expression in cattle remain unclear, the synchronized
220 upregulation of DUXA and increased accessibility of its binding sites during bovine development
221 strongly indicates a conserved role for DUX in mammalian preimplantation development, especially
222 considering that DUX family members are highly conserved and specific to placental mammals (Young *et*
223 *al*, 2013; Leidenroth & Hewitt, 2010). Further functional validation will be necessary to determine if

224 DUXA is required for bovine embryogenesis, or if is an important but non-essential regulator of EGA, as
225 in mouse (Chen & Zhang, 2019; Guo *et al*, 2019; Iaco *et al*, 2019).
226
227 The expression profile of DUXA in bovine preimplantation development closely mirrored that of another
228 TF implicated in EGA: ZSCAN4 (Figure S6a) (Ko, 2016). ZSCAN4 is a downstream target of DUX
229 factors in humans (Hendrickson *et al*, 2017) and mice (De Iaco *et al*, 2017; Whiddon *et al*, 2017), and the
230 coordinated expression of these two factors in cattle certainly suggests a similar mechanism at play.
231 Although ZSCAN4 depletion disrupts development past EGA in mice (Falco *et al*, 2007) and cattle
232 (Takahashi *et al*, 2019), its binding motifs were not enriched in open chromatin during bovine
233 preimplantation development (Figure S6b-d), suggesting that the binding motif of bovine ZSCAN4
234 differs from those in mice and humans.
235
236 Finally, the transition from 8-cell to morula also incurred extensive chromatin remodeling, with even
237 more sites gaining accessibility than during the 8-cell stage (Figure 2a). Compared to earlier stages,
238 intergenic loci that gained accessibility in morulae were primarily enriched for GATA factor binding
239 motifs (Figure 2a; Table S3) – key regulators of trophectoderm establishment and maintenance (Bai *et al*,
240 2013) – bolstering the idea that morulae are already initiating differentiation programs necessary for
241 blastocyst formation.
242
243 **Progressive establishment of maintained open chromatin sets the stage for genome activation**
244 Although many regions only experienced stage-specific gains in accessibility, a stable open chromatin
245 landscape was also progressively established after fertilization. As early as the 2-cell stage, regions began
246 to gain accessibility that was maintained until the morula stage. These regions of maintained accessibility
247 were enriched for CTCF motifs (Figure 3a; Table S4b), especially those that were first established during
248 the 4-cell stage. CTCF binding delineates chromatin loop boundaries, thus determining the genomic space
249 within which genes can interact with their regulatory elements (Ghirlando & Felsenfeld, 2016). Therefore,

250 enrichment of CTCF motifs in maintained peaks could point to a gradual establishment of a stable 3-D
251 chromatin architecture in preparation for major EGA. This proposal is consistent with 3-D chromatin
252 architecture in mouse embryos, which is greatly diminished after fertilization and then gradually re-
253 established throughout preimplantation development, facilitating long-distance chromatin contacts, i.e.,
254 promoter-enhancer interactions, in later stage embryos (Du *et al*, 2017). Indeed, transcription during
255 minor EGA in mouse is primarily driven by proximal promoters, whereas enhancers are dispensable for
256 transcription until major EGA (Abe *et al*, 2015; Wiekowski *et al*, 1993). The global reorganization of 3D
257 chromatin architecture in mouse has also been observed in bovine embryos, wherein gene-rich regions
258 switch from a random distribution to a chromosome-specific distribution during major EGA (Koehler *et*
259 *al*, 2009). Collectively, these results suggest that gradual establishment of 3D chromatin architecture is a
260 conserved feature of pre-EGA embryos, although the mechanisms regulating this restructuring remain
261 unknown.
262
263 Other than CTCF, maintained peaks were also more enriched for KLF motifs than transiently open sites
264 (Figure 3b). Although most KLFs were not expressed until the 8-cell stage, several KLFs were maternally
265 provided (Figure 3c), including KLF4, a master regulator of induced pluripotent stem cell (iPSC)
266 reprogramming (Takahashi & Yamanaka, 2006). Recent evidence suggests that KLF4 contributes to
267 reprogramming iPSC by mediating pluripotency-associated enhancer-promoter contacts (Di Giammartino
268 *et al*, 2019); thus, the concurrent establishment of open chromatin corresponding to CTCF and KLF
269 binding motifs suggests that pre-EGA embryos are priming a similar mechanism for use during major
270 EGA, especially because about 50% of activated genes during bovine major EGA contain KLF motifs in
271 their promoters (Bogliotti *et al*, 2019).
272
273 In fact, loci that remained open from the 2- and 4-cell stages onward occurred in genic regions more often
274 than stage-specific open chromatin (Figure 3d) and marked the promoters of genes that were functionally
275 enriched for housekeeping functions (Figure 3e), including 33 of the 51 genes that were upregulated in 4-

276 cell embryos compared to MII oocytes (Figure S3). Moreover, these regions of maintained accessibility
277 were strongly enriched for NFY and SP1 binding sites (Table S4a,b), which is highly reminiscent of
278 chromatin remodeling in mouse embryos, as proximal promoters enriched for NFY are the first regions to
279 gain accessibility (Lu *et al*, 2016). Of note, NFY enhances binding of the pluripotency factors POU5F1,
280 SOX2, and KL4 to their recognition motifs (Oldfield *et al*, 2014; Soufi *et al*, 2015), and is clearly
281 involved in murine EGA, as NFY knockdown embryos demonstrated impaired open chromatin
282 establishment and downregulation of gene expression (Lu *et al*, 2016). Similarly, SP1 is essential for
283 early mouse development, with knockout embryos arresting at day 11 of gestation (Marin *et al*, 1997).
284 Intriguingly, human zygotes also demonstrate KLF, SP1, and NFY motif enrichment in open chromatin
285 (Liu *et al*, 2019), overall suggesting that a conserved set of maternal regulators participates in chromatin
286 remodeling and transcription activation in pre-EGA embryos.
287
288 **Both maternal products and embryonic transcription drive chromatin reorganization**
289 Considering that embryonic transcription is extremely limited before the 8-cell stage in cattle, the
290 appearance of open chromatin in pre-EGA embryos suggests that maternal factors participate in
291 chromatin remodeling. To further dissect the maternal contribution to epigenetic reprogramming and
292 EGA, embryos were cultured in the presence of α-amanitin to inhibit POLR2-dependent transcription
293 elongation. Loss of embryonic transcription had a drastic inhibitory effect on the appearance of open
294 chromatin and intensified as development progressed (Figure 4a); 64% of loci that should have opened
295 during the 4-cell stage failed to become accessible without embryonic transcription, and in 8-cell TBE
296 embryos, 96% of loci that should have opened remained closed (Figure 4b), disrupting the chromatin state
297 of key genes, such as KLF4 (Figure 4c), and coinciding with developmental arrest.
298
299 Interestingly, inhibiting embryonic transcription did not uniformly affect chromatin remodeling genome-
300 wide, but instead preferentially disrupted establishment of stage-specific open chromatin, whereas
301 maintained open chromatin established in 2- and 4-cell embryos appeared despite embryonic transcription

inhibition (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 (Adelson *et al*, 2009) – 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 (Friedli & Trono, 2015; Hancks & Kazazian Jr, 2016). 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 (Spadafora, 2004). 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 (Chuong *et al*, 2017) and participate in 3-D chromatin architecture (Raviram *et al*, 2018).

327 Although repetitive element expression in bovine early embryos was reported a decade ago using a cDNA
328 array (Bui *et al*, 2009), a complete catalogue of repeat transcription throughout bovine preimplantation
329 development was lacking. To address this gap in knowledge, available RNA-seq data (Graf *et al*, 2014)
330 were assessed for expression of repetitive elements. Importantly, these libraries were not subjected to
331 polyA selection. As has been observed in the mouse and human (Yandım & Karakülah, 2019; Franke *et*
332 *al*, 2017), expression and accessibility of repetitive elements throughout bovine preimplantation
333 development was highly stage-specific and dynamic (Figure 5a; Figure S8).
334
335 Non-long terminal repeat (non-LTR) retrotransposons – long interspersed elements (LINEs) and short
336 interspersed elements (SINEs) – are increasingly transcribed during human preimplantation development
337 (Yandım & Karakülah, 2019; Li *et al*, 2018). In cattle, LINEs also demonstrated increasing expression
338 (Figure 5a) and accessibility (Figure 5b) starting during the 8-cell stage. Although the function of LINE
339 elements in bovine preimplantation embryos has yet to be established, their activation is crucial for mouse
340 development: perturbing LINE expression in mouse preimplantation embryos causes developmental arrest
341 at the 2-cell stage and perturbs gene expression (Beraldi *et al*, 2006). Moreover, LINE-1 activation may
342 regulate global chromatin accessibility in mouse embryos (Jachowicz *et al*, 2017). Of the SINE families,
343 MIR element expression and accessibility patterns most echoed those of L2 LINE elements, with
344 increased expression and accessibility starting at the 8-cell stage, suggesting that these elements could be
345 acting as enhancers or promoters.
346
347 In particular, LTR activation is a key feature in human (Yandım & Karakülah, 2019), mouse (Franke *et*
348 *al*, 2017), and bovine preimplantation development (Bui *et al*, 2009). Of these, mammalian LTRs and
349 ERVLs were increasingly enriched in open chromatin starting during the 8-cell stage (Figure 5b),
350 although their transcript abundance dropped throughout development (Figure 5a), indicating that other
351 mechanisms likely regulate repeat expression, e.g., DNA methylation or histone modifications (Bestor &
352 Bourc'his, 2004; Martens *et al*, 2005; Shirai *et al*, 2017). Only endogenous retroviral K (ERVK) elements

353 demonstrated both increasing expression (Figure 5a) and moderately increased prevalence in distal open
354 chromatin during the 8-cell stage (Figure 5b), suggesting that ERVK elements function as regulatory
355 elements, as observed in human preimplantation embryos(Pontis *et al*, 2019).
356
357 Mounting evidence suggests that specific types of LTR retrotransposons, especially intact elements, play
358 pivotal roles in early development. During bovine preimplantation development, the most expressed
359 retrotransposons were ERV1-1_BT and ERV1-2_BT (Figure 6a), as has been observed previously (Bui *et*
360 *al*, 2009). Upon further inspection, intact ERV1-1_BT elements often co-occurred with MER41_BT
361 repeats in a specific sequence, which demonstrated a highly reproducible pattern of transcription at
362 ERV1-1_BT elements and chromatin accessibility at MER41_BT elements (Figure S9a). Furthermore,
363 MER41_BT elements that were accessible in 8-cell embryos were enriched for the binding motifs of
364 several pluripotency factors, including POU5F1, NFY, KLF4, OTX2, and TEAD (Figure S9b),
365 suggesting that pluripotency factors are driving transposon expression.
366
367 The TF DUX has also been implicated in driving the expression of intact ERVL elements in human and
368 mouse pre-EGA embryos (Whiddon *et al*, 2017; Hendrickson *et al*, 2017; De Iaco *et al*, 2017; Liu *et al*,
369 2019). Specifically, in human embryos DUX4 appears to bind MLT2A1 elements – primate-specific
370 LTRs that flank human ERVL –activating their expression (Liu *et al*, 2019). Considering that intergenic
371 open chromatin was especially enriched for DUX binding sites in 8-cell bovine embryos, it seemed likely
372 that these sites would also correspond to retrotransposons. Indeed, several MLT elements were enriched
373 in 8-cell open chromatin harboring DUX motifs (Table S5), suggesting that bovine DUX binds LTRs
374 with sequence similarity to primate-specific MLT2A1 elements, regulating their expression. Nearly all
375 LTRs that were enriched in 8-cell open chromatin with DUX binding sites demonstrated dynamic
376 expression profiles throughout development (Figure S10). In particular, increasing accessibility at
377 MLT1A0 elements harboring DUX recognition sites (Figure 6c) was mirrored by a sharp increase in
378 transcription in 8-cell embryos (Figure 6d).

379 Whether the transcripts derived from LTRs have a functional role in bovine development, or are simply a
380 result of opportunistic expression, remains to be established. However, evidence in other species suggests
381 that retrotransposon-derived regulatory elements are often co-opted by the embryo as promoter and
382 enhancers(Chuong *et al*, 2017): a phenomenon which appears to extend to bovine embryos, as ATAC-seq
383 and RNA-seq suggest that MLT1A0 elements are co-opted as alternative promoters at several loci,
384 including CD1 (Figure 6e), ZNF41, and LPIN2. As such, an interesting balance appears to exist between
385 repetitive elements and the embryo, wherein the repeatome leverages the embryo's existing regulatory
386 network to drive transposition, while simultaneously providing new regulatory elements and TF binding
387 sites that the embryo co-opts to drive the expression of its own transcriptional programs.
388
389 **A model for mammalian genome activation**
390 Several lines of evidence suggest that the regulatory circuitry responsible for the MZT may differ
391 between mammals. First, EGA occurs in a highly species-specific fashion, with the major wave of
392 transcription in mice during the 2-cell stage (Latham *et al*, 1991; Ram & Schultz, 1993), in humans
393 during the 4-to-8-cell stage (Braude *et al*, 1988), and cattle during the 8-to-16-cell stage (Frei *et al*, 1989).
394 Second, the expression patterns of repetitive elements are not only highly-stage specific, but also species-
395 specific; primate-specific and murine ERVL elements are strongly expressed during human and mouse
396 preimplantation development, respectively (Whiddon *et al*, 2017; Hendrickson *et al*, 2017; De Iaco *et al*,
397 2017; Liu *et al*, 2019), whereas specific ERV1 elements were most prominently activated in bovine
398 embryos. Finally, the maternal programs in human and mice are divergent; human maternal programs
399 conspicuously lack the murine maternal effect transcripts POU5F1, HSF1, and DICER1, and are
400 functionally enriched for translational processing (Boroviak *et al*, 2018), reflecting the need to translate
401 maternal transcripts during the extended period before human EGA.
402
403 Maternal products are essential to support cleavage stage-embryos prior to genome activation, especially
404 in species where EGA is delayed for several cell divisions, e.g., humans and cattle. The relaxed chromatin

405 structure in early preimplantation embryos provides a unique regulatory context for maternal factors.
406 Whereas the condensed chromatin structure in somatic cells generally restricts DNA-binding proteins to
407 regions of open chromatin, the dispersed chromatin in 2-cell bovine embryos may allow maternal factors
408 to opportunistically and pervasively bind their recognition motifs. Indeed, several maternal factors appear
409 to participate in chromatin restructuring in pre-EGA embryos, leading to open chromatin establishment at
410 promoters enriched for NFY and SP1 binding sites, as well as CTCF and KLF motifs. Although not
411 maternally provided, the DUX family appears to play a conserved role in genome activation. In humans
412 and mice, DUX factors have been implicated in chromatin remodeling (Kruse *et al*, 2019; Hendrickson *et*
413 *al*, 2017) and transcription activation of cleavage-stage genes (Hendrickson *et al*, 2017; Iturbide &
414 Torres-Padilla, 2017). We find a similar pattern of DUX expression in bovine embryos, as well as
415 increased accessibility of DUX binding sites around the 8-cell stage, suggesting that DUX may also
416 modulate gene expression and chromatin accessibility during bovine EGA.
417
418 Although the chromatin landscape changes markedly upon major EGA in bovine, human (Wu *et al*, 2018;
419 Liu *et al*, 2019; Li *et al*, 2018), and mouse (Lu *et al*, 2016; Wu *et al*, 2016) embryos, the regulatory
420 circuitry that is active during this stage in development appears to differ significantly between humans
421 and mice (Wu *et al*, 2018), which suggests that regulation of mammalian EGA is highly species-specific.
422 To identify and compare putative regulators of mammalian EGA, regions that were accessible during
423 major EGA in bovine, human, and mouse embryos were evaluated for binding motif enrichment of
424 actively expressed TFs. Comparing EGA regulatory circuitry between species reflects a critical
425 divergence in regulatory and transcriptional programs that clearly separates humans and cattle from the
426 mouse (Figure S11). Compared to mouse 2-cell embryos, open chromatin in bovine and human 8-cell
427 embryos demonstrated remarkably similar patterns of sequence enrichment (Figure 7). In cattle and
428 humans, SP1, OTX2, and NFY were particularly implicated during major EGA, whereas NR5A2, RARG,
429 and ESRRB were solely enriched in mouse embryos. Nevertheless, it is unclear if the regulatory factors
430 that are enriched in open chromatin during major EGA are regulators of EGA or simply products of it.

431 For instance, although NR5A2is enriched in 2-cell mouse embryos, it is an early regulator of inner cell
432 mass and trophectoderm programs and is not essential for genome activation (Wu *et al*, 2016). Similarly,
433 OTX2 is essential for neuronal lineage specification in mice (Perea-Gomez *et al*, 2001), and has been
434 implicated in the transition from naïve to primed pluripotency (Acampora *et al*, 2013; Buecker *et al*,
435 2014; Schoorlemmer *et al*, 2014). However, OTX2 protein is clearly present in human and marmoset
436 zygotes (Boroviak *et al*, 2018), suggesting that this homeobox TF may play an as-of-yet undetermined
437 role in regulation of human and bovine EGA. Although several TFs appeared to only be important in
438 either mouse, or in cattle and humans, KLF factors were substantially enriched during EGA in all three
439 species. Considering the well-established role of KLF factors in somatic cell reprogramming and
440 establishment of pluripotency in multiple species (Takahashi & Yamanaka, 2006; Graf, 2011), KLFs may
441 play a conserved role in the MZT. Although future research will be necessary to elucidate the function of
442 specific regulatory factors, the high consistency between cattle and humans, both with respect to the
443 timing of EGA and the regulatory circuitry that accompanies it, strongly suggests that cattle are a more
444 appropriate model system for human preimplantation development than mouse.
445
446 **Conclusions**
447 Sweeping changes to chromatin structure during bovine preimplantation development suggest that 2-cell
448 embryos are characterized by globally decondensed chromatin, which is gradually compacted as
449 development progresses, which echoes similar observations in humans and mice. In particular, it is
450 tempting to speculate that a conserved set of maternal factors establish basal promoter accessibility and
451 the necessary chromatin architecture for enhancer-promoter interactions that will drive gene expression
452 during major EGA (Figure 8). However, the open chromatin landscape during major EGA clearly
453 distinguished mice from cattle and humans, suggesting that whereas maternal regulation of EGA may be
454 conserved across mammals, the transcriptional programs that are subsequently activated have diverged
455 substantially. Practically, this difference suggests that human development may be better modeled in
456 cattle than in mice. Nevertheless, the factors appear to regulate the MZT in cattle, humans, and mice

457 certainly warrant further investigation and validation, which will provide invaluable insight into the
458 regulatory framework that governs successful preimplantation development
459
460 **Materials and Methods**
461 *Oocyte collection and maturation*
462 Ovaries were procured from a local abattoir and transported to the laboratory in a warm saline solution.
463 Follicles measuring 2-10 mm were aspirated to obtain cumulus oocyte complexes (COCs). Only COCs
464 with healthy layers of cumulus cells were selected for maturation. These were washed in collection
465 medium (6:4 M199 (Sigma M7653): SOF-Hepes, supplemented with 2% fetal bovine serum (FBS;
466 Hyclone/Thermo Scientific) and transferred to maturation medium (modified M199 medium (Sigma
467 M2154) supplemented with ALA-glutamine (0.1 mM), sodium pyruvate (0.2 mM), gentamicin (5 µg/ml),
468 EGF (50 ng/ml), oFSH (50 ng/ml), bLH (3 µg/ml), cysteamine (0.1 mM), and 10% FBS.
469
470 *In vitro fertilization and embryo culture*
471 After COCs matured for 24 h, MII oocytes were washed in SOF-IVF medium (107.7 mM NaCl, 7.16 mM
472 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 $NaHCO_3$,
473 0.20 mM sodium pyruvate, 0.5 mM fructose, 1X non-essential amino acids (NEAA), 5 µg/ml gentamicin,
474 10 µg/ml heparin, 6 mg/ml fatty acid-free (FFA) BSA) and transferred to drops of SOF-IVF medium
475 under mineral oil. Frozen semen from a Holstein bull was thawed, and 10^6 sperm/ml were added to drops
476 with MII oocytes, which were incubated at 38.5°C for 12-18 h. Zygotes were then removed from
477 fertilization medium, and cumulus cells were removed by vortexing for 5 min in SOF-Hepes medium.
478 Zygotes were then transferred to culture media (KSOMaa supplemented with 4mg/mL BSA) under
479 mineral oil, and incubated at 38.5°C in 5% CO_2, 5% O_2, and 90% N_2. If embryos were to be
480 transcriptionally inhibited, the culture medium was supplemented with α-amanitin (50 µg/ml) on day-one.
481 On day-three, the culture medium was supplemented with 5% stem-cell qualified FBS (Gemini Bio 100-
482 525). Blastocyst development was evaluated at 7 days post-insemination (dpi).

483 *Collection of oocytes and preimplantation embryos for ATAC-seq*

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

490

491 *ATAC-seq library preparation*

492 Oocytes or embryos (a minimum of 500 cells) were treated with pronase (10 mg/ml) to completely
493 remove the *zona pellucida* and washed with SOF-Hepes on a warming plate. Cells were then transferred
494 to 1 ml cold SOF-Hepes, and centrifuged at 500 rcf, 4°C, for 5 min. Morulae were vortexed for 3 min in
495 cold ATAC-seq lysis buffer (10 mM Tris-Cl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$ and 0.1% IGEPAL CA-
496 630). Cell pellets were then resuspended in 1 ml cold ATAC-seq cell lysis buffer and centrifuged at 500
497 rcf, 4°C, for 10 min. Nuclear pellets were then resuspended in 50 µl transposition reaction mix (25 µl TD
498 buffer (Nextera DNA Library Prep Kit, Illumina), 2.5 µl TDE1 enzyme (Nextera DNA Library Prep Kit,
499 Illumina), 22.5 µl nuclease-free H_2O) and incubated for 60 min at 37°C, shaking at 300 rpm. The
500 transposase, which is loaded with Illumina sequencing adapters, cuts DNA where it is not sterically
501 hindered and simultaneously ligates adapters, effectively producing a library in one incubation step.
502 Transposed DNA was purified with the MinElute PCR purification kit (Qiagen, Hilden, Germany) and
503 eluted in 10 µl buffer EB. Libraries were then PCR amplified: 50 µl reactions (25 µl SsoFast Evagreen
504 supermix with low ROX (Bio-Rad, Hercules, CA), 0.6 µl 25 µM custom Nextera PCR primer 1, 0.6 µl 25
505 µM custom Nextera PCR primer 2 (with barcode; Supplementary Table 2), 13 µl nuclease-free H_2O, and
506 10 µl eluted DNA) were cycled as follows: 72°C for 5 min, 98°C for 30 s, and then thermocycling at
507 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
508 were thermocycled for 13 cycles, and 8-cell and morulae libraries were thermocycled for 11 cycles. PCR-

509 amplified libraries were again purified with the MinElute PCR purification kit (Qiagen, Hilden, Germany)
510 and eluted in 10 µl buffer EB. Libraries were then evaluated for DNA concentration and nucleosomal
511 laddering patterns using the Bioanalyzer 2100 DNA High Sensitivity chip (Agilent, Palo Alto, CA).
512 Expected nucleosomal laddering was evidenced by the presence of both small fragments, corresponding
513 to hyper-accessible DNA that was frequently transposed, and larger fragments, corresponding to DNA
514 that was wrapped around one or more nucleosomes. This study focused on mapping open chromatin;
515 therefore, the sub-nucleosomal length fraction of each library (150-250 bp) was size selected using the
516 PippinHT system (Sage Science, Beverly, MA) with 3% agarose cassettes. Size-selected libraries were
517 run on a Bioanalyzer DNA High Sensitivity chip to confirm size-selection and determine DNA
518 concentration. Final libraries were then pooled for sequencing on the Illumina NextSeq platform to
519 generate 40 bp paired end reads.
520
521 *ATAC-seq read alignment and peak calling*
522 Raw sequencing reads were trimmed with Trim_Galore, a wrapper around Cutadapt (v0.4.0) (Martin,
523 2011), to remove residual Illumina adapter sequences and low quality (q<20) ends, keeping unpaired
524 reads and reads 10 bp or longer after trimming. Trimmed reads were then aligned to either the GRCm38
525 (mouse), GRCh38 (human), or ARS-UCD1.2 (cattle) assemblies using BWA aln (-q 15) and sampe (Li &
526 Durbin, 2010). PCR duplicates were removed with PicardTools, and mitochondrial and low-quality
527 alignments (q<15) were removed with SAMtools (Li *et al*, 2009). Alignments from biological replicates
528 from each stage were merged and randomly subsampled to equivalent depth with SAMtools for detection
529 of open chromatin. To determine which regions of the genome demonstrated significant enrichment of
530 ATAC-seq signal, broad peaks were called with MACS2 (Zhang *et al*, 2008), using a q-value cutoff of
531 0.05, and settings *--nomodel --shift -100 --extsize 200*.
532
533
534

535 *RNA-seq alignment and gene expression quantification*

536 Raw sequencing reads were trimmed with Trimmomatic (v0.33) (Bolger *et al*, 2014). Low-quality leading
537 and trailing bases (3 bases) were clipped, and illumine adapter sequences were removed, allowing 2 seed
538 mismatches, a palindrome clip threshold of 30, and a simple clip threshold of 10. Sliding window
539 trimming was conducted with a window size of 4 bases, and a quality threshold of 15. Reads 36 bases or
540 longer were retained after trimming. Trimmed reads were aligned to either the GRCm38 (mouse),
541 GRCh38 (human), or ARS-UCD1.2 (cattle) assemblies with STAR (v2.7.2a) with options –
542 outFilterScoreMinOverLread 0.85 and –seedSearchStartLmax 30. Low quality alignments (q<5) were
543 removed with SAMtools. Raw counts were calculated for genes in the Ensembl 96 annotations for each
544 species with summarizeOverlaps, from the R package GenomicAlignments (v1.18.1) (Lawrence *et al*,
545 2013), using 'Union' mode and allowing fragments for paired end data. Gene counts were MLE-
546 normalized using the DESeq2 R package (v1.22.2) (Anders & Huber, 2010) and submitted to the variance
547 stabilizing transformation for some analyses. DESeq2 was also used for differential expression analysis,
548 with genes demonstrating a logFC > 2 and an adjusted p-value < 0.05 considered differentially expressed.

549

550 *Repeat expression quantification*

551 Trimmed reads were aligned to the ARS-UCD1.2 genome assembly with STAR (v2.7.0) (Dobin *et al*,
552 2013) with options –outFilterMultimapNmax 100, --winAnchorMultimapNmax 100, and --twopassMode
553 Basic. Raw expression values for individual repetitive elements were calculated for repeats in the
554 RepeatMasker annotation for the ARS-UCD1.2 assembly (download from the UCSC Genome Browser)
555 using TEtoolkit (Lerat *et al*, 2017) in 'multi' mode, which improves quantification of transposable
556 elements transcripts by including ambiguously mapped reads. Raw expression values were MLE-
557 normalized using the DESeq2 R package (v1.22.2).

558

559

560

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

562 For both ATAC-seq and RNA-seq data, alignments were converted to bigwig format using bamCoverage
563 from the DeepTools suite (Ramírez *et al*, 2014), which binned the genome into 50 bp windows and
564 calculated normalized signal (reads per kilobase million; RPKM) in each window. The plotPCA function
565 from DeepTools was then used to generate the PCA plot, with options *--transpose and --log2.* The
566 plotCorrelation function from DeepTools was used to calculate the Spearman correlation coefficient
567 between replicate libraries, based on genome-wide normalized coverage. To assess average accessibility
568 or expression at genomic intervals of interest, average ATAC-seq or RNA-seq normalized signal was
569 visualized using the Deeptools plotHeatmap function

570

571 *Comparison and classification of ATAC-seq peaks*

572 Peak sets from different stages were compared using the BEDtools intersect function (Quinlan & Hall,
573 2010), requiring a minimum of 1bp overlap to consider a peak shared by both sets. Similarly, peaks were
574 classified as genic if they overlapped either the 2 kb region upstream of a transcription start site (TSS),
575 exons, or introns by 1 bp. Otherwise, peaks were considered intergenic.

576

577 *Motif enrichment*

578 Genomic regions were evaluated for binding motif enrichment using the findMotifsGenome.pl script from
579 HOMER (v4.8) (Heinz *et al*, 2010), using the exact sizes of the input genomic intervals (-size given). The
580 most significant known or *de novo* motifs were reported, based on p-value. Known motifs that matched
581 significantly enriched *de novo* motifs were reported if their match score exceeded 0.6.

582

583 *Genome-wide motif location prediction*

584 Position-weight matrices were downloaded from the JASPAR database for TFs of interest (Fornes *et al*,
585 2019). Using the FIMO tool from the MEME suite (v5.0.4) (Bailey *et al*, 2009), TF motif locations ($p <$
586 1e-4) were predicted genome-wide in the ARS-UCD1.2 genome assembly.

587 *Repeat class enrichment in genomic intervals*

588 To determine if repetitive elements (either individual elements, families, or classes) were enriched in open
589 chromatin, the number of ATAC-seq peaks overlapping a set of repetitive elements was compared to
590 randomized intervals (ATAC-seq peak locations shuffled with BEDtools shuffle function) overlapping
591 the same set of repetitive elements, yielding a log ratio of random to observed.

592

593 *Functional annotation enrichment analysis*

594 Gene sets were submitted to DAVID (v6.8) (Huang *et al*, 2009a, 2009b) to identify enriched biological
595 functions using medium stringency. Gene ontology terms with an FDR < 0.05 were reported.

596

597 **Acknowledgements**

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

602

603 **Author contributions**

604 Conceived and designed the experiments: M.M.H., R.M.S. and P.J.R. Performed the experiments: X.M.
605 and M.M.H. Wrote the manuscript: M.M.H., R.M.S. and P.J.R.

606

607 **Conflict of interest**

608 The authors declare no competing financial interests.

609

610 **Data availability**

611 The following published NGS data sets were used, and accessed through the NCBI GEO repository: for
612 bovine oocytes and *in vitro* produced embryos, raw RNA-seq data were downloaded from accession

613 number GSE52415 (Graf *et al*, 2014). For mouse and human preimplantation embryo ATAC-seq and
614 RNA-seq data, raw sequencing files were downloaded from accession numbers GSE66390 (Wu *et al*,
615 2016) and GSE101571 (Wu *et al*, 2018), respectively. The ATAC-seq data produced in this study is
616 available via the NCBI SRA repository.
617

618 **References**

619 Abe K-I, Funaya S, Tsukioka D, Kawamura M, Suzuki Y, Suzuki MG, Schultz RM & Aoki F (2018)
620 Minor zygotic gene activation is essential for mouse preimplantation development. *Proc. Natl.*
621 *Acad. Sci. U. S. A.* **115:** E6780–E6788
622 Abe K, Yamamoto R, Franke V, Cao M, Suzuki Y, Suzuki MG, Vlahovicek K, Svoboda P, Schultz RM
623 & Aoki F (2015) The first murine zygotic transcription is promiscuous and uncoupled from splicing
624 and 3' processing. *Embo j* **34:** 1523–1537
625 Acampora D, Di Giovannantonio LG & Simeone A (2013) Otx2 is an intrinsic determinant of the
626 embryonic stem cell state and is required for transition to a stable epiblast stem cell condition.
627 *Development* **140:** 43–55 Available at: https://www.ncbi.nlm.nih.gov/pubmed/23154415
628 Adelson DL, Raison JM & Edgar RC (2009) Characterization and distribution of retrotransposons and
629 simple sequence repeats in the bovine genome
630 Ahmed K, Dehghani H, Rugg-Gunn P, Fussner E, Rossant J & Bazett-Jones DP (2010) Global chromatin
631 architecture reflects pluripotency and lineage commitment in the early mouse embryo. *PLoS One* **5:**
632 e10531–e10531 Available at: https://www.ncbi.nlm.nih.gov/pubmed/20479880
633 Anders S & Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol.* **11:**
634 R106 Available at: https://doi.org/10.1186/gb-2010-11-10-r106
635 Bai H, Sakurai T, Godkin JD & Imakawa K (2013) Expression and potential role of GATA factors in
636 trophoblast development. *J. Reprod. Dev.* **59:** 1–6 Available at:
637 https://www.ncbi.nlm.nih.gov/pubmed/23428586
638 Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW & Noble WS (2009)

639 MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res.* **37:** W202–W208
640 Available at: https://doi.org/10.1093/nar/gkp335
641 Bakhtari A & Ross PJ (2014) DPPA3 prevents cytosine hydroxymethylation of the maternal pronucleus
642 and is required for normal development in bovine embryos. *Epigenetics* **9:** 1271–1279 Available at:
643 https://www.ncbi.nlm.nih.gov/pubmed/25147917
644 Barnes FL & First NL (1991) Embryonic transcription in in vitro cultured bovine embryos. *Mol. Reprod.*
645 *Dev.* **29:** 117–123 Available at: https://www.ncbi.nlm.nih.gov/pubmed/1878221
646 Beraldi R, Pittoggi C, Sciamanna I, Mattei E & Spadafora C (2006) Expression of LINE-1 retroposons is
647 essential for murine preimplantation development. *Mol Reprod Dev* **73:** 279–287
648 Bestor TH & Bourc'his D (2004) Transposon silencing and imprint establishment in mammalian germ
649 cells. *Cold Spring Harb. Symp. Quant. Biol.* **69:** 381–387 Available at:
650 https://www.ncbi.nlm.nih.gov/pubmed/16117671
651 Bogliotti YS, Chung N, Paulson E, Chitwood J, Halstead M, Kern C, Schultz RM & Ross PJ (2019)
652 Transcript profiling of bovine embryos implicates specific transcription factors in the maternal-to-
653 embryo transition. *Biol. Reprod.* Available at: https://doi.org/10.1093/biolre/ioz209
654 Bolger AM, Lohse M & Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data.
655 *Bioinformatics* **30:** 2114–2120
656 Boroviak T, Stirparo GG, Dietmann S, Hernando-Herraez I, Mohammed H, Reik W, Smith A, Sasaki E,
657 Nichols J & Bertone P (2018) Single cell transcriptome analysis of human, marmoset and mouse
658 embryos reveals common and divergent features of preimplantation development. *Development* **145:**
659 dev167833 Available at: https://www.ncbi.nlm.nih.gov/pubmed/30413530
660 Braude P, Bolton V & Moore S (1988) Human gene expression first occurs between the four- and eight-
661 cell stages of preimplantation development. *Nature* **332:** 459–461
662 Buecker C, Srinivasan R, Wu Z, Calo E, Acampora D, Faial T, Simeone A, Tan M, Swigut T & Wysocka
663 J (2014) Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell*
664 *Stem Cell* **14:** 838–853 Available at: https://www.ncbi.nlm.nih.gov/pubmed/24905168

665 Bui LC, Evsikov A V., Khan DR, Archilla C, Peynot N, Hénaut A, Le Bourhis D, Vignon X, Renard JP
666 & Duranthon V (2009) Retrotransposon expression as a defining event of genome reprograming in
667 fertilized and cloned bovine embryos. *Reproduction* **138:** 289–299
668 Bultman SJ, Gebuhr TC, Pan H, Svoboda P, Schultz RM & Magnuson T (2006) Maternal BRG1
669 regulates zygotic genome activation in the mouse. *Genes Dev* **20:** 1744–1754
670 Chen Z & Zhang Y (2019) Loss of DUX causes minor defects in zygotic genome activation and is
671 compatible with mouse development. *Nat. Genet.* **51:** 1
672 Chuong EB, Elde NC & Feschotte C (2017) Regulatory activities of transposable elements: from conflicts
673 to benefits. *Nat. Rev. Genet.* **18:** 71–86 Available at:
674 https://www.ncbi.nlm.nih.gov/pubmed/27867194
675 Crosby IM, Gandolfi F & Moor RM (1988) Control of protein synthesis during early cleavage of sheep
676 embryos. *J. Reprod. Fertil.* **82:** 769–775
677 Davis DL (1985) Culture and storage of pig embryos. *J. Reprod. Fertil. Suppl.* **33:** 115–124 Available at:
678 https://www.ncbi.nlm.nih.gov/pubmed/3910819
679 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M & Gingeras TR
680 (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29:** 15–21
681 Du Z, Zheng H, Huang B, Ma R, Wu J, Zhang X, He J, Xiang Y, Wang Q, Li Y, Ma J, Zhang X, Zhang
682 K, Wang Y, Zhang MQ, Gao J, Dixon JR, Wang X, Zeng J & Xie W (2017) Allelic reprogramming
683 of 3D chromatin architecture during early mammalian development. *Nature* **547:** 232–235 Available
684 at: https://www.ncbi.nlm.nih.gov/pubmed/28703188
685 Eckersley-Maslin M, Alda-Catalinas C, Blotenburg M, Kreibich E, Krueger C & Reik W (2019) Dppa2
686 and Dppa4 directly regulate the Dux-driven zygotic transcriptional program. *Genes Dev.* **33:** 194–
687 208
688 Eckersley-Maslin MA, Alda-Catalinas C & Reik W (2018) Dynamics of the epigenetic landscape during
689 the maternal-to-zygotic transition. *Nat. Rev. Mol. Cell Biol.* **19:** 436–450
690 Falco G, Lee S-L, Stanghellini I, Bassey UC, Hamatani T & Ko MSH (2007) Zscan4: a novel gene

691 expressed exclusively in late 2-cell embryos and embryonic stem cells. *Dev. Biol.* **307:** 539–550
692 Available at: https://www.ncbi.nlm.nih.gov/pubmed/17553482
693 Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, Modi BP, Correard S,
694 Gheorghe M, Baranašić D, Santana-Garcia W, Tan G, Chèneby J, Ballester B, Parcy F, Sandelin A,
695 Lenhard B, Wasserman WW & Mathelier A (2019) JASPAR 2020: update of the open-access
696 database of transcription factor binding profiles. *Nucleic Acids Res.* Available at:
697 https://doi.org/10.1093/nar/gkz1001
698 Franke V, Ganesh S, Karlic R, Malik R, Pasulka J, Horvat F, Kuzman M, Fulka H, Cernohorska M,
699 Urbanova J, Svobodova E, Ma J, Suzuki Y, Aoki F, Schultz RM, Vlahovicek K & Svoboda P (2017)
700 Long terminal repeats power evolution of genes and gene expression programs in mammalian
701 oocytes and zygotes. *Genome Res.* **27:** 1384–1394
702 Frei RE, Schultz GA & Church RB (1989) Qualitative and quantitative changes in protein synthesis occur
703 at the 8-16-cell stage of embryogenesis in the cow. *J. Reprod. Fertil.* **86:** 637–641 Available at:
704 https://www.ncbi.nlm.nih.gov/pubmed/2760892
705 Friedli M & Trono D (2015) The developmental control of transposable elements and the evolution of
706 higher species. *Annu. Rev. Cell Dev. Biol.* **31:** 429–451 Available at:
707 https://www.ncbi.nlm.nih.gov/pubmed/26393776
708 Funaya S & Aoki F (2017) Regulation of zygotic gene activation by chromatin structure and epigenetic
709 factors. *J. Reprod. Dev.* **63:** 359–363 Available at: https://www.ncbi.nlm.nih.gov/pubmed/28579579
710 Gao L, Wu K, Liu Z, Yao X, Yuan S, Tao W, Yi L, Yu G, Hou Z, Fan D, Tian Y, Liu J, Chen Z-J & Liu J
711 (2018) Chromatin Accessibility Landscape in Human Early Embryos and Its Association with
712 Evolution. *Cell* **173:** 248-259.e15
713 Ghirlando R & Felsenfeld G (2016) CTCF: making the right connections. *Genes Dev* **30:** 881–891
714 Di Giammartino DC, Kloetgen A, Polyzos A, Liu Y, Kim D, Murphy D, Abuhashem A, Cavaliere P,
715 Aronson B, Shah V, Dephoure N, Stadtfeld M, Tsirigos A & Apostolou E (2019) KLF4 is involved
716 in the organization and regulation of pluripotency-associated three-dimensional enhancer networks.

717 *Nat. Cell Biol.* **21:** 1179–1190 Available at: https://doi.org/10.1038/s41556-019-0390-6

718 Graf A, Krebs S, Heininen-Brown M, Zakhartchenko V, Blum H & Wolf E (2014) Genome activation in
719 bovine embryos: review of the literature and new insights from RNA sequencing experiments. *Anim*
720 *Reprod Sci* **149:** 46–58

721 Graf T (2011) Historical origins of transdifferentiation and reprogramming. *Cell Stem Cell* **9:** 504–516
722 Available at: https://www.ncbi.nlm.nih.gov/pubmed/22136926

723 Guo M, Zhang Y, Zhou J, Bi Y, Xu J, Xu C, Kou X, Zhao Y, Li Y, Tu Z, Liu K, Lin J, Yang P, Gao S &
724 Wang Y (2019) Precise temporal regulation of Dux is important for embryo development. *Cell Res.*
725 Available at: https://doi.org/10.1038/s41422-019-0238-4

726 Hancks DC & Kazazian Jr HH (2016) Roles for retrotransposon insertions in human disease. *Mob. DNA*
727 **7:** 9 Available at: https://www.ncbi.nlm.nih.gov/pubmed/27158268

728 Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H & Glass CK
729 (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory
730 elements required for macrophage and B cell identities. *Mol Cell* **38:** 576–589

731 Hendrickson PG, Doráis JA, Grow EJ, Whiddon JL, Lim J-W, Wike CL, Weaver BD, Pflueger C, Emery
732 BR, Wilcox AL, Nix DA, Peterson CM, Tapscott SJ, Carrell DT & Cairns BR (2017) Conserved
733 roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL
734 retrotransposons. *Nat. Genet.* **49:** 925–934 Available at:
735 https://www.ncbi.nlm.nih.gov/pubmed/28459457

736 Hernandez C, Wang Z, Ramazanov B, Tang Y, Mehta S, Dambrot C, Lee Y-W, Tessema K, Kumar I,
737 Astudillo M, Neubert TA, Guo S & Ivanova NB (2018) Dppa2/4 Facilitate Epigenetic Remodeling
738 during Reprogramming to Pluripotency. *Cell Stem Cell* **23:** 396-411.e8 Available at:
739 https://www.ncbi.nlm.nih.gov/pubmed/30146411

740 Huang DW, Sherman BT & Lempicki RA (2009a) Systematic and integrative analysis of large gene lists
741 using DAVID bioinformatics resources. *Nat. Protoc.* **4:** 44–57 Available at:
742 https://www.ncbi.nlm.nih.gov/pubmed/19131956

743 Huang DW, Sherman BT & Lempicki RA (2009b) Bioinformatics enrichment tools: paths toward the
744 comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37:** 1–13 Available at:
745 https://www.ncbi.nlm.nih.gov/pubmed/19033363
746 De Iaco A, Planet E, Coluccio A, Verp S, Duc J & Trono D (2017) DUX-family transcription factors
747 regulate zygotic genome activation in placental mammals. *Nat. Genet.* **49:** 941–945 Available at:
748 https://www.ncbi.nlm.nih.gov/pubmed/28459456
749 Iaco A De, Verp S, Offner S & Trono D (2019) DUX is a non-essential synchronizer of zygotic genome
750 activation. *bioRxiv***:** 569434 Available at:
751 http://biorxiv.org/content/early/2019/03/06/569434.abstract
752 Iturbide A & Torres-Padilla M-E (2017) Starting embryonic transcription for the first time. *Nat. Genet.*
753 **49:** 820–821 Available at: https://www.ncbi.nlm.nih.gov/pubmed/28546569
754 Iurlaro M, von Meyenn F & Reik W (2017) DNA methylation homeostasis in human and mouse
755 development. *Curr. Opin. Genet. Dev.* **43:** 101–109
756 Jachowicz JW, Bing X, Pontabry J, Bošković A, Rando OJ & Torres-Padilla M-E (2017) LINE-1
757 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat.*
758 *Genet.* **49:** 1502 Available at: https://doi.org/10.1038/ng.3945
759 Jimenez R, Melo EO, Davydenko O, Ma J, Mainigi M, Franke V & Schultz RM (2015) Maternal SIN3A
760 regulates reprogramming of gene expression during mouse preimplantation development. *Biol.*
761 *Reprod.* **93:** 89 Available at: https://www.ncbi.nlm.nih.gov/pubmed/26353893
762 Ko MSH (2016) Zygotic Genome Activation Revisited: Looking Through the Expression and Function of
763 Zscan4. *Curr. Top. Dev. Biol.* **120:** 103–124 Available at:
764 https://www.ncbi.nlm.nih.gov/pubmed/27475850
765 Koehler D, Zakhartchenko V, Froenicke L, Stone G, Stanyon R, Wolf E, Cremer T & Brero A (2009)
766 Changes of higher order chromatin arrangements during major genome activation in bovine
767 preimplantation embryos. *Exp. Cell Res.* **315:** 2053–2063 Available at:
768 https://www.ncbi.nlm.nih.gov/pubmed/19254712

769 Kruse K, Díaz N, Enriquez-Gasca R, Gaume X, Torres-Padilla M-E & Vaquerizas JM (2019)
770 Transposable elements drive reorganisation of 3D chromatin during early embryogenesis. *bioRxiv***:**
771 523712 Available at: http://biorxiv.org/content/early/2019/01/17/523712.abstract
772 Latham KE, Garrels JI, Chang C & Solter D (1991) Quantitative analysis of protein synthesis in mouse
773 embryos. I. Extensive reprogramming at the one- and two-cell stages. *Development* **112:** 921 LP –
774 932 Available at: http://dev.biologists.org/content/112/4/921.abstract
775 Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT & Carey VJ (2013)
776 Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9:** e1003118–
777 e1003118 Available at: https://www.ncbi.nlm.nih.gov/pubmed/23950696
778 Leidenroth A & Hewitt JE (2010) A family history of DUX4: phylogenetic analysis of DUXA, B, C and
779 Duxbl reveals the ancestral DUX gene. *BMC Evol. Biol.* **10:** 364 Available at:
780 https://www.ncbi.nlm.nih.gov/pubmed/21110847
781 Lerat E, Fablet M, Modolo L, Lopez-Maestre H & Vieira C (2017) TEtools facilitates big data expression
782 analysis of transposable elements and reveals an antagonism between their activity and that of
783 piRNA genes. *Nucleic Acids Res.* **45:** e17–e17 Available at:
784 https://www.ncbi.nlm.nih.gov/pubmed/28204592
785 Li H & Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform.
786 *Bioinformatics* **26:** 589–595
787 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R &
788 Subgroup 1000 Genome Project Data Processing (2009) The Sequence Alignment/Map format and
789 SAMtools. *Bioinformatics* **25:** 2078–2079 Available at:
790 https://www.ncbi.nlm.nih.gov/pubmed/19505943
791 Li L, Guo F, Gao Y, Ren Y, Yuan P, Yan L, Li R, Lian Y, Li J, Hu B, Gao J, Wen L, Tang F & Qiao J
792 (2018) Single-cell multi-omics sequencing of human early embryos. *Nat. Cell Biol.* **20:** 847–858
793 Available at: https://doi.org/10.1038/s41556-018-0123-2
794 Liu L, Leng L, Liu C, Lu C, Yuan Y, Wu L, Gong F, Zhang S, Wei X, Wang M, Zhao L, Hu L, Wang J,

795 Yang H, Zhu S, Chen F, Lu G, Shang Z & Lin G (2019) An integrated chromatin accessibility and
796 transcriptome landscape of human pre-implantation embryos. *Nat. Commun.* **10:** 364 Available at:
797 https://www.ncbi.nlm.nih.gov/pubmed/30664750
798 Liu X, Wang C, Liu W, Li J, Li C, Kou X, Chen J, Zhao Y, Gao H, Wang H, Zhang Y, Gao Y & Gao S
799 (2016) Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation
800 embryos. *Nature* **537:** 558–562
801 Lu F, Liu Y, Inoue A, Suzuki T, Zhao K & Zhang Y (2016) Establishing Chromatin Regulatory
802 Landscape during Mouse Preimplantation Development. *Cell* **165:** 1375–1388
803 Lubelsky Y, Reuven N & Shaul Y (2005) Autorepression of Rfx1 Gene Expression: Functional
804 Conservation from Yeast to Humans in Response to DNA Replication Arrest. *Mol. Cell. Biol.* **25:**
805 10665–10673 Available at: http://mcb.asm.org/
806 Marin M, Karis A, Visser P, Grosveld F & Philipsen S (1997) Transcription Factor Sp1 Is Essential for
807 Early Embryonic Development but Dispensable for Cell Growth and Differentiation. *Cell* **89:** 619–
808 628 Available at: http://www.sciencedirect.com/science/article/pii/S0092867400802433
809 Martens JH, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P & Jenuwein T (2005) The
810 profile of repeat-associated histone lysine methylation states in the mouse epigenome. *Embo j* **24:**
811 800–812
812 Martin C, Beaujean N, Brochard V, Audouard C, Zink D & Debey P (2006) Genome restructuring in
813 mouse embryos during reprogramming and early development. *Dev. Biol.* **292:** 317–332 Available
814 at: http://www.sciencedirect.com/science/article/pii/S0012160606000157
815 Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17:**
816 Available at: http://journal.embnet.org/index.php/embnetjournal/article/view/200
817 Nishikimi A, Mukai J & Yamada M (1999) Nuclear translocation of nuclear factor kappa B in early 1-cell
818 mouse embryos. *Biol. Reprod.* **60:** 1536–1541 Available at:
819 https://www.ncbi.nlm.nih.gov/pubmed/10330116
820 Oldfield AJ, Yang P, Conway AE, Cinghu S, Freudenberg JM, Yellaboina S & Jothi R (2014) Histone-

821 fold domain protein NF-Y promotes chromatin accessibility for cell type-specific master
822 transcription factors. *Mol. Cell* **55:** 708–722 Available at:
823 https://www.ncbi.nlm.nih.gov/pubmed/25132174
824 Ooga M, Fulka H, Hashimoto S, Suzuki MG & Aoki F (2016) Analysis of chromatin structure in mouse
825 preimplantation embryos by fluorescent recovery after photobleaching. *Epigenetics* **11:** 85–94
826 Available at: https://www.ncbi.nlm.nih.gov/pubmed/26901819
827 Paciolla M, Boni R, Fusco F, Pescatore A, Poeta L, Ursini M V, Lioi MB & Miano MG (2011) Nuclear
828 factor-kappa-B-inhibitor alpha (NFKBIA) is a developmental marker of NF-kappaB/p65 activation
829 during in vitro oocyte maturation and early embryogenesis. *Hum. Reprod.* **26:** 1191–1201
830 Payer B, Saitou M, Barton SC, Thresher R, Dixon JPC, Zahn D, Colledge WH, Carlton MBL, Nakano T
831 & Surani MA (2003) Stella is a maternal effect gene required for normal early development in mice.
832 *Curr. Biol.* **13:** 2110–2117 Available at: https://www.ncbi.nlm.nih.gov/pubmed/14654002
833 Perea-Gomez A, Lawson KA, Rhinn M, Zakin L, Brûlet P, Mazan S & Ang SL (2001) Otx2 is required
834 for visceral endoderm movement and for the restriction of posterior signals in the epiblast of the
835 mouse embryo. *Development* **128:** 753–765 Available at:
836 https://www.ncbi.nlm.nih.gov/pubmed/11171400
837 Pontis J, Planet E, Offner S, Turelli P, Duc J, Coudray A, Theunissen TW, Jaenisch R & Trono D (2019)
838 Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic Genome
839 Activation and Control Transcription in Naive Human ESCs. *Cell Stem Cell* **24:** 724-735.e5
840 Probst A V, Okamoto I, Casanova M, El Marjou F, Le Baccon P & Almouzni G (2010) A strand-specific
841 burst in transcription of pericentric satellites is required for chromocenter formation and early mouse
842 development. *Dev Cell* **19:** 625–638
843 Probst A V, Santos F, Reik W, Almouzni G & Dean W (2007) Structural differences in centromeric
844 heterochromatin are spatially reconciled on fertilisation in the mouse zygote. *Chromosoma* **116:**
845 403–415 Available at: https://www.ncbi.nlm.nih.gov/pubmed/17447080
846 Quinlan AR & Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features.

847 *Bioinformatics* **26:** 841–842 Available at: https://www.ncbi.nlm.nih.gov/pubmed/20110278
848 Ram PT & Schultz RM (1993) Reporter gene expression in G2 of the 1-cell mouse embryo. *Dev. Biol.*
849 **156:** 552–556 Available at: https://www.ncbi.nlm.nih.gov/pubmed/8462750
850 Ramírez F, Dündar F, Diehl S, Grüning BA & Manke T (2014) deepTools: a flexible platform for
851 exploring deep-sequencing data. *Nucleic Acids Res.* **42:** W187–W191 Available at:
852 https://www.ncbi.nlm.nih.gov/pubmed/24799436
853 Raviram R, Rocha PP, Luo VM, Swanzey E, Miraldi ER, Chuong EB, Feschotte C, Bonneau R & Skok
854 JA (2018) Analysis of 3D genomic interactions identifies candidate host genes that transposable
855 elements potentially regulate. *Genome Biol.* **19:** 216 Available at: https://doi.org/10.1186/s13059-
856 018-1598-7
857 Ross PJ, Ragina NP, Rodriguez RM, Iager AE, Siripattarapravat K, Lopez-Corrales N & Cibelli JB
858 (2008) Polycomb gene expression and histone H3 lysine 27 trimethylation changes during bovine
859 preimplantation development. *Reproduction* **136:** 777–785
860 Schoorlemmer J, Pérez-Palacios R, Climent M, Guallar D & Muniesa P (2014) Regulation of Mouse
861 Retroelement MuERV-L/MERVL Expression by REX1 and Epigenetic Control of Stem Cell
862 Potency. *Front. Oncol.* **4:** 14 Available at: https://www.ncbi.nlm.nih.gov/pubmed/24567914
863 Schulz KN & Harrison MM (2019) Mechanisms regulating zygotic genome activation. *Nat. Rev. Genet.*
864 **20:** 221–234
865 Shirai A, Kawaguchi T, Shimojo H, Muramatsu D, Ishida-Yonetani M, Nishimura Y, Kimura H,
866 Nakayama J-I & Shinkai Y (2017) Impact of nucleic acid and methylated H3K9 binding activities of
867 Suv39h1 on its heterochromatin assembly. *Elife* **6:** e25317 Available at:
868 https://www.ncbi.nlm.nih.gov/pubmed/28760201
869 Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M & Zaret KS (2015) Pioneer transcription
870 factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161:** 555–568
871 Available at: https://www.ncbi.nlm.nih.gov/pubmed/25892221
872 Spadafora C (2004) Endogenous reverse transcriptase: a mediator of cell proliferation and differentiation.

873 *Cytogenet. Genome Res.* **105:** 346–350 Available at:
874 https://www.ncbi.nlm.nih.gov/pubmed/15237222
875 Takahashi K, Ross PJ & Sawai K (2019) The necessity of ZSCAN4 for preimplantation development and
876 gene expression of bovine embryos. *J. Reprod. Dev.* **65:** 319–326 Available at:
877 https://www.ncbi.nlm.nih.gov/pubmed/31019155
878 Takahashi K & Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult
879 fibroblast cultures by defined factors. *Cell* **126:** 663–676 Available at:
880 https://www.ncbi.nlm.nih.gov/pubmed/16904174
881 Vastenhouw NL, Cao WX & Lipshitz HD (2019) The maternal-to-zygotic transition revisited.
882 *Development* **146:**
883 Whiddon JL, Langford AT, Wong C-J, Zhong JW & Tapscott SJ (2017) Conservation and innovation in
884 the DUX4-family gene network. *Nat. Genet.* **49:** 935–940 Available at:
885 https://www.ncbi.nlm.nih.gov/pubmed/28459454
886 White CR, MacDonald WA & Mann MRW (2016) Conservation of DNA Methylation Programming
887 Between Mouse and Human Gametes and Preimplantation Embryos. *Biol. Reprod.* **95:** 61
888 Wiekowski M, Miranda M & DePamphilis ML (1993) Requirements for promoter activity in mouse
889 oocytes and embryos distinguish paternal pronuclei from maternal and zygotic nuclei. *Dev. Biol.*
890 **159:** 366–378 Available at: https://www.ncbi.nlm.nih.gov/pubmed/8365573
891 Wu J, Huang B, Chen H, Yin Q, Liu Y, Xiang Y, Zhang B, Liu B, Wang Q, Xia W, Li W, Li Y, Ma J,
892 Peng X, Zheng H, Ming J, Zhang W, Zhang J, Tian G, Xu F, et al (2016) The landscape of
893 accessible chromatin in mammalian preimplantation embryos. *Nature* **534:** 652–657
894 Wu J, Xu J, Liu B, Yao G, Wang P, Lin Z, Huang B, Wang X, Li tong, Shi S, Zhang N, Duan F, Ming J,
895 Zhang X, Niu W, Song W, Jin H, Guo Y, Dai S, Hu L, et al (2018) Chromatin analysis in human
896 early development reveals epigenetic transition during ZGA. Available at:
897 https://doi.org/10.1038/s41586-018-0080-8
898 Wu X, Li Y, Xue L, Wang L, Yue Y, Li K, Bou S, Li G-P & Yu H (2011) Multiple histone site epigenetic

899 modifications in nuclear transfer and in vitro fertilized bovine embryos. *Zygote* **19:** 31–45 Available
900 at: https://www.ncbi.nlm.nih.gov/pubmed/20609268
901 Xu Q & Xie W (2018) Epigenome in Early Mammalian Development: Inheritance, Reprogramming and
902 Establishment. *Trends Cell Biol.* **28:** 237–253 Available at:
903 https://www.ncbi.nlm.nih.gov/pubmed/29217127
904 Yandım C & Karakülah G (2019) Expression dynamics of repetitive DNA in early human embryonic
905 development. *BMC Genomics* **20:** 439 Available at:
906 https://www.ncbi.nlm.nih.gov/pubmed/31151386
907 Yang Y, Zhou C, Wang Y, Liu W, Liu C, Wang L, Liu Y, Shang Y, Li M, Zhou S, Wang Y, Zeng W,
908 Zhou J, Huo R & Li W (2017) The E3 ubiquitin ligase RNF114 and TAB1 degradation are required
909 for maternal-to-zygotic transition. *EMBO Rep.* **18:** 205–216 Available at:
910 https://www.ncbi.nlm.nih.gov/pubmed/28073917
911 Young JM, Whiddon JL, Yao Z, Kasinathan B, Snider L, Geng LN, Balog J, Tawil R, van der Maarel SM
912 & Tapscott SJ (2013) DUX4 binding to retroelements creates promoters that are active in FSHD
913 muscle and testis. *PLoS Genet.* **9:** e1003947–e1003947 Available at:
914 https://www.ncbi.nlm.nih.gov/pubmed/24278031
915 Zhang A, Xu B, Sun Y, Lu X, Gu R, Wu L, Feng Y & Xu C (2012) Dynamic changes of histone H3
916 trimethylated at positions K4 and K27 in human oocytes and preimplantation embryos. *Fertil. Steril.*
917 **98:** 1009–1016
918 Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, Ming J, Wu X, Zhang Y, Xu Q, Liu W, Kou X,
919 Zhao Y, He W, Li C, Chen B, Li Y, Wang Q, Ma J, Yin Q, et al (2016) Allelic reprogramming of
920 the histone modification H3K4me3 in early mammalian development. *Nature* **537:** 553–557
921 Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M,
922 Li W & Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9:** R137
923
924

925 **Figure legends**

926 Figure 1. Chromatin accessibility in bovine oocytes and in vitro preimplantation embryos. A) Schematic
927 of in vitro embryo production and ATAC-seq library preparation. B) Principal components analysis
928 (PCA) of ATAC-seq read depth normalized by reads per kilobase million (RPKM) in 50 bp windows
929 covering the whole genome. C) Normalized coverage (RPKM) of replicate ATAC-seq libraries for each
930 stage of development. D) Proportion of genome that is covered by genic and intergenic ATAC-seq peaks,
931 called from 30 million reads at each developmental stage. E) Categorization of accessible regions in 2-, 4-
932 , and 8-cell embryos into stage-specific and maintained peaks (accessibility maintained up until the
933 morula stage). Maintained peaks are carried over from latter stages to show cumulative maintained peaks.
934

935 Figure 2. Gradual establishment of open chromatin enriched in regulatory motifs. A) Regions that
936 significantly lost accessibility from the GV oocyte to 2-cell stage, and regions that significantly gained
937 accessibility from the 2- to 4-cell, 4- to 8-cell, and 8-cell to morula stages, according to 30 million reads
938 per stage. Accessibility at each region, scaled to average width ± 1kb, was determined by normalized
939 ATAC-seq read depth (reads per kilobase million; RPKM), based on 20 million reads per stage. B) Top
940 five enriched *de novo* motifs enriched in intergenic peaks with matching known motifs (match score >
941 0.6) of TFs that were expressed at the given stage. C) Top seven known motifs enriched in 4-cell specific
942 peaks. D) DESeq2 normalized expression of TFs corresponding to enriched motifs in 4-cell specific open
943 chromatin. E) Top seven known motifs enriched in 8-cell specific peaks. F) DESeq2 normalized
944 expression of TFs corresponding to enriched motifs in 8-cell specific open chromatin. G) Proportion of
945 intergenic 8-cell peaks overlapping the de novo motif most closely matching the DUXA motif, relative to
946 background. H) DESeq2-normalized expression of DUXA throughout development. RNA-seq data from
947 Graf *et al* (2014).
948

949 Figure 3. Binding motif enrichment in open chromatin opened at the 2-, 4-, or 8-cell stages and
950 maintained up to morula stage. Proportion of peaks with A) CTCF or B) KLF5, binding motifs. C)

951 DESeq2-normalized expression of KLF genes. RNA-seq data from Graf *et al* (2014). D) Proportion of
952 maintained and stage-specific peaks that were genic or intergenic. Maintained peaks were established at
953 the indicated stage and maintained up to the morula stage. E) Gene ontology term enrichment of the 9,456
954 genes that were marked by maintained genic open chromatin that was first established in 2- or 4-cell
955 embryos.
956
957 Figure 4. Effect of transcription inhibition on chromatin remodeling. A) Transcription inhibition effect on
958 loci that should have opened or closed between consecutive stages. Peaks called based on 20 million
959 reads per stage and categorized as either genic or intergenic. B) Normalized read depth (RPKM) at loci
960 which opened during the 4- to 8-cell transition in control 8-cell embryos and transcriptionally inhibited 8-
961 cell embryos. Regions scaled to average width ± 1 kb. C) Normalized ATAC-seq and RNA-seq signal
962 (RPKM) in 8-cell and 8-cell TBE at the KLF4 locus. RNA-seq data from Bogliotti *et al* (2019). D)
963 Proportion of stage-specific and maintained peaks that should appear at each stage, but which fail to open
964 in TBE.
965
966 Figure 5. Expression and accessibility dynamics of repeat families during bovine preimplantation
967 development. A) DESeq2 normalized expression profiles of select LTR, LINE, and SINE repeat families.
968 RNA-seq data from Graf *et al* (2014). B) Enrichment of several transposable element (TE) families in
969 ATAC-seq peaks, called from 30M reads. Promoter peaks fell within the 2kb region upstream of TSS.
970 Intergenic peaks did not overlap the 2kb regions upstream of TSS, exons, or introns.
971
972 Figure 6. Activity of LTR elements. A) Open chromatin enrichment of LTR elements with increased
973 accessibility in 4-cell, 8-cell, and/or morulae, and VST-normalized expression of the same LTR elements;
974 three replicates per stage. B) ATAC-seq and RNA-seq read coverage at a highly expressed ERV1-2 repeat
975 sequence. ATAC-seq tracks show coverage from 30 million reads per library; RNA-seq tracks show
976 combined coverage from three replicates. TF motifs predicted from JASPAR binding motifs MA0039.3

977 (KLF4) and MA0712.1 (OTX2), respectively. C) Average normalized ATAC-seq and D) RNA-seq signal
978 (RPKM) at MLT1A0 repeats overlapped by 8-cell intergenic open chromatin harboring DUXA motifs,
979 predicted based on JASPAR motif MA0468.1. E) Co-option of an accessible MLT1A0 element with a
980 DUXA binding motif as an alternative promoter upstream of the C1D locus. RNA-seq data from Graf *et*
981 *al* (2014).
982
983 Figure 7. Inference of key regulatory factors during EGA in cattle, human, and mouse, based on
984 enrichment for TF binding factors in open chromatin and expression of the corresponding TF. Bovine
985 RNA-seq data from Graf *et al* (2014); human ATAC-seq and RNA-seq data from Wu *et al* (2018); mouse
986 ATAC-seq and RNA-seq from Wu *et al* (2016).
987
988 Figure 8. Potential mechanistic model depicting events leading to major EGA. Chromatin structure is
989 globally decondensed following fertilization, allowing opportunistic binding of maternal factors, which
990 initiate a minor wave of transcription and begin to establish 3D chromatin architecture. This sets the stage
991 for major EGA, wherein maternal products, minor EGA products, and promoter-enhancer contacts
992 mediate the first major wave of gene expression and continue to refine 3D chromatin structure.
993
994
995
996
997
998
999
1000
1001
1002

Figure 1.

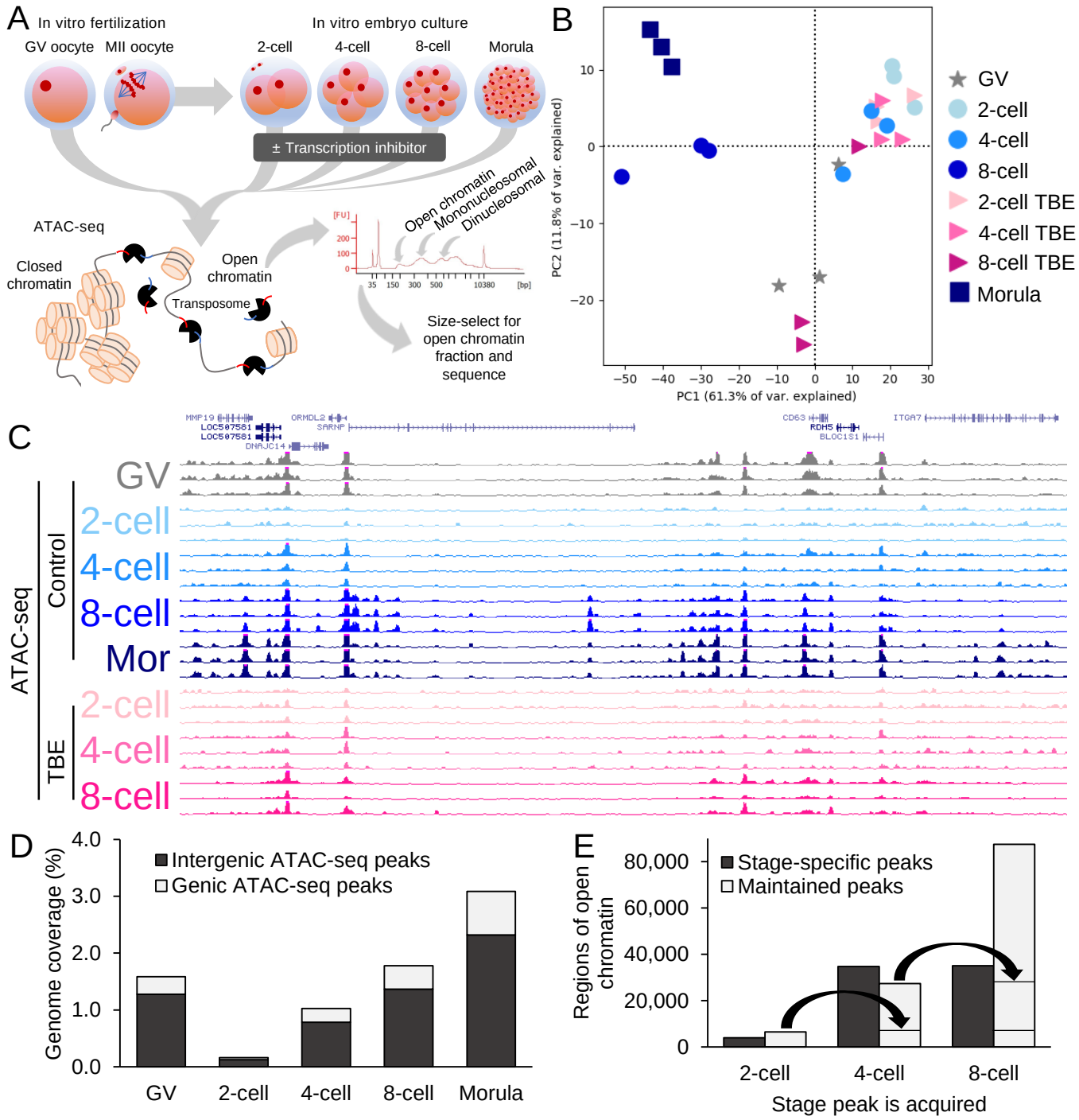
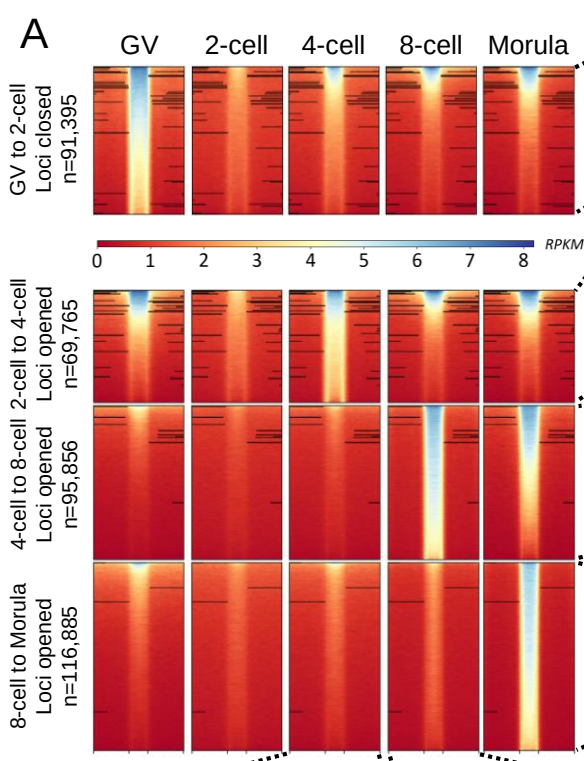

A
In vitro fertilization
GV oocyte
MII oocyte
In vitro embryo culture
2-cell
4-cell
8-cell
Morula
± Transcription inhibitor
ATAC-seq
Closed chromatin
Open chromatin
Transposome
Open chromatin
Mononucleosomal
Dinucleosomal
Size-select for open chromatin fraction and sequence
B
PC2 (11.8% of var. explained)
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.



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

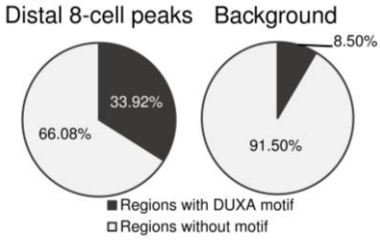
Motif	P-value	Peaks with motif (%)
NFkB-p65-Rel	1e-83	4.13
NFkB-p50,p52	1e-73	6.85
NFkB-p65	1e-67	30.46
NF-E2	1e-58	4.86
RFX5	1e-34	10.14
JUN-AP1	1e-28	7.47
BACH2	1e-27	6.66

E

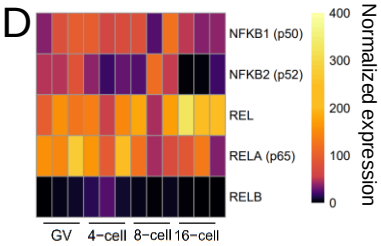
8-cell-specific open chromatin

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

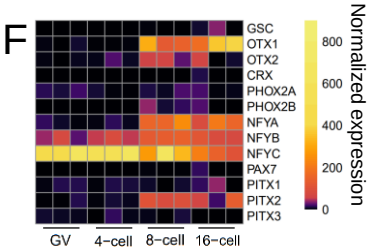
G



D



F



H

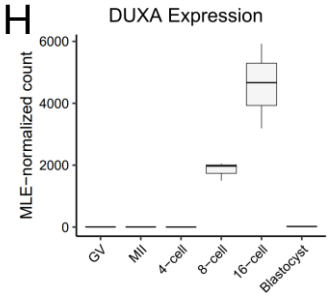


Figure 3.

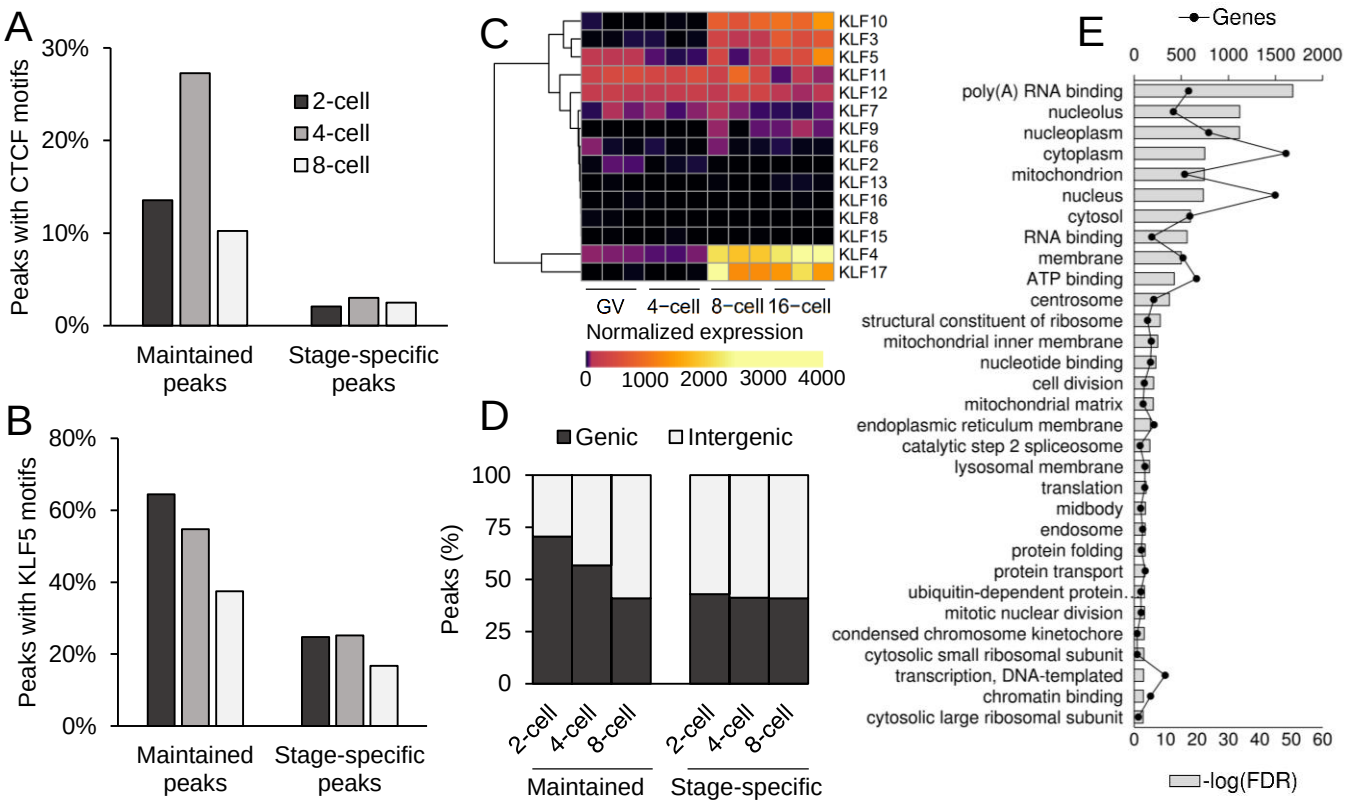

A
Peaks with CTCF motifs
30%
20%
10%
0%
2-cell
4-cell
8-cell
Maintained peaks
Stage-specific peaks
B
Peaks with KLF5 motifs
80%
60%
40%
20%
0%
Maintained peaks
Stage-specific peaks
C
KLF10
KLF3
KLF5
KLF11
KLF12
KLF7
KLF9
KLF6
KLF2
KLF13
KLF16
KLF8
KLF15
KLF4
KLF17
GV
4−cell
8−cell
16−cell
Normalized 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.

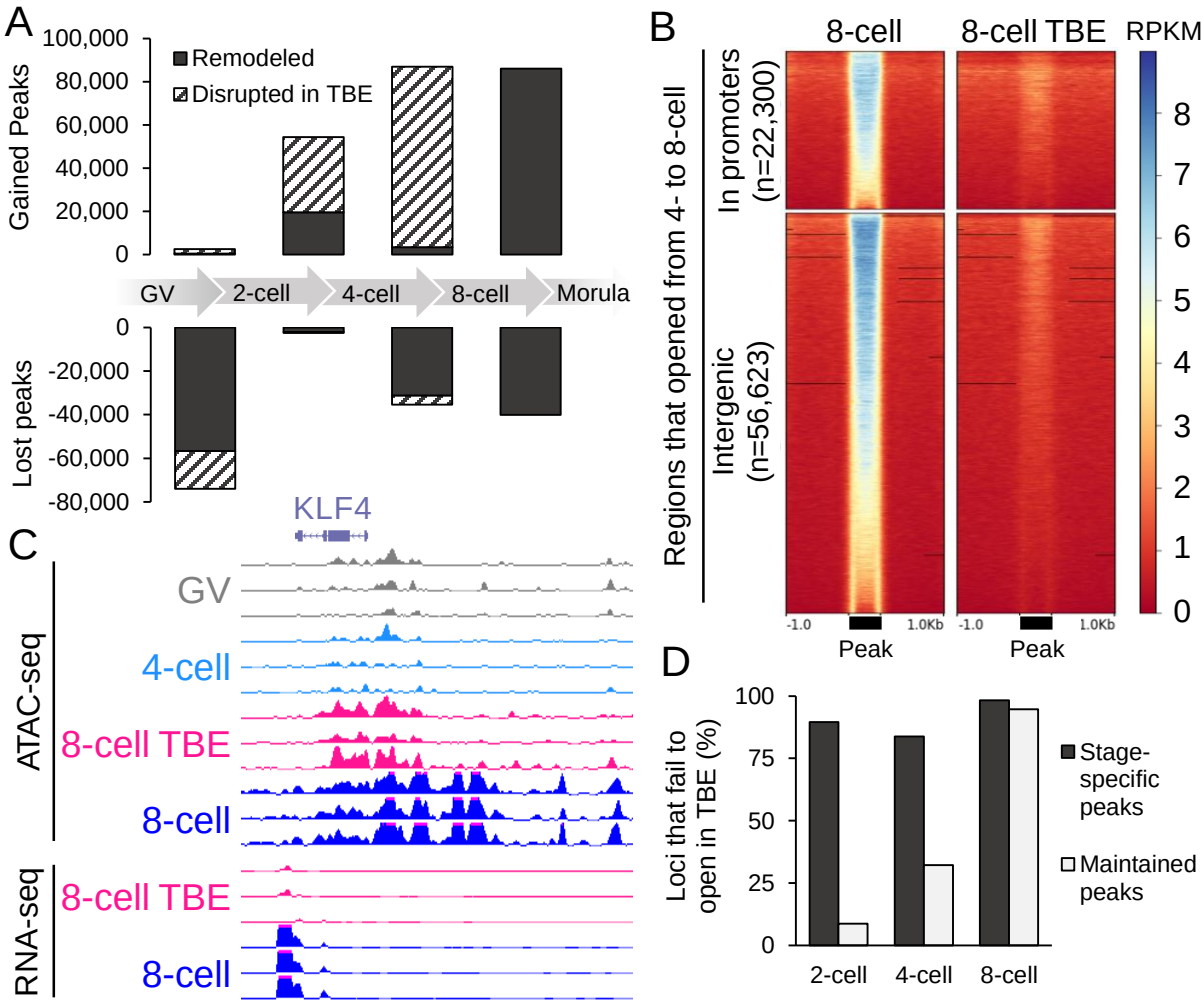

A
Gained Peaks
100,000
80,000
60,000
40,000
20,000
0
Remodeled
Disrupted in TBE
GV
2-cell
4-cell
8-cell
Morula
Lost peaks
0
-20,000
-40,000
-60,000
-80,000
B
8-cell
8-cell TBE
RPKM
Regions that opened from 4- to 8-cell
In promoters (n=22,300)
Intergenic (n=56,623)
8
7
6
5
4
3
2
1
0
-1.0
1.0Kb
Peak
C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

Figure 5.

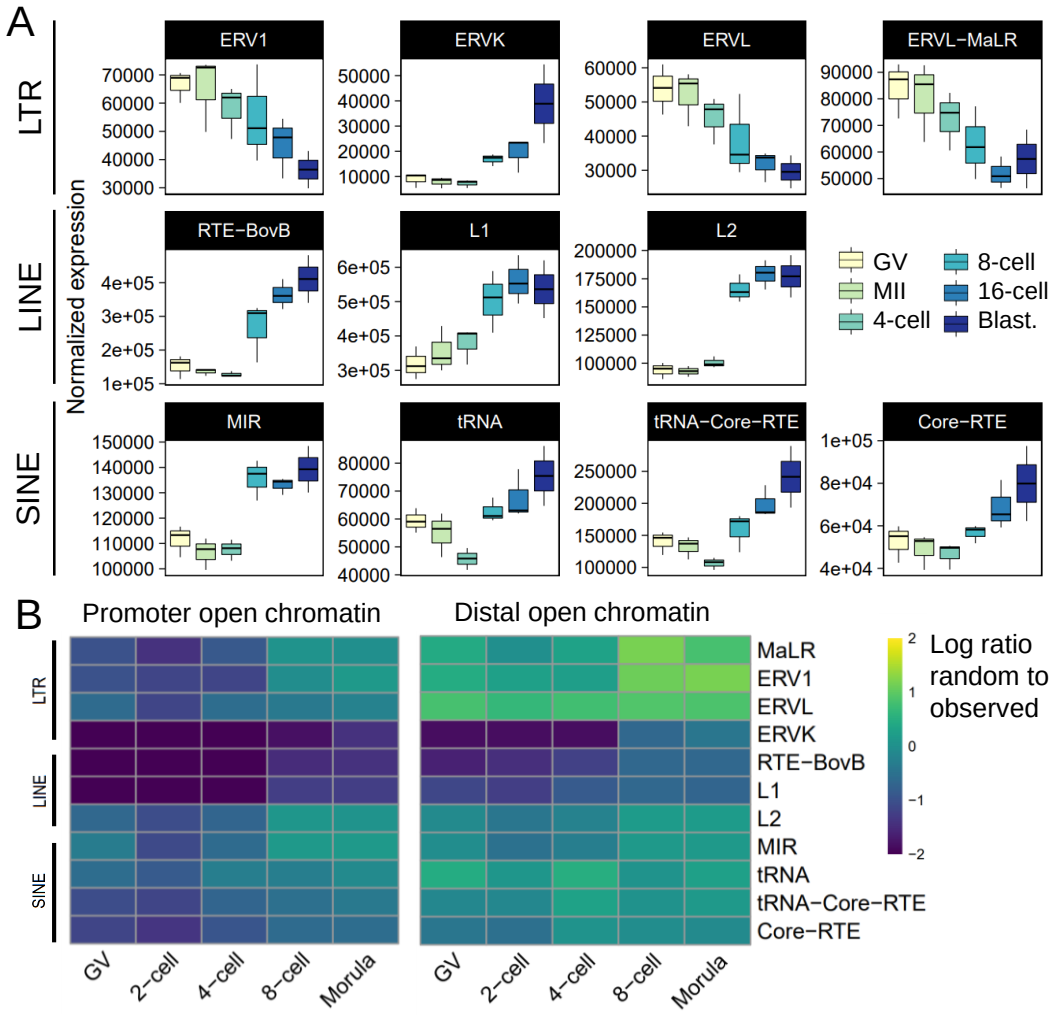

A
LTR
LINE
SINE
Normalized expression
ERV1
ERVK
ERVL
ERVL−MaLR
RTE−BovB
L1
L2
MIR
tRNA
tRNA−Core−RTE
Core−RTE
GV
MII
4-cell
8-cell
16-cell
Blast.
B
Promoter open chromatin
Distal open chromatin
MaLR
ERV1
ERVL
ERVK
RTE−BovB
L1
L2
MIR
tRNA
tRNA−Core−RTE
Core−RTE
GV
2−cell
4−cell
8−cell
Morula
Log ratio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

Figure 6.

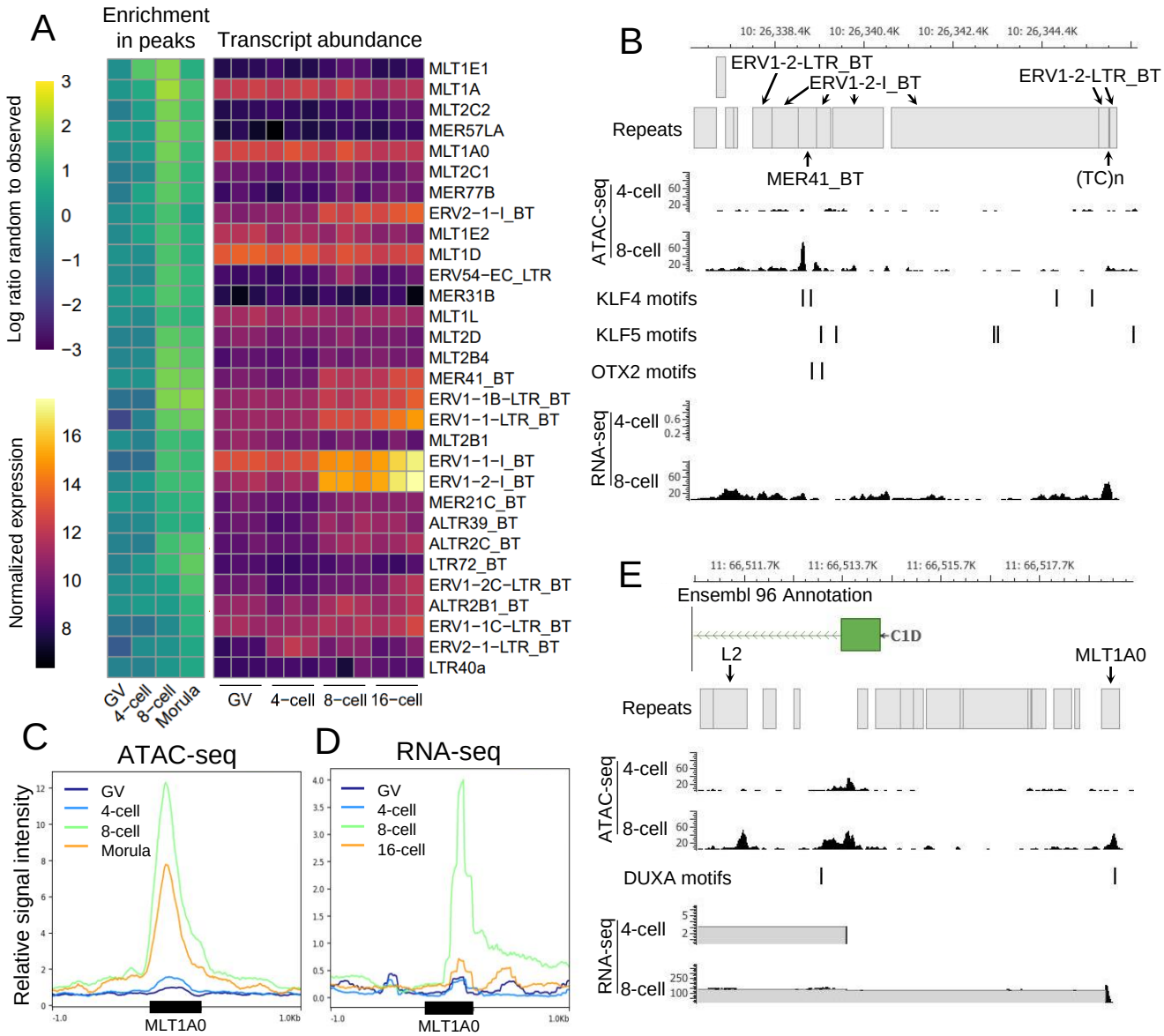

A
Enrichment in peaks
Transcript abundance
Log ratio random to observed
Normalized expression
MLT1E1
MLT1A
MLT2C2
MER57LA
MLT1A0
MLT2C1
MER77B
ERV2−1−I_BT
MLT1E2
MLT1D
ERV54−EC_LTR
MER31B
MLT1L
MLT2D
MLT2B4
MER41_BT
ERV1−1B−LTR_BT
ERV1−1−LTR_BT
MLT2B1
ERV1−1−I_BT
ERV1−2−I_BT
MER21C_BT
ALTR39_BT
ALTR2C_BT
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.

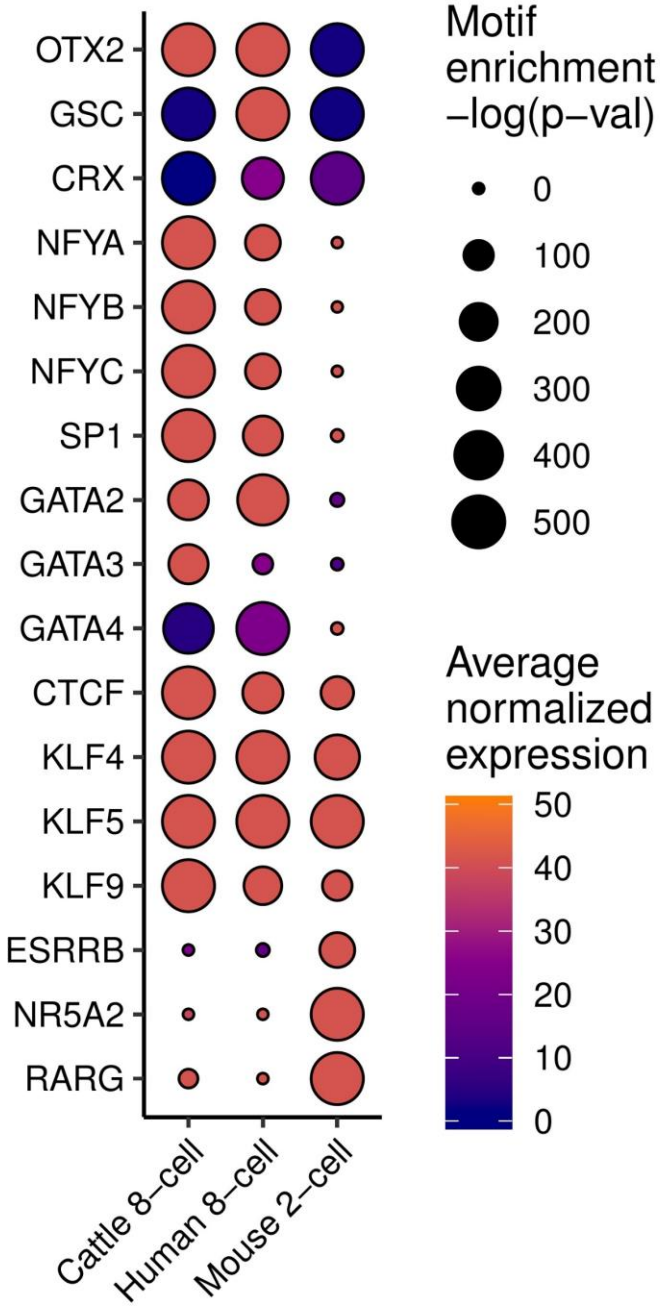
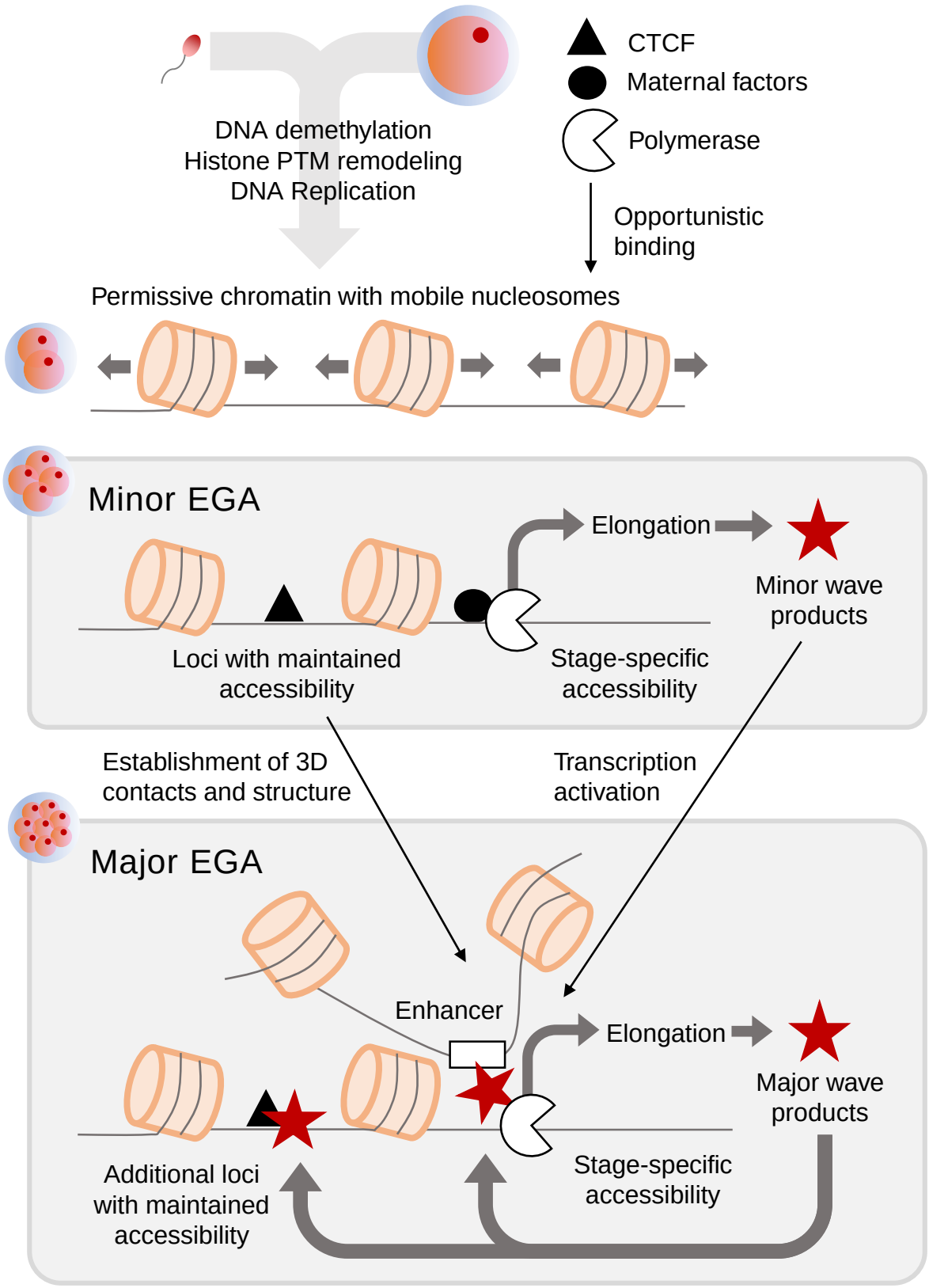

OTX2
GSC
CRX
NFYA
NFYB
NFYC
SP1
GATA2
GATA3
GATA4
CTCF
KLF4
KLF5
KLF9
ESRRB
NR5A2
RARG
Cattle 8–cell
Human 8–cell
Mouse 2–cell
Motif enrichment –log(p–val)
0
100
200
300
400
500
Average normalized expression
50
40
30
20
10
0

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