The transcription factor GABPA is a master regulator of naïve pluripotency

Chengjie Zhou[1,2,†], Meng Wang[1,2,†], Chunxia Zhang[1,2,†, #], Yi Zhang[1,2,3,4,5,*]

Affiliations:

[1]Howard Hughes Medical Institute, Boston Children's Hospital, Boston, MA 02115, USA.

[2]Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA 02115, USA.

[3]Division of Hematology/Oncology, Department of Pediatrics, Boston Children's Hospital, Boston, MA 02115, USA

[4]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA.

[5]Harvard Stem Cell Institute, Boston, MA 02115, USA.

* Corresponding author. Email: yzhang@genetics.med.harvard.edu

† These authors contributed equally to this work.

Current address: State Key Laboratory of Molecular Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China.

Abstract

Mouse preimplantation development harbors a series of molecular events that include zygotic genome activation (ZGA) and cell lineage specification which concurrent with totipotency to pluripotency transition. The establishment of pluripotency is a continuous process involving two cell lineage specification events, the generation of inner cell mass (ICM) and trophectoderm at E3.5 from morula embryo, and the generation of the epiblast (EPI) with naïve pluripotency and primitive endoderm (PrE) at E4.5 from ICM. Recent studies have revealed key transcription factors (TFs) for ICM formation, but the TFs for EPI initiation remain elusive. In this study, using a targeted rapid protein degradation system, we identified and demonstrated that the transcription factor GABPA is not only a regulator of major ZGA, but also a master EPI specifier required for naïve pluripotency establishment during E3.5 to E4.5 transition. Importantly, GABPA regulates 47% of EPI genes to control EPI formation. Chromatin binding dynamics analysis suggests that GABPA controls EPI formation at least partly by binding to the ICM gene promoters bound by the pluripotency regulators TFAP2C and SOX2 at E3.5 to establish naïve pluripotency at E4.5. Our study not only uncovers GABPA as a master pluripotency master regulator, but also support the notion that mammalian pluripotency establishment requires a dynamic and stepwise multi-TFs regulatory network.

Introduction

The development of mouse embryos starts with maternal-to-zygotic transition (MZT), which is accompanied by maternal RNA degradation and zygotic genome activation (ZGA) [1,2]. After a few rounds of cleavage, the totipotent embryos go through the first cell lineage differentiation to generate inner cell mass (ICM) and trophectoderm (TE) at E3.5 [3,4]. Subsequently, the ICM cells go through the second cell lineage differentiation to generate epiblast (EPI) and primitive endoderm (PrE) at E4.5 [5].

Transcription factors (TFs) play important roles in cell lineage specification and pluripotency acquisition. At E3.5, the ICM cells express both pluripotency factors such as *Oct4*, *Sox2* and *Nanog*, and PrE factors such as *Gata4* and *Gata6* [6]. After completion of the second cell fate specification at E4.5, the expression of lineage-specific TFs become restricted with *Gata4* and *Gata6* confined to the PrE, while *Nanog* and *Sox2* restricted to the EPI, marking the establishment of naïve pluripotency in EPI [6]. Given that embryonic stem cells (ESCs) cannot be fully established from E3.5 ICM [7], and E3.5 ICM cells have different transcriptome and chromatin accessibility from that of the E4.5 EPI [8], E3.5 ICM is considered to be at a 'prepluripotency' state [8]. Accumulating evidence support a continuous totipotency-prepluripotency-naïve pluripotency transition during preimplantation development. Recent studies support the transcription factors NR5A2 and TFAP2C mediate totipotency to pluripotency transition by activating prepluripotency genes [9]. However, neither of them is responsible for activating naïve pluripotency genes [10]. Moreover, although the pluripotency factors NANOG and SOX2 are essential for maintaining the naïve pluripotency state, they are not required for initiating the naïve pluripotency *in vivo* [11,12]. Thus, the TFs that drive ICM to naïve pluripotency transition remain elusive.

Identification of the TFs regulating pluripotency establishment *in vivo* is hindered by their potential roles in earlier developmental stages before the pluripotency establishment. Thus, conventional TF knockout mouse models could not separate their potential roles in regulating totipotency or E3.5 ICM formation from that in regulating pluripotent EPI formation at E4.5. The development of targeted protein degradation with degron technologies like AID [13] and dTAG [14,15], which enable rapid degradation of target proteins at a specific time-window, making the study of gene function at a specific developmental stage possible. In this study, by generating dTAG mice and combined

with RNA-seq and low-input CUT&RUN assay, we identified and demonstrated that the transcription factor GA repeat binding protein alpha (GABPA), encoded by the minor ZGA gene *Gabpa*, plays an essential role in naïve pluripotency establishment *in vivo*. Our stage-specific degradation results revealed that GABPA is not only important for major ZGA, but also critical for ICM to naïve pluripotency transition by regulating a large set of pluripotency genes through binding to their promoters. Importantly, we found that during E3.5 ICM to E4.5 ICM transition, the decreased binding of TFAP2C and SOX2 at the promoters of certain ICM genes concomitant with the increased binding of GABPA at the promoters of ICM genes, indicating a switch in the key TFs that regulate pluripotency genes expression. These results support a dynamic and stepwise regulatory model for naïve pluripotency establishment during pre-implantation development.

Results

Identification of GABPA as a potential pluripotency regulator and generation of its dTAG mice

To identify candidate TFs potentially involved in ICM to naïve EPI differentiation, we performed integrative analyses using public RNA-seq and ATAC-seq datasets of mouse pre-implantation embryos [16-19]. Compared to OCT4, NANOG, and SOX2 binding motif, GABPA binding motif is highly enriched in the open promoters of E4.5 ICM, suggesting GABPA may have a role in regulating E4.5 ICM formation (Fig. 1a). Consistent with a previous study [16], GABPA binding motif is not enriched in distal open chromatin (Fig. 1a). In addition, single-cell gene expression correlation analysis between various TFs and the ICM genes (higher expression in ICM than TE, fold change≥2, $q<0.05$, see method) or TE genes (higher expression in TE than ICM, fold change≥2, $q<0.05$, see method) (Supplementary Table 1) revealed that *Gabpa* expression is positively correlated with expression of ICM genes in blastocysts, with higher correlations at the late blastocyst stage similar to that of *Oct4* and *Nanog* (Fig. 1b), indicating potential involvement of *Gabpa* in pluripotency regulation. Furthermore, *Gabpa* transcription starts from zygote to early 2-cell stage, and reaches the highest at late 2-cell and 4-cell stages, and then become differentially expressed in ICM and TE at blastocyst stage (Fig. 1c). Immunostaining indicated that GABPA can

be detected from zygote to blastocyst (Fig. 1d). All these data are consistent with a potential role of GABPA in pluripotency regulation in pre-implantation embryos.

GABPA is a member of the ETS TFs family, which forms a tetrameric complex with GABPB to regulate the transcriptional activation [20]. Previous studies have shown that *Gabpa* knockout resulted in embryonic lethal before blastocyst stage [21], but its role in cell fate specification has not been investigated. To assess its precise function in pre-implantation development, particularly in ICM and EPI specification, we utilized the dTAG protein degradation system [14]. Given the potential instability of dTAG fusion protein [22], we first tested GABPA dTAG system in mouse ES cells (Extended Data Fig. 1a). Western blot analysis showed that the GABPA-FKBP-HA and the endogenous GABPA proteins are at similar levels (Fig. 1e), indicating the FKBP-HA fusion did not affect GABPA stability. Importantly, the presence of the small molecule dTAG13 resulted in complete GABPA-FKBP-HA degradation within 30 min (Fig. 1e). We thus proceeded to generate the mouse harboring the *Gabpa*^{dTAG} allele using the CRISPR technology [23]. After confirming the Fkbp-HA knock-in efficiency (9 of 20 blastocyst showed Fkbp-HA knock-in), we performed 2-cell embryo transplantation (Fig. 1f and Extended Data Fig. 1a). The offspring produced by *Gabpa*$^{dTAG/+}$ mating followed the Mendel's ratio (Extended Data Fig. 1b). Importantly, addition of dTAG13 to the cultured embryos can efficiently degrade GABPA at zygote, 2-cell, 4-cell, morula, and blastocyst stages in both short- and long-time window treatments (Fig. 1g and Extended Data Fig. 1c). Moreover, dTAG13 treatment did not affect embryo development of control embryos without the knock-in (Extended Data Fig. 1d). Collectively, these results demonstrate the successful generation of a GABPA-dTAG mouse model.

GABPA activates a group of major ZGA genes by binding to their promoters

Gabpa KO leads to embryonic lethal before blastocyst [21], but the mechanism underlying the lethality is not clear. To understand when and how GABPA degradation affects preimplantation development, we performed dTAG13 treatment at different time windows. Given that *Gabpa* starts to express after fertilization (Fig. 1c), we confirmed that GABPA is a minor ZGA gene with a detectable protein level at 10 hours post fertilization (hpf) (Fig. 1d). We therefore performed GABPA degradation starting at 6 hpf with continuous dTAG13 presence and checked the effect on embryo development (Fig. 2a, #1). Results showed that most embryos were arrested at 4-cell

124 or 8-cell stage (Fig. 2b, #1, arrows), suggesting GABPA plays important roles before this stage.
125 Since GABPA is highly expressed at late 2-cell stage (Fig. 1c, d), we hypothesized that GABPA
126 may regulate major ZGA. To test this possibility, we treated the embryos with dTAG13 from 6 to
127 42 hpf, which covers the time window of ZGA, then washed out dTAG13 (Fig. 2a, b, #2). We
128 found this treatment affected embryo development similarly to that when dTAG13 is continuously
129 present from 6-112 hpf. Moreover, GABPA degradation after major ZGA (42-112 hpf) resulted in
130 a much weaker phenotype (Fig. 2a, b, #3). These results suggest that GABPA may play an
131 important role in ZGA.

132 To study GABPA's role in regulating major ZGA, we treated the *Gabpa*$^{dTAG/dTAG}$ embryos with
133 dTAG13 from mid-zygote (6 hpf) to late 2-cell stage (30 hpf) and collected embryos at 30 hpf for
134 RNA-seq (Fig. 2c and Extended Data Fig. 2a). We found the treatment resulted in 589 down-
135 regulated and 78 up-regulated genes (Fig. 2c). GO term enrichment analysis revealed that the
136 down-regulated genes are involved in processes such as ribosome biogenesis, mitochondrial gene
137 expression, and rRNA processing etc. (Extended Data Fig. 2b, c), explaining the embryo arrest
138 phenotype. Further analysis of the effect on major ZGA genes identified 104 down-regulated genes,
139 indicating GABPA has a role in activating these major ZGA genes (Fig. 2c and Supplementary
140 Table 1, 2).

141 To determine whether GABPA directly binds and regulates the down-regulated genes, we
142 performed low-input CUT&RUN assay on GABPA [24]. We first tested GABPA CUT&RUN in
143 mouse ESCs and showed that the use of 500 ESCs generated similar results as that of using 20k
144 cells (Extended Data Fig. 3a, b). We then performed GABPA CUT&RUN using late 2-cell
145 embryos (Extended Data Fig. 3b). Analysis of the CUT&RUN data indicated that most of the
146 GABPA binding peaks were located in the nucleosome-depleted-regions (NDR) of promoters (Fig.
147 2d and Extended Data Fig. 3c). Interestingly, the GABPA binding regions were enriched for the
148 GABPA motif but not the motifs of other reported murine ZGA regulators such as OBOX, DUX
149 and NR5A2 [25-27], suggesting direct GABPA binding at these regions (Extended Data Fig. 3c). For
150 most (449 out of 589) of the down-regulated genes, a clear GABPA binding signal around their
151 transcription start sites (TSSs) can be observed, but no binding for the up-regulated genes (Fig. 2
152 e, f). Importantly, most (83 out of 104) of the down-regulated major ZGA genes also showed direct
153 GABPA promoter binding (Fig. 2f, g). The fact that GABPA degradation caused loss of promoter

154 binding, resulted in down-regulation of most GABPA direct targets (Extended Data Fig. 3d, e),
155 including some major ZGA genes (Fig. 2c, g), support that the minor ZGA gene *Gabpa* plays an
156 important role in activating a group of major ZGA genes by directly binding to their promoters.

157 **GABPA controls EPI specification by activating pluripotency genes**

158 Next, we asked whether GABPA has a role in the first cell lineage specification, the generation of
159 ICM and TE. To avoid ZGA defect, we treated *Gabpa*$^{dTAG/dTAG}$ embryos with dTAG13 starting
160 from 4-cell stage, and collected early (84 hpf) and late blastocyst (112 hpf) for immunostaining
161 (Fig. 3a). We found that neither ICM nor TE cell numbers in early blastocysts were affected by
162 the treatment (Fig. 3b, c), suggesting GABPA is not necessary for the first cell lineage specification.
163 However, for the late blastocyst, the EPI cells were dramatically decreased while the PrE cells
164 were not affected by the treatment (Fig. 3d, e). RNA-seq analysis of morula embryos revealed a
165 limited transcriptome change upon the treatment (Extended Data Fig. 4a-c and Supplementary
166 Table 3), consistent with no major role of GABPA from 4-cell to morula embryos. To avoid any
167 potential contribution from defects between 4-cell to early blastocyst to the EPI defects, we started
168 the treatment in early blastocyst when the first cell lineage specification has already finished, and
169 then checked the effect on EPI and PrE at late blastocysts (Fig. 3f). We observed a similar effect
170 of this treatment to that of the treatment started at 4-cell embryos (Fig. 3g, h). These results indicate
171 that GABPA has a direct role in EPI specification.

172 To understand how GABPA regulates EPI specification, we performed RNA-seq. To capture the
173 earlier molecular changes that lead to EPI defects, we collected the ICM at the mid-blastocyst
174 stage (100 hpf) when the number of EPI cells has not yet shown significant difference between
175 DMSO and dTAG13 treated groups (Extended Data Fig. 4d-f). Transcriptome analyses revealed
176 1,605 up- and 1,886 down-regulated genes in response to GABPA degradation (Fig. 3i, Extended
177 Data Fig. 4g and Supplementary Table 4). Since some pluripotency factors are already expressed
178 at E3.5 ICM when we started to degrade GABPA (Extended Data Fig. 4d), EPI could form at
179 middle blastocyst, but these EPI cells are defective in their transcriptome in response to GABPA
180 degradation (Fig. 3i). Importantly, 223 of the 1,886 (11.8%) down-regulated genes belong to the
181 ICM genes, while 246 of the 1,605 (15.3%) up-regulated genes belong to the TE genes (Fig. 3j).
182 Further analysis revealed that 27% of the early ICM genes and 29.2% of the late ICM genes were

183 down-regulated, while 21% of the early TE genes and 29.3% of the late TE genes were up-
184 regulated by GABPA degradation, indicating that GABPA plays an essential role in the
185 pluripotency establishment (Fig. 3k and Extended Data Fig. 4h). Next, we asked what percentage
186 of the EPI genes (expressed higher in EPI than PrE, fold change≥4, q<0.05, see method) and PrE
187 genes (expressed higher in PrE than EPI, fold change≥4, q<0.05, see method) (Extended Data Fig.
188 4i and Supplementary Table1) were affected. We found 46.8% of EPI genes (e.g., pluripotency
189 factors *Nanog*, *Sox2, Fgf4* and *Prdm14*) were down-regulated (Fig. 3l, m). In contrast, much
190 smaller percentage of PrE genes (1.6%) were down-regulated, and importantly most PrE marker
191 genes, such as *Sox17* and *Sox7,* did not show significant change (Fig. 3l, m and Extended Data Fig.
192 4j), consistent with EPI but not PrE were affected upon GABPA degradation (Fig. 3d-h). The
193 increase of TE genes expression, such as *Cdx2, Krt8, Gata3* and *Krt18* might be due to the loss of
194 pluripotency as knockout of pluripotency factors have been reported to up-regulate TE genes [8,28].
195 Since EPI specification is still ongoing at middle blastocyst stage [29], yet GABPA degradation has
196 already led to ICM and EPI gene defects, which prevent the transition of E3.5 ICM to E4.5 EPI.
197 Collectively, our data support the notion that GABPA determines EPI but not PrE specification by
198 regulating EPI genes.

199 **GABPA regulates EPI gene expression by both promoter and distal binding**

200 To understand how GABPA regulates the ICM genes in E4.5 embryos, we performed GABPA
201 CUT&RUN using E4.5 ICM cells (Extended Data Fig. 5a). Similar to that in 2-cell embryos
202 (Extended Data Fig. 3c), GABPA mainly occupies the promoter regions at E4.5 ICM (Fig. 4a and
203 Extended Data Fig. 5b). Interestingly, motif analysis revealed that these regions were enriched for
204 the motif of GABPA, but not other lineage regulators such as OCT4, SOX2, NANOG, NR5A2,
205 TFAP2C or GATA3, indicating that GABPA plays a direct regulatory role by binding to these
206 regions (Extended Data Fig. 5b). This specific motif enrichment underscores GABPA's unique
207 regulatory role that distinguishes it from other factors involved in pluripotency regulation.
208 Integrative analyses of the GABPA CUT&RUN and RNA-seq data revealed that almost half of
209 the down-regulated genes (904 out of 1,886) were bound by GABPA, which is consistent with its
210 motif enrichment in these promoters (Fig. 4b). In contrast, the up-regulated genes were not directly
211 bound by GABPA (Fig. 4b). These results indicate that GABPA directly binds to promoters to
212 activate these genes in E4.5 ICM.

213 By separating E4.5 ICM genes into early and late ICM genes, we found that GABPA mainly bound
214 to the promoters of early ICM genes (e.g., *Tdgf1*, *Etv5, Dppa4* and *Uhrf1*), but not that of the late
215 ICM genes (Fig. 4c, d). In addition to promoter, GABPA also exhibited distal binding, which were
216 putative enhancer regions. We asked whether ICM gene expression is related to distal GABPA
217 binding by calculating the distance between TSS of ICM/TE genes and their nearest distal GABPA
218 peaks. We found that the TSS of the ICM genes were closer to the distal GABPA peaks than that
219 of the TE genes (Fig. 4e), indicating a potential role of these distal GABPA bindings in regulating
220 ICM genes. Indeed, we detected binding of GABPA to the known super-enhancer of ICM gene
221 *Sox2* [30] (Extended Data Fig. 5c).

222 To investigate whether GABPA would affect chromatin accessibility and/or promoter/enhancer
223 activity, we performed ATAC-seq as well as H3K27ac CUT&RUN in E4.5 ICM with or without
224 GABPA degradation (Extended Data Fig. 5d). Interestingly, we found that GABPA degradation
225 resulted in a widespread decrease of H3K27ac with little effect on chromatin accessibility (Fig. 4f,
226 g and Extended Data Fig. 5c). This result indicates that while GABPA is not responsible for
227 chromatin opening at its bound regions, it is important for the transcriptional activity of the bound
228 genes. Previous reports of the interaction between GABPA and p300 [31,32] and the fact that GABPA
229 degradation resulted in a decrease in H3K27ac level in GABP bound regions suggest that GABPA
230 may activate its target genes through recruiting the acetyltransferase p300.

231 **GABPA regulates a common sets of EPI genes in E4.5 ICM and naïve pluripotent stem cells**

232 Since E4.5 ICM is composed of EPI and PrE cells, and it is technically challenging to obtain pure
233 EPI cells to evaluate the role of GABPA in EPI gene expression. To overcome this challenge, we
234 used 2iESCs as an alternative because 2iESCs is believed to resemble the E4.5 EPI [33]. Since *Gabp*a
235 knockout leads to defects in ESCs survival [34], which is consistent with the role of GABPA in EPI
236 generation (Fig. 3d-h), we used the *Gabpa*$^{dTAG/dTAG}$ ESCs treated with dTAG13 to evaluate the
237 acute effect of GABPA degradation (Extended Data Fig. 6a). GABPA degradation with dTAG13
238 was confirmed by Western blot analysis (Fig. 1e) and immunostaining (Extended Data Fig. 6b).
239 RNA-seq analysis revealed 2,265 down-regulated genes and 1,563 up-regulated genes in response
240 to GABPA degradation for 24h (Extended Data Fig. 6c, d and Supplementary Table 5).
241 Importantly, GABPA CUT&RUN analyses revealed that 1,238 (54.7%) of the down-regulated

242 genes exhibited GABPA promoter binding (Fig. 4h and Extended Data Fig. 6e). Similar to that of
243 the E4.5 ICM, GABPA binding motif was enriched in the GABPA peaks and its removal did not
244 affect chromatin accessibility accessed by ATAC-seq (Extended Data Fig. 6f, g). In contrast, the
245 up-regulated genes did not have GABPA direct promoter binding (Fig. 4h). Comparative analysis
246 confirmed that GABPA binding in E4.5 ICM and ESCs were highly similar (Fig. 4i), suggesting
247 that GABPA may regulate the same set of genes in ESCs and E4.5 ICM. Indeed, analysis of the
248 differentially expressed genes (DEGs) in E4.5 ICM and ESCs in response to GABPA degradation
249 revealed significant overlap with 933 commonly down-regulated (41.19% in ESCs, 49.47% in
250 E4.5 ICM), and 356 commonly up-regulated genes (29.49% in ESCs, 28.50% in E4.5 ICM) (Fig.
251 4j). These results indicate that GABPA regulates a similar set of genes in E4.5 ICM and 2i ESCs.

252 We next analyzed the effect of GABPA degradation on EPI and PrE genes separately and found
253 26.3% EPI genes were down-regulated and 12.7% PrE genes were up-regulated (Extended Data
254 Fig. 6h), suggesting GABPA plays an important role in pluripotency regulation in ESCs. Further
255 analysis of the down-regulated ICM and EPI genes revealed 80 ICM genes in common (e.g., *Spic,*
256 *Utf1, Prdm14* and *Utf1*) (Fig. 4k), including 60 early ICM genes and 20 late ICM genes (Extended
257 Data Fig. 6i). The analysis also revealed 38 EPI genes in common (e.g., *Nanog, Eras* and *Pim2*)
258 (Fig. 4l). Collectively, data from E4.5 ICM and 2iESCs support the notion that GABPA plays a
259 crucial role in regulating naïve pluripotency both *in vivo* and *in vitro*.

260 **Stepwise pluripotency establishment controlled by TFAP2C, SOX2 and GABPA**

261 The first and second cell fate specification in mouse embryo occurs at E3.5 and E4.5 with
262 pluripotency gene expression restricted within ICM and EPI, respectively [35-37]. However, the
263 expression of some ICM genes starts as early as the 2-cell stage (Fig. 1b). A recent study has
264 reported the role of the transcription factors NR5A2 and TFAP2C in activating ICM genes at the
265 8-cell stage [10]. However, the expression of *Nr5a2* and *Tfap2c* are respectively silenced at E3.5
266 ICM and E4.5 EPI (Extended Data Fig. 7a), indicating other transcription factors are responsible
267 for the ICM gene activation at these stages. Our data shown above suggest that GABPA is at least
268 one of the key TFs responsible for EPI gene activation at E4.5.

269 To understand how GABPA participates in this process, we generated additional GABPA
270 CUT&RUN dataset in 8-cell embryos and E3.5 ICM (Extended Data Fig. 7b, c). Comparative

271 analysis of GABPA binding profiles at 2-cell, 8-cell, E3.5 ICM and E4.5 ICM indicate strong
272 GABPA binding at promoters has already occurred at 2-cell stage and is maintained through E4.5
273 ICM (Fig. 5a and Extended Data Fig. 7d), while weak promoter bindings at 2-cell became stronger
274 at E4.5 ICM stage (Fig. 5a). Compared to promoters, GABPA occupancy at intron and distal
275 regions were more dynamic (Fig. 5a). Increased GABPA binding to promoters of ICM genes at
276 E4.5 ICM likely contributes to GABPA's stage-specific functions (Fig. 5a and Extended Data Fig.
277 7e).

278 Next, we analyzed the relationship between GABPA and NR5A2 by comparing their binding
279 profiles at 2-cell and 8-cell stages. Interestingly, GABPA tended to bind to promoters, while
280 NR5A2 preferentially bind to putative enhancers (introns and distal regions) (Fig. 5b), suggesting
281 that they have distinct molecular mechanisms in regulating gene expression. We also compared
282 the binding profiles of GABPA and TFAP2C at 2-cell, 8-cell and E3.5 ICM stages and found
283 TFAP2C co-occupied with GABPA on a portion of ICM genes promoters at 8-cell and E3.5 ICM
284 (Extended Data Fig. 7f), which may explain why GABPA degradation at these stages does not
285 show phenotype. Given that TFAP2C is not expressed in E4.5 EPI (Extended Data Fig. 7a) [9], while
286 GABPA binding is increased from E3.5 ICM to E4.5 ICM, GABPA may take over TFAP2C's
287 function on their commonly occupied ICM gene promoters (Fig. 5c, d). To identify the commonly
288 regulated genes taken over by GABPA, we compared the down-regulated genes in GABPA-dTAG
289 E4.5 ICM with the down-regulated genes in *Tfap2c* maternal-zygotic KO (mz-KO) E3.5 ICM and
290 identified 7 out of the 13 *Tfap2c* KO down-regulated ICM genes and all the 5 *Tfap2c* KO down-
291 regulated EPI genes showed overlap (Extended Data Fig. 7g). Since *Tfap2c* mz-KO at 8-cell
292 showed more down-regulated ICM and EPI genes than that in E3.5 ICM (Extended Data Fig. 7h),
293 TFAP2C plays more important role in activating pluripotency genes at 8-cell than in E3.5, which
294 is consistent with the observation that *Tfap2c* expression is almost not expressed at E3.5 ICM
295 (Extended Data Fig. 7a) [9].

296 *Sox2* starts to express at E3.5 ICM and plays an important role in regulating pluripotency
297 (Extended Data Fig. 7a) [8]. We thus compared the binding profiles of GABPA and SOX2 in E3.5
298 ICM and E4.5 ICM. We found a global increase in GABPA binding concomitant with a global
299 decrease in SOX2 binding, especially on promoters from E3.5 ICM to E4.5 ICM (Fig. 5e).
300 Consistently, GABPA binding on ICM gene promoters was enhanced, while SOX2 binding on

301 ICM gene promoters was lost (Fig. 5f, g), indicating the regulation of ICM genes by GABPA was
302 enhanced at E4.5 stage. To explore the possibility that GABPA takes over the role of SOX2 in
303 activating the commonly regulated ICM genes at E4.5, we compared the down-regulated genes in
304 GABPA-dTAG E4.5 ICM and *Sox2*-mzKO E3.5 ICM and found 18 out of the 55 *Sox2*-KO down-
305 regulated ICM genes and 8 out of the 9 *Sox2*-KO down-regulated EPI genes showed overlap
306 (Extended Data Fig. 7i), supporting the notion that GABPA takes over SOX2's role on these ICM
307 gene promoters at E4.5 ICM. To gain further support for this notion, we compared the down-
308 regulated genes of GABPA-dTAG E4.5 ICM with that of the SOX2-dTAG 2iESCs because a
309 direct comparison of the effect of the two genes on E4.5 ICM is not possible due to the fact that
310 the *Sox2* mz-KO has already affected E3.5 ICM gene expression. Although SOX2 degradation in
311 2iESC only down-regulated a small number of ICM (27) and EPI (5) genes, all the SOX2 down-
312 regulated EPI genes were also down-regulated by GABPA and more than half of the SOX2 down-
313 regulated ICM genes were also down-regulated by GABPA (Fig. 5h). Collectively, these data
314 support our notion that TFAP2C and SOX2 are responsible for activation of pluripotency genes at
315 8-cell and E3.5 ICM, respectively, while GABPA mainly activates pluripotency genes during E3.5
316 ICM to E4.5 EPI transition. GABPA appears to take over the role of TFAP2C and SOX2 at the
317 promoter of pluripotency genes in E4.5 EPI in addition to activate its unique targets.

318

319 **Discussion**

320 Zygotic genome activation and pluripotency acquisition are two of the most important events
321 during preimplantation development. Understanding the molecular mechanisms underlying these
322 events are not only important in the field of development, but also critical for regenerative
323 medicine. Here we identified and demonstrated that a minor ZGA factor GABPA not only
324 regulates major ZGA, but also plays a critical role in pluripotency establishment by activating a
325 large group of pluripotency genes in E4.5 EPI. Despite the observation that GABPA KO down-
326 regulates pluripotency factors *Nanog* and *Oct4* in ESCs [34,38], its role as a key pluripotency regulator
327 was not recognized due to the lack of direct genome binding data. Additionally, its KO embryonic
328 lethality phenotype before blastocyst formation [21] prevented its role in pluripotency *in vivo* being
329 addressed using the conventional KO mouse model. In this study, by taking advantage of the dTAG
330 rapid protein degradation system, we were able to reveal the developmental stage-specific function

331 of GABPA during mouse pre-implantation development. Our data support that GABPA plays a
332 non-dispensable role in ZGA and as a master regulator of pluripotency establishment at E4.5 EPI.

333 Identification of TFs important for mammalian ZGA has been a hot topic in the past several years.
334 Previous studies have been mainly focused on maternal factors that regulate mouse ZGA such as
335 OBOX [25], NR5A2 [27], DUX [26], NFYA [39] and KLF17 [40], while the contributions of zygotic early
336 transcribed genes are neglected although published work have showed that minor ZGA is required
337 for major ZGA in mice [41]. Here, we provide evidence demonstrating that the minor ZGA factor
338 GABPA regulates a subset of major ZGA genes and pre-implantation development by directly
339 binding to the promoter of 83 major ZGA genes to regulate their transcription (Fig. 2g). We
340 provided genetic and molecular evidence demonstrating a non-dispensable role of GABPA in ZGA
341 and mouse pre-implantation development using our *Gabpa*$^{dTAG/dTAG}$ mice (Fig. 2). Our study
342 indicates that minor ZGA gene product can regulate major ZGA by directly binding to and
343 activating major ZGA. Further studies of the other minor ZGA genes are warranted to fully
344 understand the role of minor ZGA genes in regulating major ZGA and pre-implantation
345 development in mammals.

346 Another important finding of this study is that we identify GABPA as a key TF driving the E4.5
347 EPI specification. We showed that GABPA degradation during E3.5 to E4.5 transition affected
348 EPI formation, but has little effect on PrE (Fig. 3d-h), which could be explained by the down-
349 regulation of pluripotency factors including *Nanog* and *Sox2* (Fig. 3m) that are needed for
350 maintaining EPI, but not PrE, *in vivo* [11,42]. Different from *Nanog* and *Sox2*, which are exclusively
351 expressed in EPI, *Gabpa* is expressed in both EPI and PrE. Such expression pattern raises an
352 intriguing question about why GABPA selectively activates EPI genes, but not PrE genes (Fig. 3l,
353 m). One possibility is that binding of GABPA to the pluripotency gene promoters requires another
354 factor that only expressed in EPI. Alternatively, PrE may express a protein that can mask GABPA
355 binding to promoters. Future studies should confirm or refusal these possibilities.

356 Establishment of pluripotency is a continuous and complex process. Although the expression of
357 pluripotency genes such as *Nanog*, *Sox2* and *Oct4* are critical for maintaining pluripotency [43-46],
358 their role in pluripotency establishment has not been shown. At 8-cell and E3.5 ICM stages, TFs
359 such as NR5A2, TFAP2C and SOX2 are believed to modulate the activity of these pluripotency

genes [8,9], they are unlikely have a major role in EPI specification as they are either not expressed in EPI or they only occupy the promoters of very few pluripotency genes in EPI. Our data indicate that GABPA can fulfil such a role as it regulates 46.8% and 26.3% of EPI genes in E4.5 and 2iESC, respectively (Fig. 5i). Although the lack of a dTAG mouse model for SOX2 prevented a direct comparison of the role of GABPA and SOX2 in EPI formation, a comparison of dTAG13-mediated degradation of GABPA or SOX2 in 2iESC showed a clear difference as GABPA degradation down-regulated many more ICM genes and EPI genes than that of SOX2 (Fig. 5i). Our study, together with previous studies, supports a stepwise model for naïve pluripotency establishment (Fig. 5j). At 8-cell stage, TFAP2C binds to gene promoters and enhancers to active pluripotency genes. At E3.5, with the decrease of TFAP2C expression in ICM, its role on pluripotency gene expression decreases, while SOX2 starts to function. By E4.5, TFAP2C is no longer expressed in ICM (Extended Data Fig. 7a) [9], and the binding of SOX2 to the promoters of ICM genes also disappeared, while GABPA now occupies and activates almost half of the pluripotency gene in EPI. Our study thus not only identifies GABPA as an important TF for ZGA, but also demonstrates its crucial role as a master regulator of naïve pluripotency in E4.5 embryos.

376 **Methods**

377 **mESCs culture and establishment of *Gabpa*$^{dTAG/dTAG}$ cell line.**

378 The ES-E14 cells [47] were cultured on 0.1% gelatin coated plates with 2i/LIF condition. Cells were grown
379 in DMEM (Gibco, 11960069), supplemented with 15% fetal bovine serum (FBS) (Sigma-Aldrich, F6178),
380 2 mM GlutaMAX (Gibco, 35050061), 1 mM sodium pyruvate (Gibco, 11360), 1× MEM NEAA (Gibco,
381 11140050), 0.084 mM 2-mercaptoethanol (Gibco, 21985023), 1 mM sodium pyruvate (Gibco, 11360070),
382 100 U/ml penicillin-streptomycin (Gibco, 15140122), 1000 IU/ml LIF (Millipore, ESG1107), 0.5 μM
383 PD0325901 (Tocris, 4192) and 3 μM CHIR99021 (Tocris, 4423).

384 To establish *Gabpa*$^{dTAG/dTAG}$ cell line, *Gabpa*-HAL-FKBP^{F36V}-2xHA-HAR and px330 were transfected into
385 mESCs with Lipofectamine™ 2000 (Thermo, 11668030). 24 hrs later, cells were selected with puromycin
386 (Gibco, A1113803) for another 48 hrs. Then cells were cultured in puromycin-free medium for one week.
387 Single clones were picked for genotyping and further analysis.

388 ***Gabpa*$^{dTAG/dTAG}$ mice generation**

389 All animal experiments were performed in accordance with the protocols of the Institutional Animal Care
390 and Use Committee at Harvard Medical School. All mice were kept under specific pathogen-free conditions
391 within an environment controlled for temperature (20-22°C) and humidity (40-70%), and were subjected
392 to a 12-hour light/dark cycle. Generation of knock-in mice was as described previously with some
393 modification [23]. Briefly, 2-cell embryo (20 hpf) were injected with *Gabpa* donor DNA (30 ng/μl), Cas9
394 mRNA (100 ng/μl) and sgRNA (50 ng/μl) using a Piezo impact-driven micromanipulator (Primer Tech,
395 Ibaraki, Japan). Then 2-cell embryos were incubated in KSOM for 2 hrs before transferred into oviducts of
396 pseudo-pregnant ICR strain mothers (Charles River). F0 chimera mice was backcross with wild-type
397 C57BL/6J mice for at least two generation. Genotyping was performed using mouse tail lysed in lysis buffer
398 (50 mM Tirs-HCl, 0.5% Triton and 400 μg/ml Proteinase K) at 55°C overnight. For F0 and F1 mice
399 genotyping, the primers outside the homology arm are used. For genotyping of F2 and beyond, the inner
400 primers are used. The primers are listed in table S6.

401 Droplet Digital PCR (ddPCR) was used for detecting the copy number of *Gabpa*-dTAG knock-in allele in
402 F1 mice. Briefly, 250 ng purified DNA templates were digested by incubation with Haelll Enzyme (NEB)
403 at 37°C for 1h, and then inactivated at 80°C for 5 mins. A final 30 ng DNA was used as the templates for
404 PCR. *Fkbp* was used for knock-in detecting, *mRPP30* was used as a control. The primers are included in
405 table S6. Only the mice with single *Fkbp* copy (table S7) were used for further mating.

406 ***In vitro* fertilization and embryo culture**

407 Female mice (7-8 weeks) were superovulated through an initial injection of 7.5 IU pregnant mare serum
408 gonadotropin (PMSG, BioVendor, RP1782725000), followed 48 hours later with a 7.5 IU injection of
409 human chorionic gonadotropin (hCG, Sigma, C1063). Oocyte-cumulus complexes (OCCs) were collected
410 14 hrs post hCG injection. Sperm was harvested from the cauda epididymis of adult male mice (8-12 weeks)
411 1 h before OCCs collection. The sperm suspension was capacitated for 1 h in 200 μl HTF medium (Millipore,
412 MR-070-D). Subsequently, OCCs were exposed to spermatozoa for a 6-hour incubation. The time when
413 sperm were added to OCCs was considered as 0 hpf. Two-nuclear zygotes were cultured in the KSOM
414 medium (Millipore, MR-106-D) under a humidified atmosphere of 5% CO_2 at 37°C for further development.

415 **Western blot**

416 1.5×10^6 cells were lysed in 100 μl RIPA lysis buffer and incubated on ice for 30 mins. 90 μl supernatant
417 were mixed with 12.5 μl 5× loading buffer, and heated at 98 °C for 15 min. Samples were run on NuPAGE
418 4-12% gel (Invitrogen, NP0322BOX) and transferred onto PVDF Transfer Membrane. Primary antibodies
419 used included anti-GABPA (1:4000, Proteintech), anti-HA (1:1000, CST), and anti-β-actin (1:5000, CST).
420 Protein bands were detected with ECL kit (Thermo Fisher Scientific, 32209) and imaged by Tanon 4600SF
421 Imaging System (Tanon).

422 **Immunostaining and confocal microscope**

423 Embryos were fixed with 4% paraformaldehyde/0.5% triton for 20 min, followed by three times of washing
424 with PBS/0.01% triton, then blocked in PBS/1% BSA/0.01% triton for 1 h. Embryos were incubated
425 overnight at 4 °C with primary antibodies: anti-GABPA (1:200, Proteintech, 21542-1-AP, Lot#00018047),
426 anti-HA (1:200, CST, 2367S), anti-GATA4 (1:200, R&D Systems, MAB2606-SP), anti-NANOG (1:200,
427 Abcam, ab80892) and anti-CDX2 (1:500, R&D Systems, AF3665-SP). After three times of washing, the
428 embryos were incubated with secondary antibodies at room temperature (RT). DNA was stained with
429 10 μg/ml Hoechst 33342 (Sigma). The confocal microscope (Zeiss, LSM800) was used for fluorescence
430 detecting.

431 **dTAG13 treatment**

432 dTAG13 (Tocris, 6605) was reconstituted in DMSO to a 5 mM stock. For *Gabpa*$^{dTAG/dTAG}$ mESCs treatment,
433 dTAG13 was dilute in mESCs culture medium to 0.5 μM. For *Gabpa*$^{dTAG/dTAG}$ embryos treatment, dTAG13
434 was dilute in KSOM to 1 μM. Embryos were washed with KSOM with dTAG13 for at least three times,
435 then culture in KSOM with dTAG13 for further development.

436 **CUT&RUN, ATAC and RNA-seq library preparation and sequencing**

437 For mESCs CUT&RUN with more than 10K, cells were resuspended in 50 µl washing buffer (20 mM
438 HEPES/pH=7.5, 150 mM NaCl, 0.5 mM spermidine and 1× protease inhibitor) with activated Concanavalin
439 A Magnetic Beads (Polysciences, 86057-3) for 10 mins at RT, then samples were incubated with anti-
440 GABPA (1:40, Proteintech, 21542-1-AP, lot#00018047) overnight at 4°C. Samples were incubated with
441 2.8 ng/µl pA-MNase (home-made) for 2 hrs at 4 °C. Subsequently, samples were incubated with 200 µl pre-
442 cooled 0.5 µM $CaCl_2$ for 20 mins at 4 °C and quench by adding 23 µl 10×stop buffer (1700 mM NaCl, 20
443 mM EGTA, 100 mM EDTA, 0.02% Digitonin, 250 µg/ml glycogen and 250 µg/ml RNase A). DNA
444 Fragments were released by incubation at 37°C for 15 mins. 2.5 µl 10% SDS and 2.5 µl 20 mg/ml Protease
445 K (Thermo Fisher) was added and incubated at 55 °C for at least 1 h. DNA was extracted by phenol-
446 chloroform followed by ethanol precipitation. For low input mESCs and embryo CUT&RUN, some
447 modifications were made. Briefly, mESCs, zona-free embryos or isolated ICM were resuspended in 50 µl
448 washing buffer with activated Concanavalin A Magnetic Beads for 10 mins at RT, then samples were fixed
449 with 1% formaldehyde for 1 min. After three times of wash with washing buffer, samples were incubated
450 with anti-GABPA overnight at 4°C. Subsequent procedure is the same as described above. Sequencing
451 libraries were prepared with NEBNext Ultra II DNA library preparation kit for Illumina (New England
452 Biolabs, E7645S).

453 ATAC-seq was performed as previously described with some modifications [48]. Briefly, ESCs and isolated
454 ICM were digested with adapter-loaded Tn5 for 15 mins at 37°C, and stopped by stop buffer (100 mM
455 Tris/pH=8.0, 100 mM NaCl, 40 µg/ml Proteinase K and 0.4% SDS) and incubated overnight at 55°C. 5 µl
456 of 25% Tween-20 was added to quench SDS. Sequencing libraries were prepared with NEBNext High-
457 Fidelity 2×PCR Master Mix (NEB, M0541S).

458 For RNA-seq, fresh ESCs or embryos were collected. Reverse transcription and cDNA amplification were
459 performed with SMART-Seq™ v4 Ultra™ Low Input RNA Kit (Clontech, 634890), followed by cDNA
460 fragmentation, adaptor ligation, and amplification using Nextera® XT DNA Sample Preparation Kit kit
461 (Illumina, FC-131-1024).

462 All libraries were sequenced by NextSeq 550 system (Illumina) with paired-ended 75-bp reads.

463 **Immunosurgery**

464 ICMs were isolated as previously described [49]. Briefly, blastocysts at different stages were collected by
465 removing the zona pellucida with Acidic Tyrode's solution (Millipore). Embryos were then treated with

466 anti-mouse serum antibody (Sigma-Aldrich, 1:5 dilution in KSOM) for 30 min at 37 °C. After washed for
467 three times with KSOM, embryos were treated with guinea pig complement (Millipore, 1:5 dilution in
468 KSOM) for another 20 min at 37 °C. Then, the trophectoderm cells were removed by a glass pipette (the
469 inner diameter is around 40-50 μm).

470 **RNA-seq data analysis**

471 The raw sequencing reads were trimmed with Trimmomatic [50] (v0.39) to remove sequencing adaptors. Then,
472 the reads were mapped to GRCm38 genome using STAR [51] (v2.7.8a). Gene expression levels were
473 quantified with RSEM [52] (v1.3.1). To identify differentially expression genes, DESeq2 [53] (v1.32.0) package
474 in R was used. The significantly differential expressed genes were called with an adjusted P-value cutoff
475 of 0.05, fold change cutoff of 2, and mean FPKM cutoff of 1. GO enrichment was performed using R
476 package clusterProfiler [54]. GSEA analysis was performed using R clusterProfiler [54] and enrichplot.

477 **CUT&RUN data analysis**

478 The raw reads were trimmed with Trimmomatic [50] (v0.39) to remove sequencing adaptors then mapped to
479 GRCm38 reference genome using bowtie2 [55] (v2.4.2). PCR duplicates were removed with Picard
480 MarkDuplicates (v2.23.4). Reads with mapping quality less than 30 were removed. The mapped reads were
481 further filtered to only retain proper paired reads with fragment length between 10 and 120. Peaks were
482 called with MACS2 [56] (v2.2.7.1). Reproducible peaks were generated with the IDR framework [57] using two
483 replicates, with IDR threshold of 0.05. For E4.5 ICM and ESCs, we further filtered the peaks to keep the
484 ones with q-value $\leq 10^{-30}$ to remove weak peaks. The signal tracks were generated with deeptools [58]
485 bamCoverage (v3.5.1) with bin size of 1 and normalized by CPM. For z-score normalized signal tracks, we
486 first used bamCoverage with bin size of 100 to generate FPKM signals then used a customized script to
487 calculate the z-score of each bin. For late 2-cell and 8-cell GABPA ultra-low-input CUT&RUN data, we
488 noticed low mapping rates of the raw data. Further examination of the unmapped reads suggested they were
489 environmental DNA from human, bacteria, vectors etc. and due to the ultra-low-input cells and small
490 number of GABPA binding regions, the ratio of mapped reads from GABPA bound DNA versus unmapped
491 reads arising from environmental DNA were low. But this would not affect the identification of GABPA
492 peaks, since 1) the discarded reads were unmappable to mouse reference genome, and 2) upon GABPA
493 degradation these peaks disappeared.

494 The peaks were annotated with R package ChIPseeker [59]. Peaks within -1000 to +500 around TSS were
495 considered as promoter peaks.

496 The heatmaps of binding profiles were calculated with deeptools [58] computeMatrix (v3.5.1) using bigwig
497 signal tracks as input and bin size of 10, and visualized in R with packages profileplyr and EnrichedHeatmap
498 [60].

499 **ATAC-seq data analysis**

500 ATAC-seq data were analyzed with the ENCODE [61] ATAC-seq pipeline with default parameters (v2.1.2,
501 https://github.com/ENCODE-DCC/atac-seq-pipeline).

502 **Motif enrichment analysis**

503 Motif enrichment analysis was performed with HOMER [62] (v4.11) findMotifsGenome.pl with mm10
504 reference and parameter -size 200, using peaks file as input.

505 **Motif occurrence analysis**

506 Motif occurrence analysis was performed with HOMER [62] (v4.11) annotatePeaks.pl with parameters:
507 "mm10 -size -2000,2000 -hist 20 -ghist" for peaks regions, and parameters: "mm10 -size -500,500 -hist 20
508 -ghist" for regions around genes TSS. The motif files were download from JASPAR database [63] and
509 manually converted to HOMER motif format. We used the log odds detection threshold of 6.0 for all TFs
510 we analyzed. The motif occurrence matrix were visualized in R with package EnrichedHeatmap [60].

511 The JASPAR motif IDs for the TFs we analyzed were: GABPA - MA0062.2; OBOX - PH0121.1; DUX -
512 MA0611.1; NFYA - MA0060.1; NR5A2 - MA0505.1; OCT4 - MA1115.1; SOX2 - MA0143.1; NANOG
513 - MA2339.1; TFAP2C - MA0524.2; GATA3 - MA0037.4.

514 **Identification of ICM/TE genes**

515 ICM/TE genes were identified using bulk RNA-seq data of E3.5 ICM and E3.5 TE from Zhang et al. [18].
516 DESeq2 [53] (v1.32.0) package in R was used to find the differential expressed genes between ICM and TE,
517 with an adjusted P-value cutoff of 0.05 and fold change cutoff of 2. Genes with 8-cell FPKM $\geq$ 1 were
518 defined as early ICM/TE genes, while genes with 8-cell FPKM < 1 were defined as late ICM/TE genes.

519 **Identification of EPI/PrE genes**

520 EPI/PrE genes were identified using single-cell RNA-seq data of E4.5 embryos from Mohammed et al. [64].
521 R package Seurat [65] (v5.0.1) function FindMarkers was used to identify the differential expressed genes
522 between E4.5 EPI and E4.5 PrE, with an adjusted P-value cutoff of 0.05, fold change cutoff of 4 and cell

523 expression percentage cutoff of 0.8. Genes with expression in at least 50% E3.5 single cells were considered
524 as early EPI/PrE genes, while the others were considered as late EPI/PrE genes.

525 **Public data used in this study**

526 RNA-seq of mouse MII oocyte to 8-cell [17]: GSE71434. RNA-seq of mouse E3.5 ICM and TE [18]: GSE76505.
527 RNA-seq of mouse E4.5 TE [9]: GSE216256. scRNA-seq of mouse E4.5 EPI and PrE [66]: GSE159030.
528 scRNA-seq of mouse early embryos [19]: GSE45719. scRNA-seq of mouse E3.5 and E4.5 embryos [64]:
529 GSE100597. ATAC-seq of mouse early embryos [16]: GSE66390. NR5A2 binding in mouse embryos [10]:
530 GSE229740. TFAP2C binding in mouse embryos [9]: GSE216256. SOX2 binding in mouse embryos [8]:
531 GSE203194.

532

533 **References**

534 1 Schulz, K. N. & Harrison, M. M. Mechanisms regulating zygotic genome activation.
535 *Nature reviews. Genetics* **20**, 221-234, doi:10.1038/s41576-018-0087-x (2019).

536 2 Fu, X., Zhang, C. & Zhang, Y. Epigenetic regulation of mouse preimplantation embryo
537 development. *Current opinion in genetics & development* **64**, 13-20,
538 doi:10.1016/j.gde.2020.05.015 (2020).

539 3 Zernicka-Goetz, M., Morris, S. A. & Bruce, A. W. Making a firm decision: multifaceted
540 regulation of cell fate in the early mouse embryo. *Nat Rev Genet* **10**, 467-477,
541 doi:10.1038/nrg2564 (2009).

542 4 Chen, Q., Shi, J., Tao, Y. & Zernicka-Goetz, M. Tracing the origin of heterogeneity and
543 symmetry breaking in the early mammalian embryo. *Nat Commun* **9**, 1819,
544 doi:10.1038/s41467-018-04155-2 (2018).

545 5 Rossant, J. Genetic Control of Early Cell Lineages in the Mammalian Embryo. *Annu Rev*
546 *Genet* **52**, 185-201, doi:10.1146/annurev-genet-120116-024544 (2018).

547 6 Pera, M. F. & Rossant, J. The exploration of pluripotency space: Charting cell state
548 transitions in peri-implantation development. *Cell Stem Cell* **28**, 1896-1906,
549 doi:10.1016/j.stem.2021.10.001 (2021).

550 7 Boroviak, T., Loos, R., Bertone, P., Smith, A. & Nichols, J. The ability of inner-cell-mass
551 cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat*
552 *Cell Biol* **16**, 516-528, doi:10.1038/ncb2965 (2014).

553 8 Li, L. *et al.* Multifaceted SOX2-chromatin interaction underpins pluripotency progression
554 in early embryos. *Science* **382**, eadi5516, doi:10.1126/science.adi5516 (2023).

555 9 Li, L. *et al.* Lineage regulators TFAP2C and NR5A2 function as bipotency activators in
556 totipotent embryos. *Nat Struct Mol Biol*, doi:10.1038/s41594-023-01199-x (2024).

557 10 Lai, F. *et al.* NR5A2 connects zygotic genome activation to the first lineage segregation in
558 totipotent embryos. *Cell Res* **33**, 952-966, doi:10.1038/s41422-023-00887-z (2023).

559 11 Wicklow, E. *et al.* HIPPO pathway members restrict SOX2 to the inner cell mass where it
560 promotes ICM fates in the mouse blastocyst. *PLoS Genet* **10**, e1004618,
561 doi:10.1371/journal.pgen.1004618 (2014).

562 12 Silva, J. *et al.* Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722-737,
563 doi:10.1016/j.cell.2009.07.039 (2009).

564 13 Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T. & Kanemaki, M. An auxin-
565 based degron system for the rapid depletion of proteins in nonplant cells. *Nat Methods* **6**,
566 917-922, doi:10.1038/nmeth.1401 (2009).

567 14 Nabet, B. *et al.* The dTAG system for immediate and target-specific protein degradation.
568 *Nat Chem Biol* **14**, 431-441, doi:10.1038/s41589-018-0021-8 (2018).

569 15 Abuhashem, A., Lee, A. S., Joyner, A. L. & Hadjantonakis, A. K. Rapid and efficient
570 degradation of endogenous proteins in vivo identifies stage-specific roles of RNA Pol II
571 pausing in mammalian development. *Dev Cell* **57**, 1068-1080 e1066,
572 doi:10.1016/j.devcel.2022.03.013 (2022).

573 16 Wu, J. *et al.* The landscape of accessible chromatin in mammalian preimplantation
574 embryos. *Nature* **534**, 652-657, doi:10.1038/nature18606 (2016).

575 17 Zhang, B. *et al.* Allelic reprogramming of the histone modification H3K4me3 in early
576 mammalian development. *Nature* **537**, 553-557, doi:10.1038/nature19361 (2016).

577 18 Zhang, Y. *et al.* Dynamic epigenomic landscapes during early lineage specification in
578 mouse embryos. *Nature genetics* **50**, 96-105, doi:10.1038/s41588-017-0003-x (2018).

579 19 Deng, Q., Ramskold, D., Reinius, B. & Sandberg, R. Single-cell RNA-seq reveals dynamic,
580 random monoallelic gene expression in mammalian cells. *Science* **343**, 193-196,
581 doi:10.1126/science.1245316 (2014).

582 20 Rosmarin, A. G., Resendes, K. K., Yang, Z., McMillan, J. N. & Fleming, S. L. GA-binding
583 protein transcription factor: a review of GABP as an integrator of intracellular signaling
584 and protein-protein interactions. *Blood Cells Mol Dis* **32**, 143-154,
585 doi:10.1016/j.bcmd.2003.09.005 (2004).

586 21 Ristevski, S. *et al.* The ETS transcription factor GABPalpha is essential for early
587 embryogenesis. *Mol Cell Biol* **24**, 5844-5849, doi:10.1128/MCB.24.13.5844-5849.2004
588 (2004).

589 22 Liu, N. Q. *et al.* WAPL maintains a cohesin loading cycle to preserve cell-type-specific
590 distal gene regulation. *Nat Genet* **53**, 100-109, doi:10.1038/s41588-020-00744-4 (2021).

591 23 Gu, B., Posfai, E. & Rossant, J. Efficient generation of targeted large insertions by
592 microinjection into two-cell-stage mouse embryos. *Nat Biotechnol* **36**, 632-637,
593 doi:10.1038/nbt.4166 (2018).

594 24 Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution
595 mapping of DNA binding sites. *Elife* **6**, doi:10.7554/eLife.21856 (2017).

596 25 Ji, S. *et al.* OBOX regulates mouse zygotic genome activation and early development.
597 *Nature* **620**, 1047-1053, doi:10.1038/s41586-023-06428-3 (2023).

598 26 Chen, Z. & Zhang, Y. Loss of DUX causes minor defects in zygotic genome activation and
599 is compatible with mouse development. *Nat Genet* **51**, 947-951, doi:10.1038/s41588-019-
600 0418-7 (2019).

601 27 Gassler, J. *et al.* Zygotic genome activation by the totipotency pioneer factor Nr5a2.
602 *Science* **378**, 1305-1315, doi:10.1126/science.abn7478 (2022).

603 28 Jia, J. *et al.* Regulation of pluripotency and self- renewal of ESCs through epigenetic-
604 threshold modulation and mRNA pruning. *Cell* **151**, 576-589,
605 doi:10.1016/j.cell.2012.09.023 (2012).

606 29 Nakai-Futatsugi, Y. & Niwa, H. Epiblast and primitive endoderm differentiation: fragile
607 specification ensures stable commitment. *Cell Stem Cell* **16**, 346-347,
608 doi:10.1016/j.stem.2015.03.005 (2015).

609 30 Zhou, H. Y. *et al.* A Sox2 distal enhancer cluster regulates embryonic stem cell
610 differentiation potential. *Genes Dev* **28**, 2699-2711, doi:10.1101/gad.248526.114 (2014).

611 31 Bannert, N., Avots, A., Baier, M., Serfling, E. & Kurth, R. GA-binding protein factors, in
612 concert with the coactivator CREB binding protein/p300, control the induction of the
613 interleukin 16 promoter in T lymphocytes. *Proc Natl Acad Sci U S A* **96**, 1541-1546,
614 doi:10.1073/pnas.96.4.1541 (1999).

615 32 Bush, T. S., St Coeur, M., Resendes, K. K. & Rosmarin, A. G. GA-binding protein (GABP)
616 and Sp1 are required, along with retinoid receptors, to mediate retinoic acid responsiveness
617 of CD18 (beta 2 leukocyte integrin): a novel mechanism of transcriptional regulation in
618 myeloid cells. *Blood* **101**, 311-317, doi:10.1182/blood.V101.1.311 (2003).

619 33 Boroviak, T. *et al.* Lineage-Specific Profiling Delineates the Emergence and Progression
620 of Naive Pluripotency in Mammalian Embryogenesis. *Dev Cell* **35**, 366-382,
621 doi:10.1016/j.devcel.2015.10.011 (2015).

622 34 Ueda, A., Akagi, T. & Yokota, T. GA-Binding Protein Alpha Is Involved in the Survival
623 of Mouse Embryonic Stem Cells. *Stem Cells* **35**, 2229-2238, doi:10.1002/stem.2673 (2017).

624 35 Lim, H. Y. G. & Plachta, N. Cytoskeletal control of early mammalian development. *Nat*
625 *Rev Mol Cell Biol* **22**, 548-562, doi:10.1038/s41580-021-00363-9 (2021).

626 36 Zhu, M. & Zernicka-Goetz, M. Principles of Self-Organization of the Mammalian Embryo.
627 *Cell* **183**, 1467-1478, doi:10.1016/j.cell.2020.11.003 (2020).

628 37 Zhang, H. T. & Hiiragi, T. Symmetry Breaking in the Mammalian Embryo. *Annu Rev Cell*
629 *Dev Biol* **34**, 405-426, doi:10.1146/annurev-cellbio-100617-062616 (2018).

630 38 Kinoshita, K. *et al.* GABPalpha regulates Oct-3/4 expression in mouse embryonic stem
631 cells. *Biochem Biophys Res Commun* **353**, 686-691, doi:10.1016/j.bbrc.2006.12.071 (2007).

632 39 Lu, F. *et al.* Establishing Chromatin Regulatory Landscape during Mouse Preimplantation
633 Development. *Cell* **165**, 1375-1388, doi:10.1016/j.cell.2016.05.050 (2016).

634 40 Hu, Y. *et al.* Maternal KLF17 controls zygotic genome activation by acting as a messenger
635 for RNA Pol II recruitment in mouse embryos. *Dev Cell* **59**, 613-626 e616,
636 doi:10.1016/j.devcel.2024.01.013 (2024).

637 41 Abe, K. I. *et al.* Minor zygotic gene activation is essential for mouse preimplantation
638 development. *Proceedings of the National Academy of Sciences of the United States of*
639 *America* **115**, E6780-E6788, doi:10.1073/pnas.1804309115 (2018).

640 42 Mitsui, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in
641 mouse epiblast and ES cells. *Cell* **113**, 631-642, doi:10.1016/s0092-8674(03)00393-3
642 (2003).

643 43 Ding, J., Xu, H., Faiola, F., Ma'ayan, A. & Wang, J. Oct4 links multiple epigenetic
644 pathways to the pluripotency network. *Cell Res* **22**, 155-167, doi:10.1038/cr.2011.179
645 (2012).

646 44 Murakami, K. *et al.* NANOG alone induces germ cells in primed epiblast in vitro by
647 activation of enhancers. *Nature* **529**, 403-407, doi:10.1038/nature16480 (2016).

648 45 Surani, M. A., Hayashi, K. & Hajkova, P. Genetic and epigenetic regulators of pluripotency.
649 *Cell* **128**, 747-762, doi:10.1016/j.cell.2007.02.010 (2007).

650 46 Masui, S. *et al.* Pluripotency governed by Sox2 via regulation of Oct3/4 expression in
651 mouse embryonic stem cells. *Nat Cell Biol* **9**, 625-635, doi:10.1038/ncb1589 (2007).

652 47 Fu, X., Djekidel, M. N. & Zhang, Y. A transcriptional roadmap for 2C-like-to-pluripotent
653 state transition. *Sci Adv* **6**, eaay5181, doi:10.1126/sciadv.aay5181 (2020).

654 48 Zhang, C. *et al.* The chromatin remodeler Snf2h is essential for oocyte meiotic cell cycle
655 progression. *Genes Dev* **34**, 166-178, doi:10.1101/gad.331157.119 (2020).

656 49 Chen, Z., Djekidel, M. N. & Zhang, Y. Distinct dynamics and functions of H2AK119ub1
657 and H3K27me3 in mouse preimplantation embryos. *Nat Genet* **53**, 551-563,
658 doi:10.1038/s41588-021-00821-2 (2021).

659 50 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina
660 sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).

661 51 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21,
662 doi:10.1093/bioinformatics/bts635 (2013).

663 52 Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with
664 or without a reference genome. *BMC bioinformatics* **12**, 323, doi:10.1186/1471-2105-12-
665 323 (2011).

666 53 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion
667 for RNA-seq data with DESeq2. *Genome biology* **15**, 550, doi:10.1186/s13059-014-0550-
668 8 (2014).

669 54 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing
670 biological themes among gene clusters. *Omics : a journal of integrative biology* **16**, 284-
671 287, doi:10.1089/omi.2011.0118 (2012).

672 55 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature*
673 *methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).

674 56 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, R137,
675 doi:10.1186/gb-2008-9-9-r137 (2008).

676 57 Landt, S. G. *et al.* ChIP-seq guidelines and practices of the ENCODE and modENCODE
677 consortia. *Genome research* **22**, 1813-1831, doi:10.1101/gr.136184.111 (2012).

678 58 Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data
679 analysis. *Nucleic acids research* **44**, W160-165, doi:10.1093/nar/gkw257 (2016).

680 59 Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak
681 annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383,
682 doi:10.1093/bioinformatics/btv145 (2015).

683 60 Gu, Z., Eils, R., Schlesner, M. & Ishaque, N. EnrichedHeatmap: an R/Bioconductor
684 package for comprehensive visualization of genomic signal associations. *BMC genomics*
685 **19**, 234, doi:10.1186/s12864-018-4625-x (2018).

686 61 Encode Project Consortium *et al.* Expanded encyclopaedias of DNA elements in the human
687 and mouse genomes. *Nature* **583**, 699-710, doi:10.1038/s41586-020-2493-4 (2020).

688 62 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime
689 cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**,
690 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).

691 63 Castro-Mondragon, J. A. *et al.* JASPAR 2022: the 9th release of the open-access database
692 of transcription factor binding profiles. *Nucleic acids research* **50**, D165-D173,
693 doi:10.1093/nar/gkab1113 (2022).

694 64 Mohammed, H. *et al.* Single-Cell Landscape of Transcriptional Heterogeneity and Cell
695 Fate Decisions during Mouse Early Gastrulation. *Cell reports* **20**, 1215-1228,
696 doi:10.1016/j.celrep.2017.07.009 (2017).

697 65 Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell
698 analysis. *Nature biotechnology* **42**, 293-304, doi:10.1038/s41587-023-01767-y (2024).

699 66 Stirparo, G. G. *et al.* OCT4 induces embryonic pluripotency via STAT3 signaling and
700 metabolic mechanisms. *Proceedings of the National Academy of Sciences of the United*
701 *States of America* **118**, doi:10.1073/pnas.2008890118 (2021).

702

Acknowledgements

We thank Dr. Yota Hagihara and Dr. Qianying Yang for their valuable discussions on the project and comments on the manuscript. This project was supported by the NIH (R01HD092465) and the HHMI. YZ is an investigator of the Howard Hughes Medical Institute. This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

Author contributions

Y.Z. conceived and supervised the project. C-J.Z. performed the experiments including RNA-seq, CUT&RUN, ATAC-seq etc. M.W. performed all the bioinformatic analysis. C-X.Z. established the GABPA-dTAG mice and GABPA-dTAG ESCs and performed WB verification. C-J.Z., M.W. and Y.Z. wrote the manuscript. All authors interpreted the data and reviewed the manuscript.

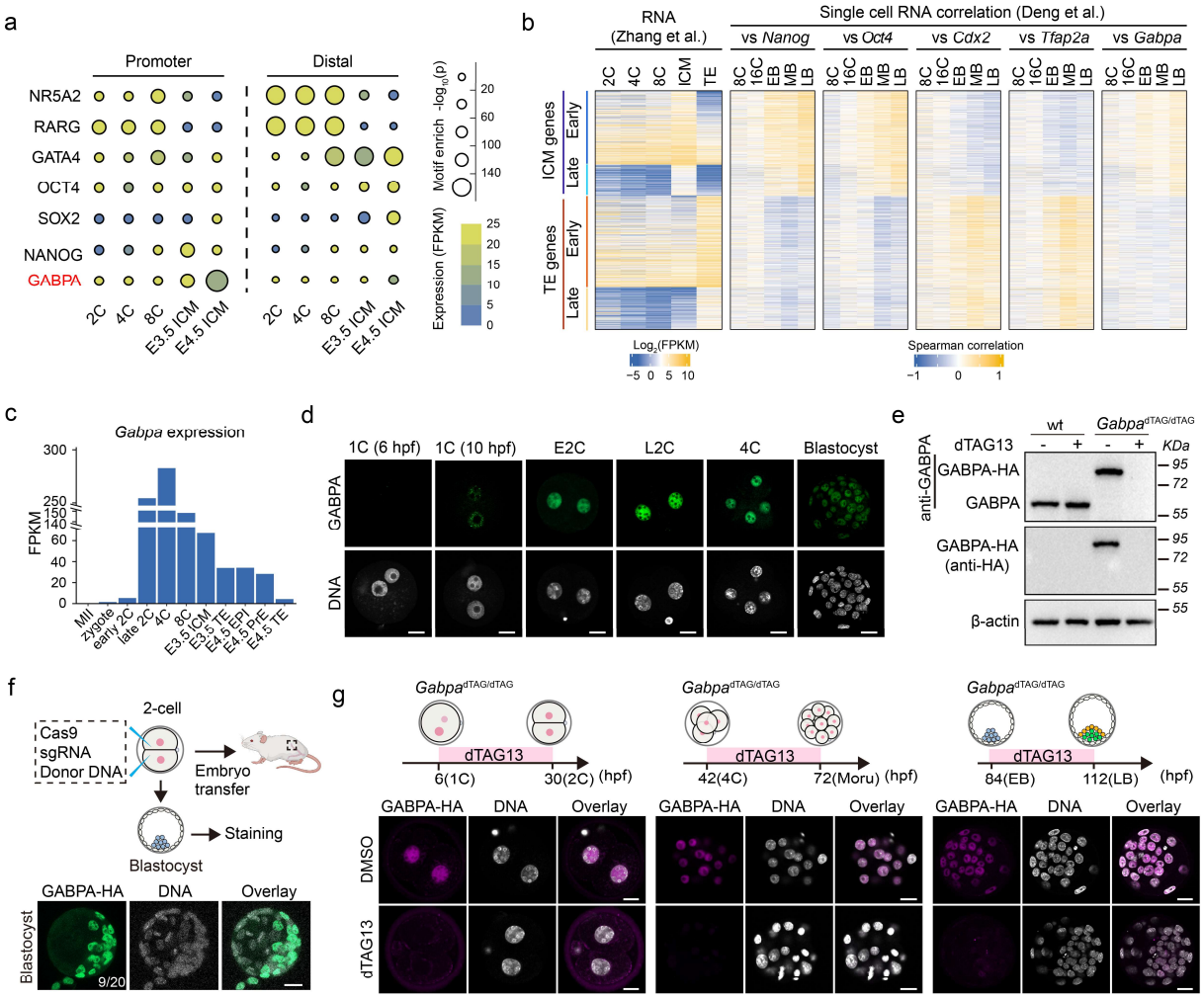
Competing interests

The authors declare no competing interests.

Data and materials availability

All data generated in this study have been deposited to the NCBI Gene Expression Omnibus (GEO) with accession number GSE263171. (Token for reviewers' access: ojineoicnnaftkz).

722 **Fig.1**



723

724 **Fig. 1. Identification of *Gabpa* as a potential pluripotency regulator and generation the**
725 ***Gabpa*$^{dTAG/dTAG}$ mice.** (**a**) Enrichment of transcription factors motifs at promoter (left) and distal
726 (right) ATAC-seq peaks during mouse preimplantation development. (**b**) Expression correlation
727 of ICM genes and TE genes with the expression of *Gabpa*, and other known ICM markers (*Nanog*
728 and *Oct4*), TE markers (*Cdx2* and *Tfap2a*) at each stage of 8-cell (8C), 16-cell (16C), early
729 blastocyst (EB), middle blastocyst (MB) and late blastocyst (LB) at single cell level. Early ICM/TE
730 genes are genes with 8-cell expression (FPKM≥1), late ICM/TE genes are genes starting
731 expression at E3.5 with no 8-cell expression (FPKM<1). (**c**) *Gabpa* expression levels during mouse
732 preimplantation development. (**d**) Immunostaining of GABPA (green) during mouse
733 preimplantation development. DNA, Hoechst 33342. Scale bar, 20 μm. (**e**) Western blot

734 confirming GABPA degradation with dTAG13 in ESCs. The blots were incubated with anti-
735 GABPA and anti-HA, respectively. β-actin was used as a loading control. (**f**) Strategy used for
736 *Gabpa*$^{dTAG/dTAG}$ mouse generation. HA staining indicates the dTAG knock-in cells in blastocyst.
737 DNA, Hoechst 33342. Scale bar, 20 μm. (**g**) Immunostaining confirms GABPA degradation with
738 dTAG13 treatment at different preimplantation developmental stages. anti-HA antibody was used
739 for endogenous GABPA-HA fusion protein staining. DNA, Hoechst 33342. Scale bar, 20 μm.

740

Fig.2

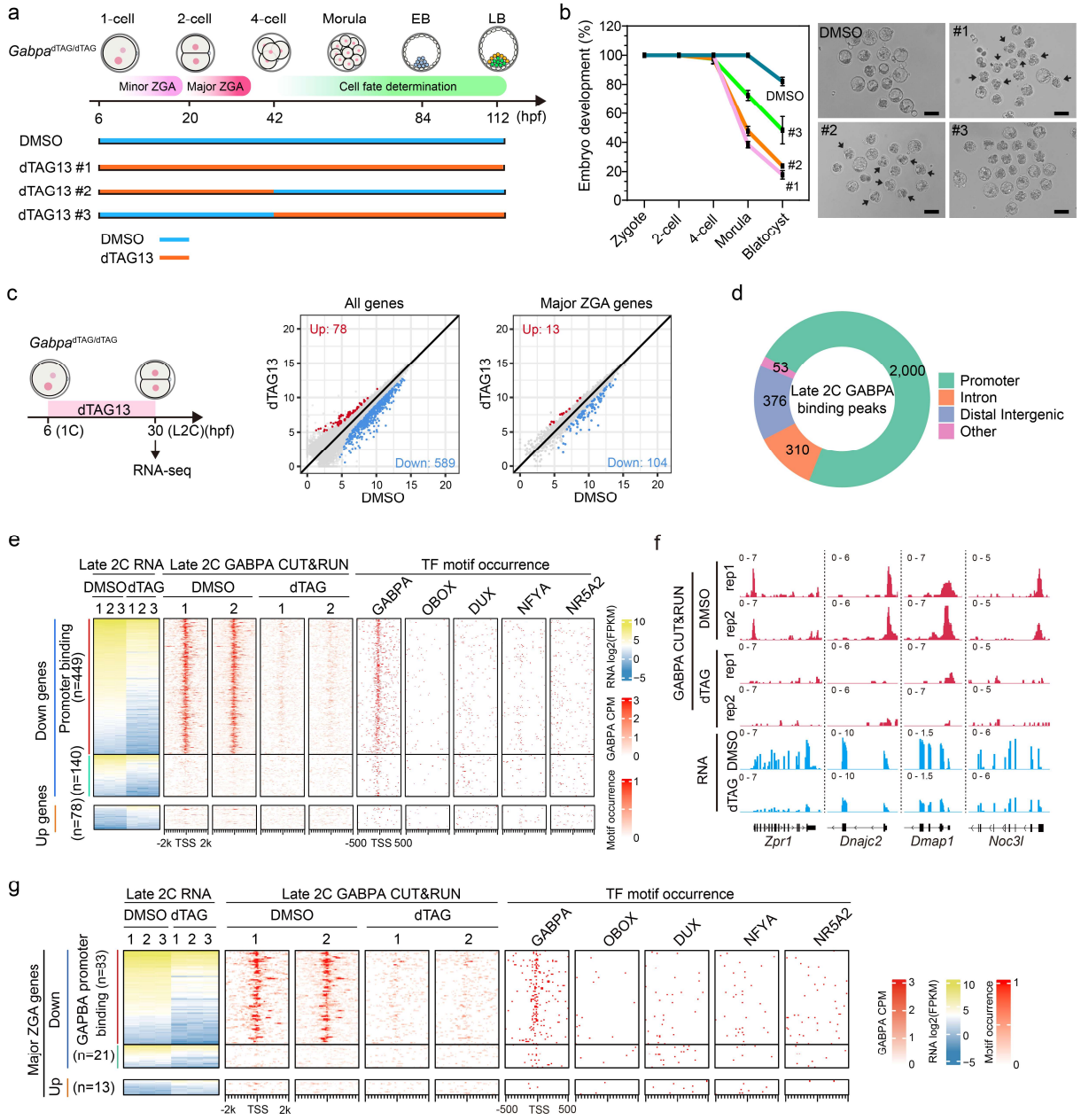
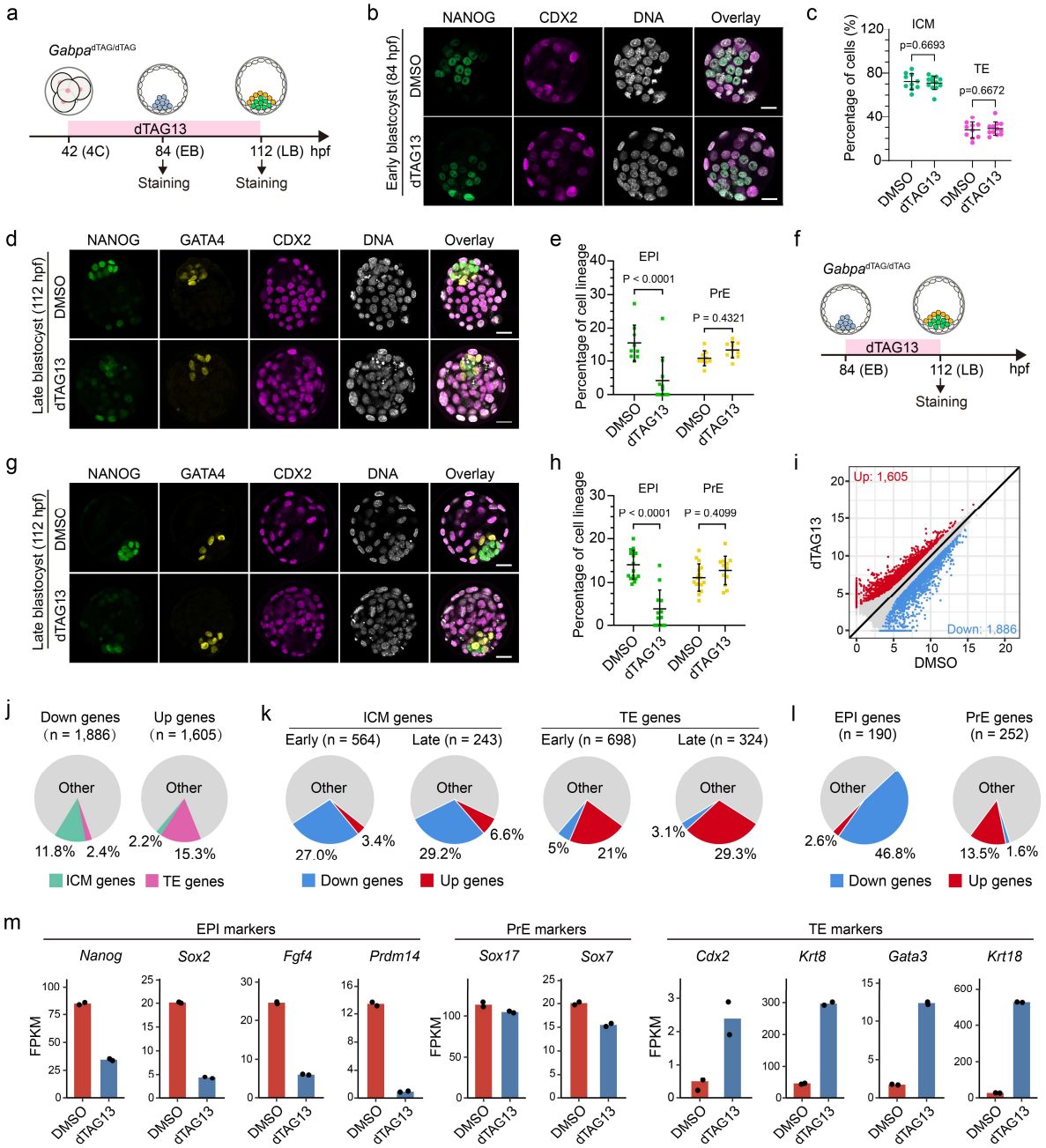


Fig. 2. GABPA activates a group of major ZGA genes by binding to their promoters. (**a**) Diagram showing dTAG13 triggered GABPA degradation at different time window of mouse pre-implantation development. dTAG13 #1: dTAG13 treatment from zygote (6 hpf) to late blastocyst (112 hpf) stage; dTAG13 #2: dTAG13 treatment from zygote to 4-cell (42 hpf) stage (covers ZGA stage). dTAG13 #3: dTAG13 treatment from 4-cell to late blastocyst (after ZGA). (**b**) Embryo

748 development rate (left) and representative images of blastocyst stage embryos (right) after GABPA
749 degradation. Black arrows indicate the 4-cell or 8-cell arrested embryos. Scale bar, 100 μm. (**c**)
750 RNA-seq comparison of late 2-cell embryos treated with dTAG13 or DMSO from zygote (6 hpf)
751 to late 2-cell (30 hpf). The x and y axis of the dot plots are log2 normalized counts from RNA-
752 seq. (**d**) Genomic distribution of GABPA binding peaks generated by GABPA CUT&RUN at late
753 2-cell stage. (**e**) Heatmaps showing the differentially expressed genes (DEGs) at late 2-cell upon
754 GABPA degradation, and the GABPA binding, TFs motif occurrence around the TSS of
755 corresponding genes. (**f**) Genome browser examples of GABPA target genes at late 2-cell stage.
756 (**g**) Heatmaps showing the differentially expressed major ZGA genes and GABPA binding profile
757 in late 2-cell embryos after GABPA degradation. TFs motif occurrences around the TSS of
758 corresponding ZGA genes are shown.

759

760 **Fig.3**



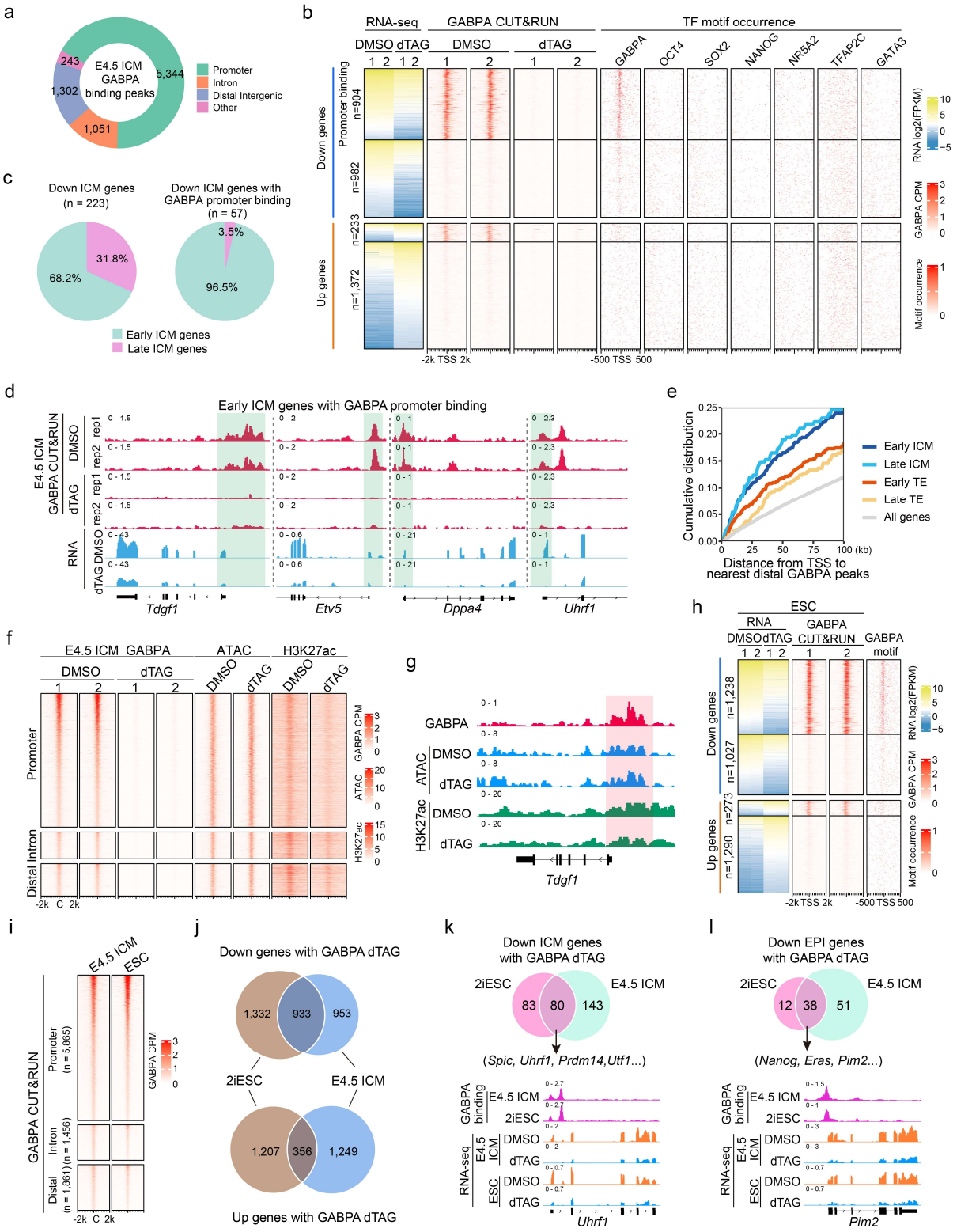
761

762 **Fig. 3. GABPA controls EPI specification by activating pluripotency genes.** (**a**) Diagram
763 showing the dTAG13 treatment after major ZGA and sample collection at early (84 hpf) and late
764 (112 hpf) blastocyst. (**b**) Immunostaining of NANOG (green) and CDX2 (red) at early blastocyst
765 with or without dTAG13 treatment. Scale bar, 20 μm. (**c**) Percentage of ICM and TE cells
766 quantified based on panel b. ICM: $NANOG^{+}/CDX2^{-}$. TE: $CDX2^{+}$. p values were calculated with

767 student's t-test. (**d**) Immunostaining of NANOG (green), GATA4 (yellow) and CDX2 (red) at
768 middle blastocyst stage with or without dTAG13 treatment. Scale bar, 20 μm. (**e**) Percentage of
769 EPI and PrE cells quantified based on panel d. EPI: $NANOG^{+}/GATA4^{-}/CDX2^{-}$. PrE: $GATA4^{+}/$
770 $NANOG^{-}/CDX2^{-}$. p values were calculated with Student's t-test. (**f**) Diagram showing dTAG13
771 treatment from early to late blastocyst. (**g**) Same as panel d except the dTAG13 treatment time is
772 as indicated in panel f. Scale bar, 20 μm. (**h**) Percentage of EPI and PrE cells quantified based on
773 panel g. EPI: $NANOG^{+}/GATA4^{-}/CDX2^{-}$. PrE: $GATA4^{+}/$ $NANOG^{-}/CDX2^{-}$. p values were
774 calculated with Student's t-test. (**i**) DEGs from RNA-seq of E4.5 (late blastocyst) ICM with or
775 without dTAG13 treatment. (**j**) Percentage of differential expressed ICM and TE genes. (**k**)
776 Percentage of affected early/late ICM and TE genes after dTAG13 treatment. (**l**) Percentage of up-
777 and down-regulated EPI and PrE genes after dTAG13 treatment. (**m**) Expression levels of EPI,
778 PrE and TE marker genes with or without dTAG13 treatment.

779

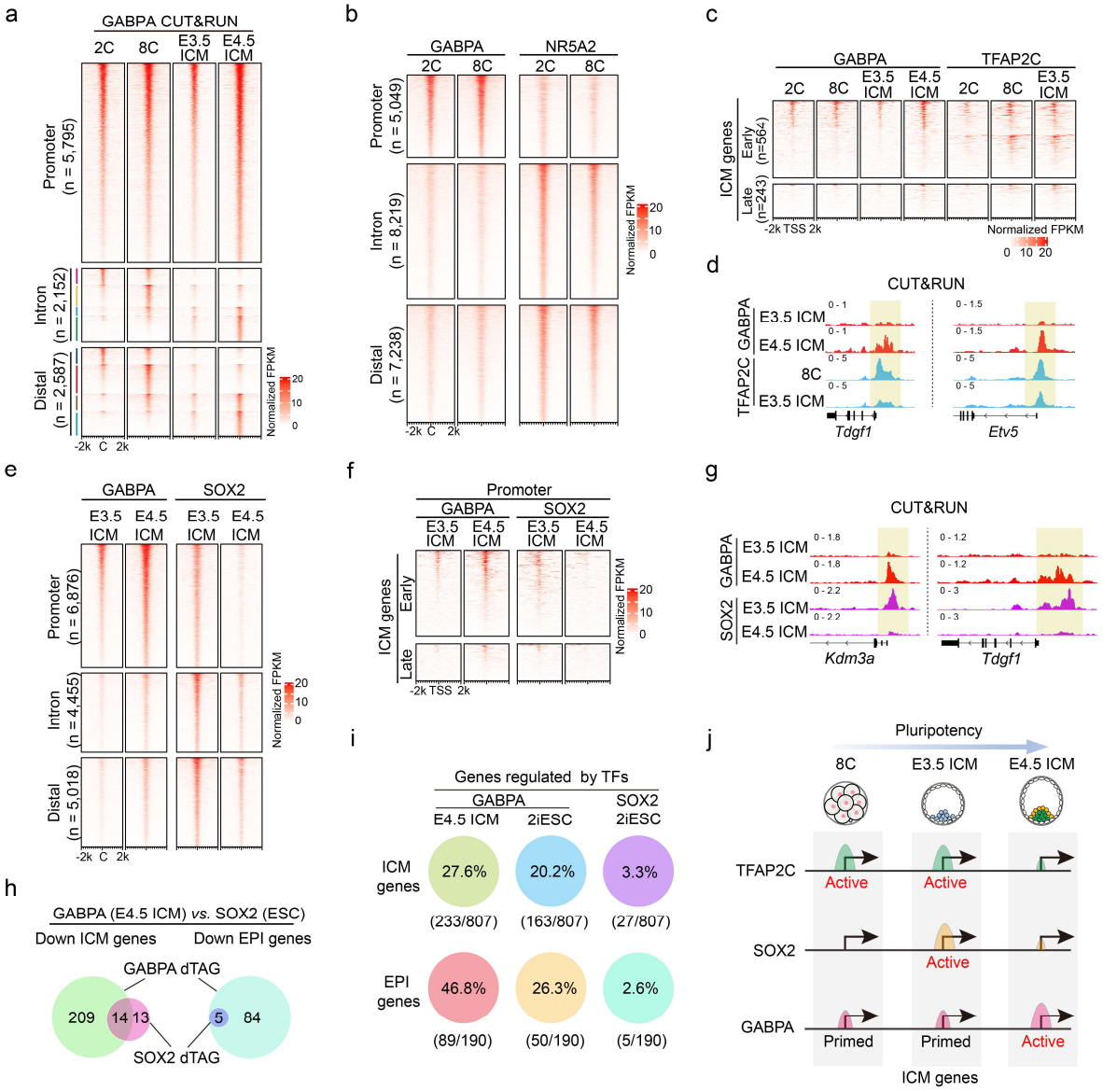
Fig.4


a
E4.5 ICM GABPA binding peaks
5,344
1,051
1,302
243
Promoter
Intron
Distal Intergenic
Other
b
RNA-seq
GABPA CUT&RUN
TF motif occurrence
DMSO dTAG
DMSO
dTAG
GABPA
OCT4
SOX2
NANOG
NR5A2
TFAP2C
GATA3
Down genes
Promoter binding
n=904
n=982
Up genes
n=233
n=1,372
-2k TSS 2k
-500 TSS 500
RNA log2(FPKM)
GABPA CPM
Motif occurrence
c
Down ICM genes (n = 223)
Down ICM genes with GABPA promoter binding (n = 57)
31.8%
68.2%
3.5%
96.5%
Early ICM genes
Late ICM genes
d
Early ICM genes with GABPA promoter binding
E4.5 ICM GABPA CUT&RUN
DMSO
dTAG
rep1
rep2
RNA
Tdgf1
Etv5
Dppa4
Uhrf1
e
Cumulative distribution
Early ICM
Late ICM
Early TE
Late TE
All genes
Distance from TSS to nearest distal GABPA peaks
f
E4.5 ICM GABPA
ATAC
H3K27ac
Promoter
Distal Intron
GABPA CPM
ATAC
H3K27ac
g
GABPA
ATAC
H3K27ac
Tdgf1
h
ESC
RNA
GABPA CUT&RUN
GABPA motif
Down genes
n=1,238
n=1,027
Up genes
n=273
n=1,290
i
E4.5 ICM
ESC
GABPA CUT&RUN
Promoter (n = 5,865)
Intron (n = 1,456)
Distal (n = 1,861)
GABPA CPM
j
Down genes with GABPA dTAG
1,332
933
953
2iESC
E4.5 ICM
1,207
356
1,249
Up genes with GABPA dTAG
k
Down ICM genes with GABPA dTAG
2iESC
E4.5 ICM
83
80
143
(Spic, Uhrf1, Prdm14,Utf1...)
GABPA binding
RNA-seq
Uhrf1
l
Down EPI genes with GABPA dTAG
2iESC
E4.5 ICM
12
38
51
(Nanog, Eras, Pim2...)
Pim2

782 **Fig. 4. GABPA regulates EPI genes in E4.5 ICM and ESCs by both promoter and distal**
783 **binding.** (**a**) Genomic distribution of GABPA binding peaks in E4.5 ICM based on CUT&RUN
784 data. (**b**) Heatmaps showing the DEGs at E4.5 ICM after dTAG13 treatment, as well as the
785 GABPA binding and several TFs motif occurrence around the TSS of the indicated gene groups.
786 (**c**) Percentage of the down-regulated early and late ICM genes (left), and down-regulated ICM
787 genes with GABPA promoter binding (right) at E4.5 ICM. (**d**) Examples of early ICM genes with
788 GABPA promoter binding (shaded) at E4.5 ICM. (**e**) Cumulative distributions of the distance from
789 TSS to the nearest distal GABPA peaks for each group of genes at E4.5 ICM. (**f**) Heatmap showing
790 all the GABPA peaks at promoters, introns and distal regions in E4.5 ICM, and ATAC-seq and
791 H3K27ac signals at these GABPA peak regions with or without dTAG13 treatment. For ATAC
792 and H3K27ac, two replicates in each condition were merged. C: peak center. (**g**) An example of
793 genome browser view of GABPA binding (shaded), ATAC and H3K27ac signals in E4.5 ICM
794 with or without dTAG13 treatment. (**h**) Heatmaps showing the DEGs in 2iESCs after dTAG
795 treatment for 24h, and the GABPA binding, GABPA motif occurrence around the TSS of the
796 indicated gene groups. (**i**) Heatmaps showing similar GABPA binding peaks in E4.5 ICM and
797 2iESCs. Two replicates were merged for GABPA binding analysis in E4.5 ICM and ESCs.
798 Comparison of all DEGs (**j**), down-regulated ICM genes (**k**), and down-regulated EPI genes (**l**)
799 between E4.5 ICM and 2iESCs after dTAG13 treatment.

800

801 **Fig. 5**



802

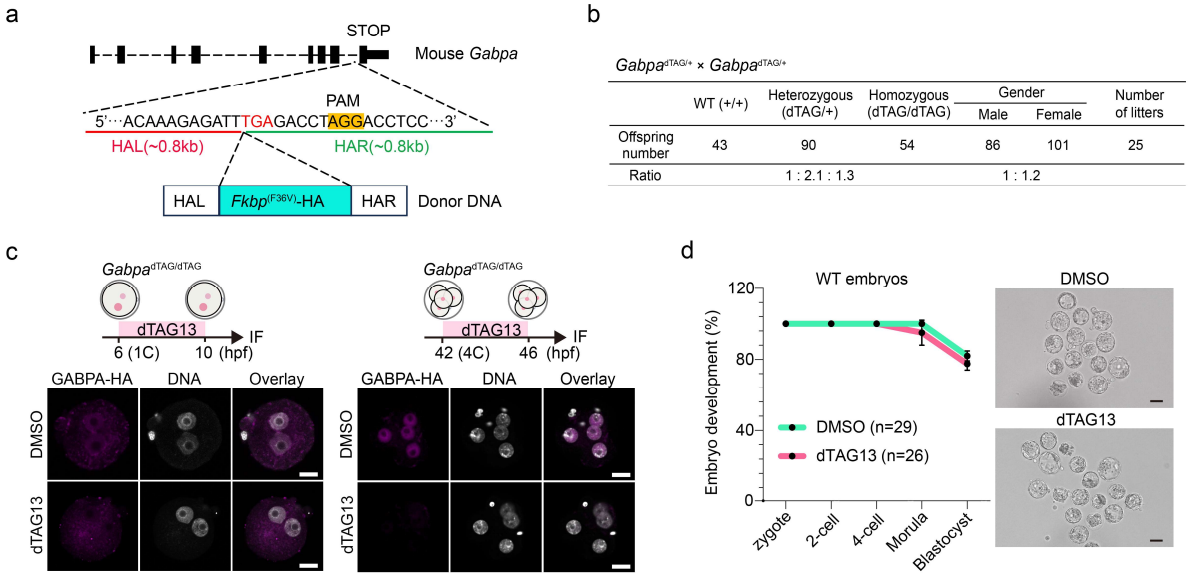
803 **Fig. 5. GABPA genomic binding dynamics during pluripotency establishment and its**
804 **relationship with other pluripotency regulators.** (**a**) Heatmaps showing the dynamics of
805 GABPA genomic binding profile from 2-cell, 8-cell and E3.5 ICM to E4.5 ICM at promoter, intron
806 and distal regions. C: peak center. (**b**) Comparison of GABPA and NR5A2 binding profiles at
807 promoter, intron and distal regions at 2-cell and 8-cell stages. (**c**) GABPA and TFAP2C binding
808 profiles at promoters of early and late ICM genes in 2-cell, 8-cell and E3.5 ICM. (**d**) Genome

809 browser view of two examples showing GABPA and TFAP2C binding dynamics. Promoter
810 binding regions are shaded. (**e**) GABPA and SOX2 binding profiles at promoter, intron and distal
811 regions in E3.5 ICM and E4.5 ICM. (**f**) GABPA and SOX2 binding profiles at promoters of early
812 and late ICM genes in E3.5 ICM and E4.5 ICM. (**g**) Genome browser view of two examples
813 showing the binding (shaded) of GABPA and SOX2 in E3.5 ICM and E4.5 ICM. (**h**) Comparisons
814 of GABPA (E4.5 ICM) and SOX2 (ESC) affected ICM genes (left) and EPI genes (right) after
815 dTAG13 treatments. (**i**) Percentage of ICM and EPI genes regulated by GABPA and SOX2 in E4.5
816 ICM and ESCs. (**j**) A model illustrating different TFs including TFAP2C, SOX2 and GABPA play
817 major roles at different stages during the pluripotency establishment (from 8-cell stage to E4.5
818 ICM stage).

819

820 **Extended data**

821

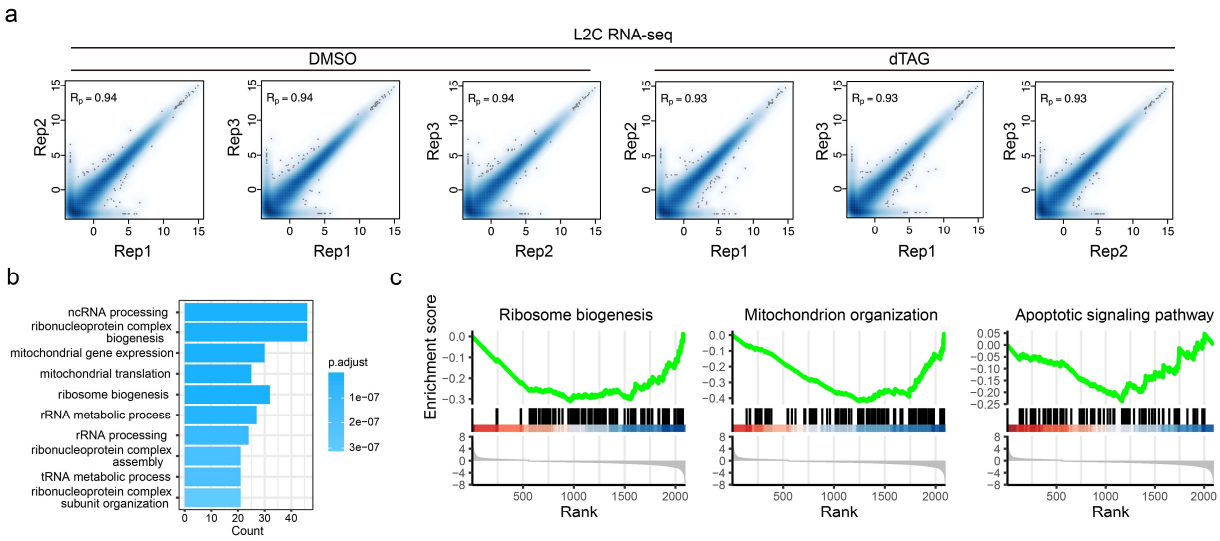


	WT (+/+)	Heterozygous (dTAG/+)	Homozygous (dTAG/dTAG)	Gender		Number of litters
				Male	Female	
Offspring number	43	90	54	86	101	25
Ratio	1 : 2.1 : 1.3			1 : 1.2		

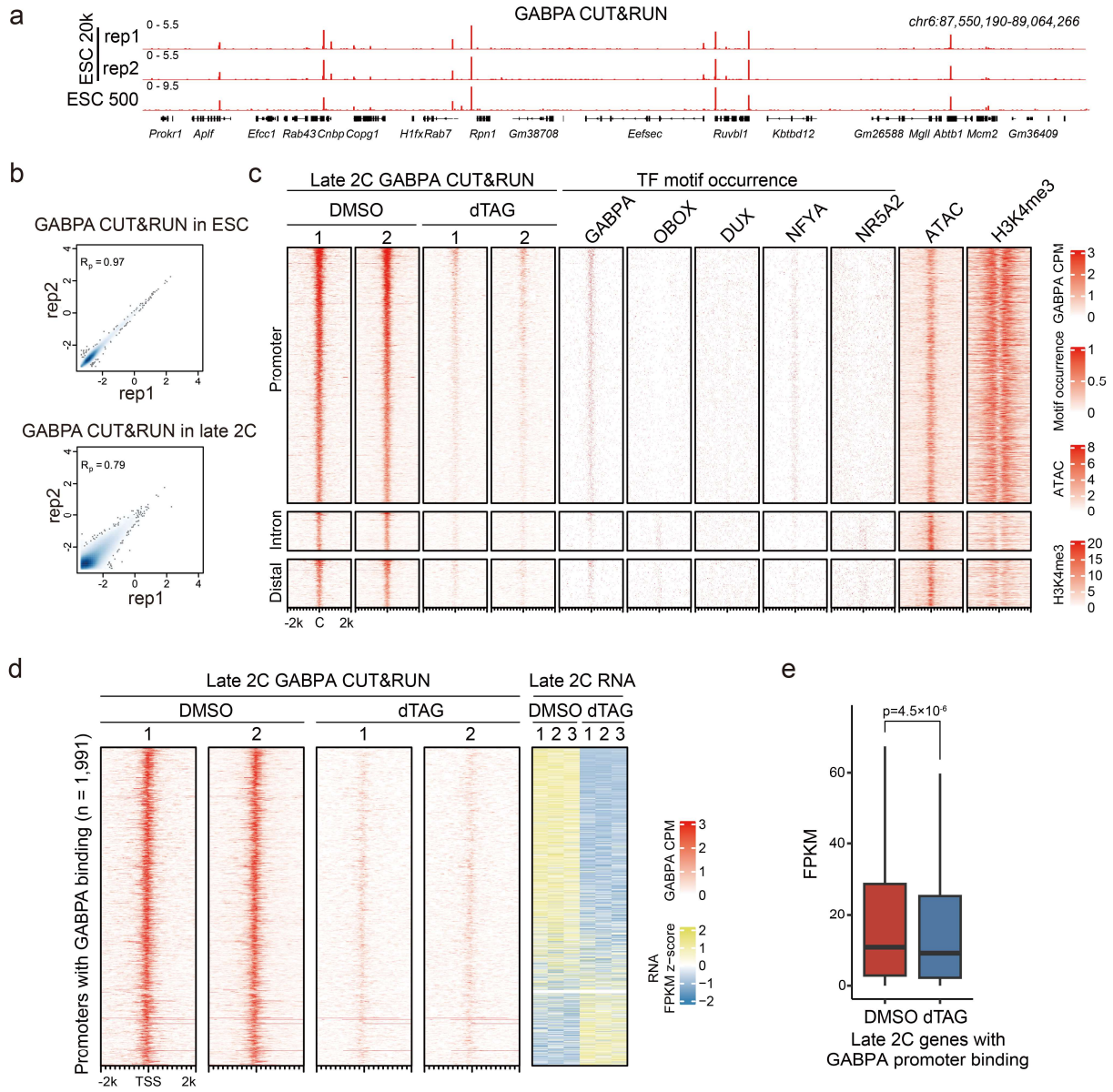
822

823 **Extended Data Fig. 1. Generation of *Gabpa*$^{dTAG/dTAG}$ mice.** (**a**) Diagram of *Fkbp*$^{(F36V)}$-HA
824 knock-in at the 3' end of *Gabpa* gene. (**b**) Genotypes of pups from heterozygous *Gabpa*$^{dTAG/+}$×
825 *Gabpa*$^{dTAG/+}$ crosses. (**c**) Confirmation of GABPA degradation with 4h-dTAG13 treatment in
826 zygote and 4-cell. anti-HA antibody was used for GABPA-HA fusion protein staining. DNA,
827 Hoechst 33342. Scale bar, 20 μm. (**d**) Embryo development of wild-type embryos with or without
828 dTAG 13 treatment.

829



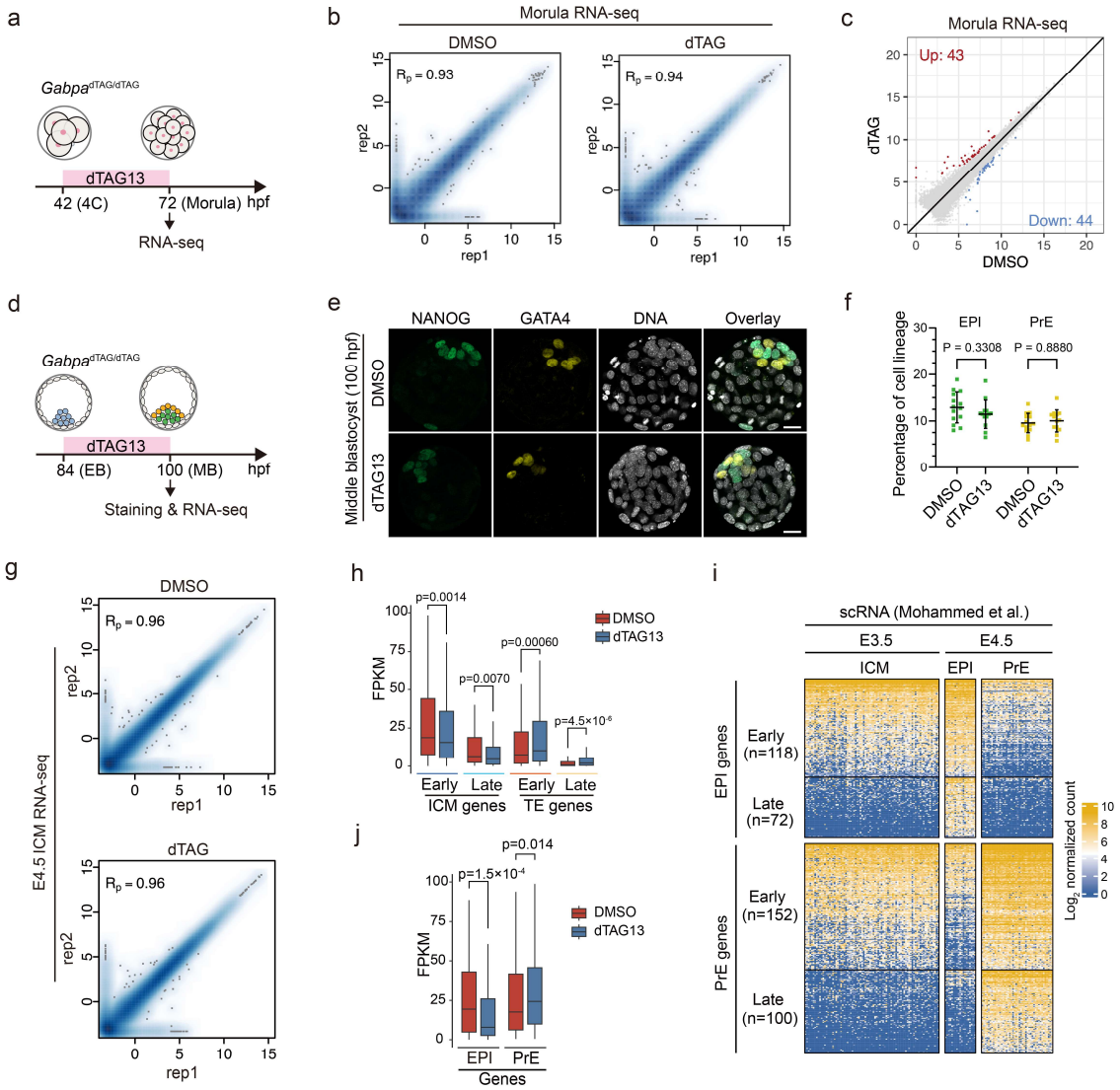
Extended Data Fig. 2. Genes affected by GABPA degradation at late 2-cell. (**a**) Correlations of the replicates of the RNA-seq at late 2-cell stage. R_p: Pearson correlation. (**b**) Gene Ontology (GO) terms enriched for all the down-regulated genes after dTAG13 treatment at late 2-cell stage. (**c**) Gene set enrichment analysis (GSEA) of all the down-regulated genes at late 2-cell stage after dTAG13 treatment for ribosome biogenesis (left), mitochondrion organization (middle) and apoptotic signaling pathway (right).



Extended Data Fig. 3. GABPA binding profiles at the late 2-cell stage. (**a**) Genome browser view showing examples of GABPA CUT&RUN profiles in ESCs using 500 and 20k cells. (**b**) Correlation of replicates for GABPA CUT&RUN with ESCs (top) and late 2-cell embryos (bottom). (**c**) Heatmaps showing the GABPA binding profiles at promoter, intron and distal regions in late 2-cell embryos. TFs motif occurrence, ATAC and H3K4me3 signals are shown at the corresponding regions. C: peak center. (**d**) Heatmaps showing GABPA promoter binding at late 2-cell stage and the corresponding genes expression level with or without the dTAG13 treatment.

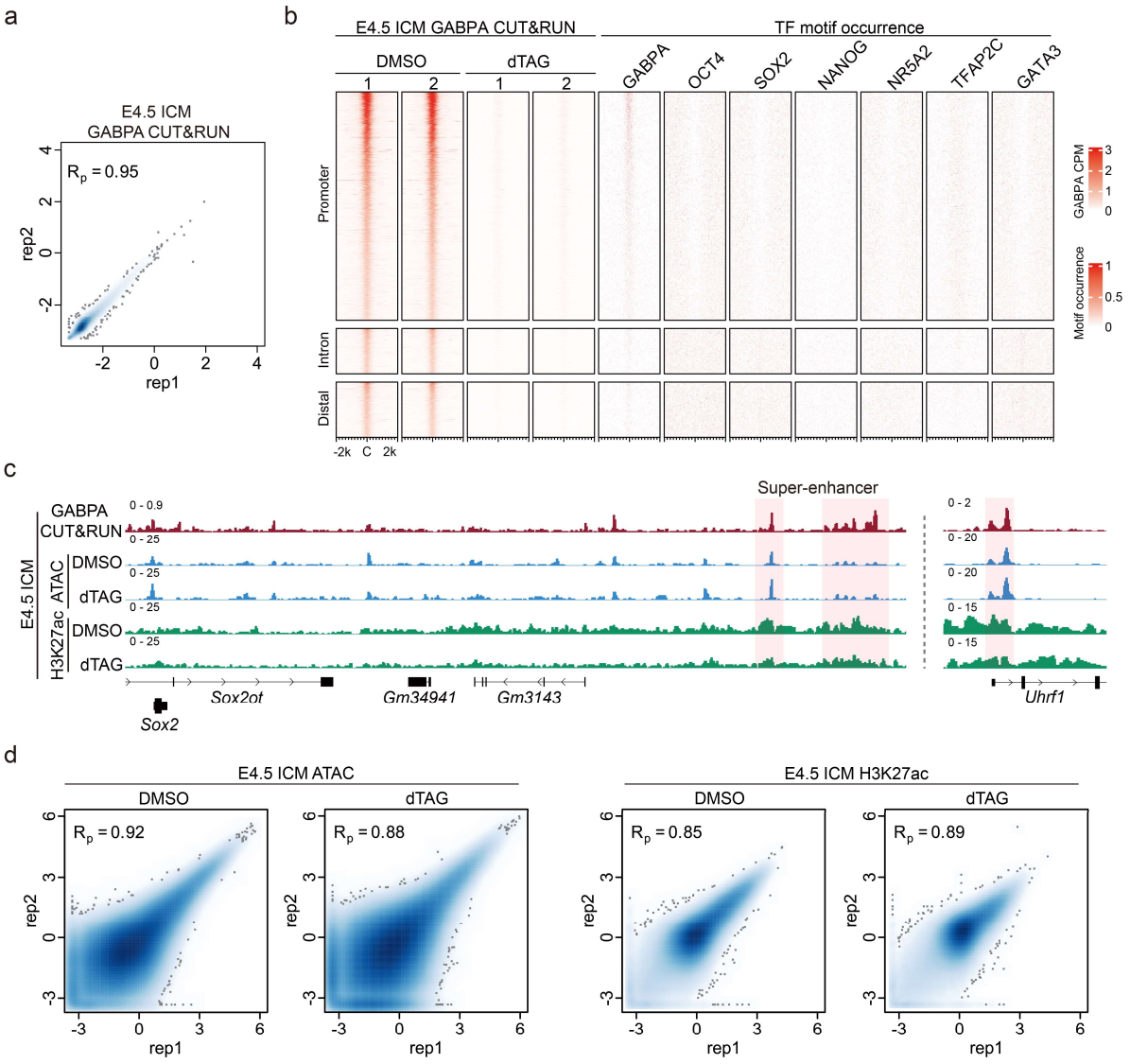
846 (**e**) Boxplot comparing the expression levels of GABPA targeted genes in DMSO and dTAG-13
847 treated late 2-cell. p values were calculated with two-sided Mann-Whitney U test.

848

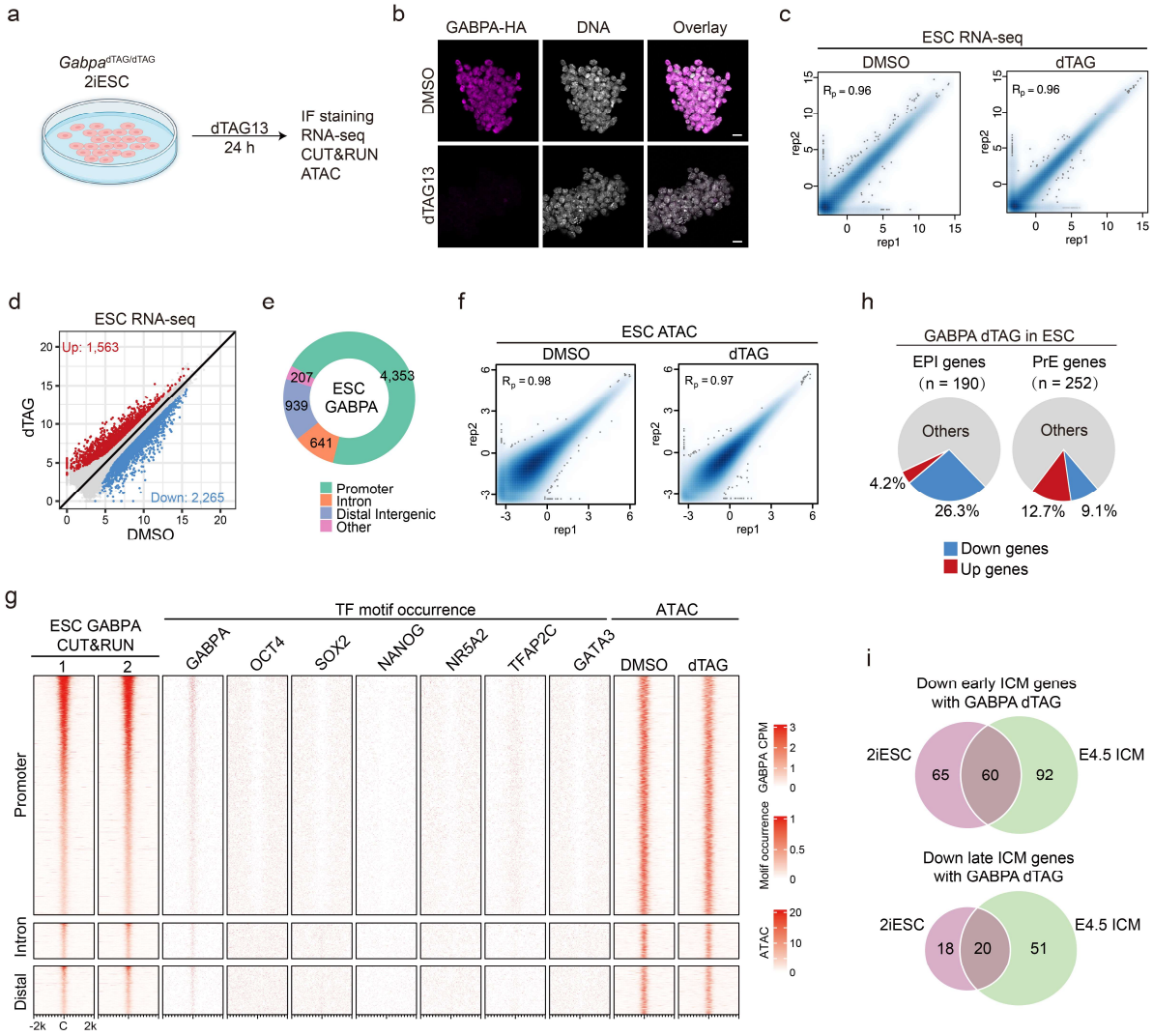


Extended Data Fig. 4. RNA-seq in morula and E4.5 ICM in response to dTAG13 treatment after major ZGA. (**a**) Diagram showing dTAG13 treatment time window and the time of sample collection for RNA-seq. (**b**) Correlation of RNA-seq replicates at morula stage. R_p: Pearson correlation. (**c**) Dot plot showing the DEGs in morula after dTAG13 treatment shown in panel a. (**d**) Diagram showing the times of dTAG13 treatment and sample collection. EB: early blastocyst, MB: middle blastocyst. (**e**) Immunostaining of NANOG (green) and GATA4 (yellow) for middle blastocyst in control (DMSO) and GABPA degradation (dTAG13) conditions. (**f**) Percentage of EPI and PrE cells of middle blastocyst in control (DMSO) and GABPA degradation (dTAG13) conditions. p values were calculated with Student's t-test. (**g**) Correlation of replicates of RNA-seq in E4.5 ICM. (**h**) Boxplot comparing the expression levels of early/late ICM and early/late TE

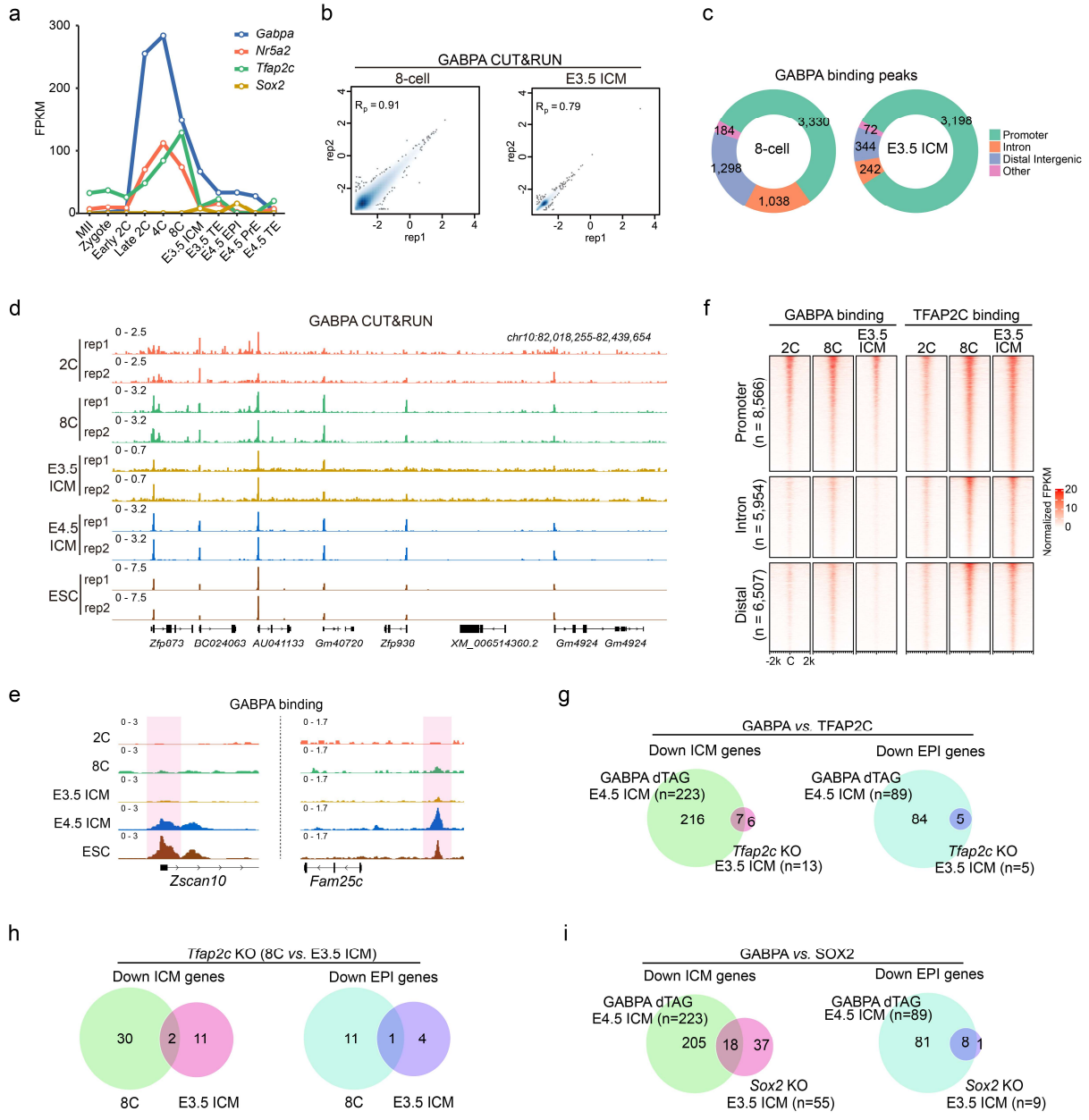
genes at E4.5 ICM with or without dTAG13 treatment. p values were calculated with two-sided Mann-Whitney U test. (**i**) EPI genes and PrE genes identified from single-cell RNA-seq data from Mohammed et al. (**j**) Boxplot comparing the expression levels of EPI and PrE genes at E4.5 ICM with or without dTAG13 treatment. p values were calculated with two-sided Mann-Whitney U test.



Extended Data Fig. 5. GABPA CUT&RUN in E4.5 ICM. (**a**) Correlation of replicates for GABPA CUT&RUN in E4.5 ICM. R_p: Pearson correlation. (**b**) Heatmaps showing the GABPA binding profiles at promoter, intron and distal regions in E4.5 ICM, and the TFs motif occurrence at the corresponding regions. C: peak center. (**c**) Genome browser view of examples of GABPA binding, ATAC and H3K27ac signals at E4.5 ICM. The binding sites are shaded. (**d**) Correlation of replicates for ATAC-seq and H3K27ac CUT&RUN with or without dTAG13 treatment in E4.5 ICM. R_p: Pearson correlation.



Extended Data Fig. 6. GABPA degradation in ESCs. (**a**) Diagram illustrating the GABPA degradation by dTAG13 treatment in *Gabpa*$^{dTAG/dTAG}$ ESCs. (**b**) Immunostaining confirming GABPA degradation after dTAG13 treatment in ESCs. (**c**) Correlations of replicates RNA-seq in ESCs. R_p: Pearson correlation. (**d**) Dot plot showing the DEGs in ESCs after GABPA degradation. (**e**) Genomic distribution of GABPA binding peaks in ESCs. (**f**) Correlation of replicates for ATAC-seq with or without dTAG13 treatment in ESCs. R_p: Pearson correlation. (**g**) Heatmaps showing the GABPA binding profiles at promoter, intron and distal regions in ESCs, and the TFs motif occurrence and ATAC signals at the corresponding regions. C: peak center. (**h**) Percentage of down-regulated EPI and PrE genes after dTAG13 treatment in ESCs. (**i**) Comparisons of the down-regulated early and late ICM genes upon dTAG13 treatment in ESCs and E4.5 ICM.



Extended Data Fig. 7. GABPA binding dynamics and comparisons with other pluripotency regulators. (**a**) Expression dynamics of *Gabpa*, *Nr5a2*, *Tfap2c* and *Sox2* during mouse preimplantation development. (**b**) Correlation of replicates for GABPA CUT&RUN in 8-cell embryos (left) and E3.5 ICM (right). R_p: Pearson correlation. (**c**) Genomic distribution of GABPA binding peaks in 8-cell embryos (left) and E3.5 ICM (right). (**d**) Genome browser view showing examples of the binding profiles of GABPA in 2-cell, 8-cell, E3.5 ICM, E4.5 ICM and ESCs. (**e**)

893 Examples of genome browser view illustrating the GABPA binding (shaded) dynamics during
894 preimplantation development. (**f**) Heatmap comparison of GABPA and TFAP2C binding profiles
895 at promoter, intron and distal regions in 2-cell, 8-cell and E3.5 ICM. (**g**) Comparisons of GABPA
896 (E4.5 ICM) and TFAP2C (E3.5 ICM) affected ICM genes (left) and EPI genes (right). (**h**)
897 Comparisons of TFAP2C affected ICM genes (left) and EPI genes (right) in 8-cell and E3.5 ICM.
898 (**i**) Comparisons of GABPA (E4.5 ICM) and SOX2 (E3.5 ICM) affected ICM genes (left) and EPI
899 genes (right).

900

901 **Supplementary Tables**

902 **Supplementary Table 1.** Gene lists used in this study.

903 **Supplementary Table 2.** Differential expressed genes of late 2-cell treated with GABPA
904 dTAG13 versus that treated with DMSO.

905 **Supplementary Table 3.** Differential expressed genes of morula treated with GABPA dTAG13
906 versus that treated with DMSO.

907 **Supplementary Table 4.** Differential expressed genes of E4.5 ICM treated with GABPA
908 dTAG13 versus that treated with DMSO.

909 **Supplementary Table 5.** Differential expressed genes of ESC treated with GABPA dTAG13
910 versus that treated with DMSO.

911 **Supplementary Table 6.** Oligos used in this study.

912 **Supplementary Table 7.** ddPCR validate *Gabpa*^{dTAG} knock-in copy number

913