The ETS Transcription Factor ERF controls the exit from the naïve pluripotent state

M. Vega-Sendino,[1] T. Olbrich,[1] D. Tillo,[2] A. D. Tran,[3] C. N. Domingo,[1] M. Franco,[1] P. C. FitzGerald,[4] M. J. Kruhlak,[3] and S. Ruiz[1, *]

[1] Laboratory of Genome Integrity, [2] Genetics Branch, [3] Laboratory of Cancer Biology and Genetics and [4] Genome Analysis Unit, CCR, NCI, NIH, Bethesda, MD, USA.

*Corresponding author: sergio.ruizmacias@nih.gov, @sruizmacias

The naïve epiblast undergoes a transition to a pluripotent primed state during embryo implantation. Despite the relevance of the FGF pathway during this period, little is known about the downstream effectors regulating this signaling. Here, we examined the molecular mechanisms coordinating the naïve to primed transition by using inducible ESC to genetically eliminate all RAS proteins. We show that differentiated RAS^{KO} ESC remain trapped in an intermediate state of pluripotency with naïve-associated features. Elimination of the transcription factor ERF overcomes the developmental blockage of RAS-deficient cells by naïve enhancer decommissioning. Mechanistically, ERF regulates NANOG expression and ensures naïve pluripotency by strengthening naïve transcription factor binding at ESC enhancers. Moreover, ERF negatively regulates the expression of the *de novo* methyltransferase DNMT3B, which participates in the extinction of the naïve transcriptional program. Collectively, we demonstrated an essential role for ERF controlling the exit from naïve pluripotency during the progression to primed pluripotency.

Teaser: ERF is the MAPK-dependent switch controlling the transition between naïve and primed pluripotency during embryonic development.

INTRODUCTION:

Embryonic cells residing within the inner cell mass (ICM) are pluripotent as they have the potential to generate all cell lineages of the organism. The state of pluripotency found in the preimplantation epiblast, prior to any lineage specification, is usually defined as naïve (*1*). However, as the embryo expands and develops after implantation, epiblast cells become individually fated, although still retain pluripotent features. This state of pluripotency associated to the post-implantation epiblast is defined as primed (*2, 3*). More than distinct pluripotent states, naïve and primed can be viewed as different phases of a coordinated developmental progression where naïve unbiased cells differentiate responding to inductive cues to initiate a multi-lineage decision commitment at gastrulation (*2, 3*).

Naïve and primed pluripotent states can be captured *in vitro* by defined culture conditions. Mouse embryonic stem cells (ESC) can be propagated in naïve conditions by using a combination of a MEK inhibitor (PD0325091) plus a glycogen synthase kinase-3 inhibitor (CHIR99021; hereafter referred as the 2i condition) (*4*). ESC grown under these conditions resemble embryonic cells residing in the pre-implantation embryonic (E) post-fertilization day E3.5-E4.25 embryo and have been extensively used to study the ground state of pluripotency. In addition, cultures of embryonic cells that retain primed pluripotency can also be established from post-implantation embryos (*5*). These epiblast stem cells (EpiSC) resemble transcriptionally the late epiblast layer of the post-implantation embryo at E6.0-E6.5. Both types of cells, naïve and primed, share the expression of core pluripotent transcription factors (TF), such as OCT4 and SOX2. However, 2i-ESC express specific naïve TF (REX1, NANOG or KLF4), which are absent in EpiSC, whereas EpiSC express epiblast specific genes (OCT6 or OTX2), absent in ESC. Importantly, these pluripotent cell types are interconvertible by modifying the culture conditions or expressing specific TF (*3*). Recently, a distinct intermediate state of pluripotency has been identified, the rosette pluripotent state (*6*). Rosette embryonic stem cells (RSC) co-express naïve TF and the primed marker OTX2. Inhibition of WNT signaling drives the pluripotent transition from naïve to the rosette state while further activation of the RAS/MAPK signaling promotes the progression to primed pluripotency (*6*). Furthermore, additional states of intermediate pluripotency characterized by germ cell specification capacity have also been described, formative stem cells

59 (FSC) and chimera pluripotent stem cells (XPSC) (*7, 8*). These intermediates rely on exogenous

60 (XPSC) or autocrine (FSC) FGF signaling for their self-renewal in contrast to the requirements for

61 RSC. In summary, pluripotency can be considered as a dynamic property associated to different

62 stem cell states supported by defined pluripotent transcriptional networks.

63 Exit from naïve pluripotency during implantation is essential to fate epiblast cells residing in the

64 ICM with inductive signals prior to gastrulation. However, the molecular mechanisms

65 coordinating the naïve to primed transition are not fully understood. Moreover, despite the

66 relevance of the RAS pathway during this critical period, little is known about the downstream

67 effectors regulating the MAPK transcriptional program. We recently identified the transcriptional

68 repressor ERF, a member of the E26 transformation specific (ETS) family, as an important

69 regulator downstream of the RAS pathway in ESC (*9*). ERF is a TF that shuffles between the nuclei

70 and cytoplasm in ESC in a phosphorylation-dependent manner (*10*). In the absence of RAS/MAPK

71 signaling, ERF remains unphosphorylated and chromatin-bound in the nuclei while growth factor

72 stimulation keeps ERF inactive in the cytoplasm by ERK-dependent phosphorylation (*10-12*).

73 Importantly, the precise role of ERF during early embryonic development and pluripotent

74 transitions is unclear. Here, we show that ERF plays a dual role in the transition from naïve to

75 primed pluripotency. First, ERF binds to ESC super-enhancers to ensure optimal expression of

76 naïve TF in the absence of FGF signaling. Second, activation of MAPK signals induces the release

77 of chromatin-bound ERF, an event that is necessary and sufficient to trigger full commitment to

78 primed pluripotency. In addition, we also show that ERF controls the expression of the *de novo*

79 methyltransferase DNMT3B through LIN28 regulation, leading to the inactivation of the naïve

80 transcriptional program. Our data demonstrate that ERF is the MAPK-dependent switch that

81 controls the progression to primed pluripotency.

82

83

84

85

86

87

88   **RESULTS:**

89

90   **ERF is expressed in the naïve pluripotent epiblast**

91   We first sought to determine the precise expression timing for ERF during embryonic

92   development. For this, we collected mouse embryos at different developmental stages and

93   performed immunofluorescence analyses. We determined that ERF expression peaks around

94   E3.5-E4.0 and is coincidental with the naïve pluripotent epiblast (Fig. 1A-C). ERF is expressed in

95   the naïve inner cell mass (ICM) and co-expressed with the naïve pluripotent markers NANOG and

96   KLF4 (Fig. 1B, C). In addition, we also detected ERF in the trophectoderm (TE), consistent with its

97   reported requirement for chorionic trophoblast differentiation (Fig. 1B, C) (*13*). Exit from naïve

98   pluripotency *in vivo* and downregulation of naïve associated markers strongly correlated with

99   decreased expression of ERF (Fig. 1B, C). We confirmed this expression pattern by analyzing

100   single-cell RNAseq datasets from embryos at different stages (fig. S1A) (*14*). These results showed

101   that ERF is upregulated during ICM formation and it is quickly downregulated before

102   implantation, suggesting a role for ERF in the naïve epiblast.

103   To validate our observations and assess the role of the MAPK pathway in ERF levels, we used an

104   ESC model deficient in H-RAS and N-RAS with a conditional knockout K-RAS triggered by the

105   addition of 4-hydroxytamoxifen (OHT) (RAS$^{lox/lox}$ hereafter) (*9, 15*). These ESC allowed us to

106   generate cells devoid of all RAS proteins. We next applied a well-established *in vitro* protocol

107   using FGF2 and Activin-A (FA, hereafter) to differentiate naïve ESC, growing in 2i+LIF (2iL)

108   conditions, to post-implantation epiblast like stem cells (EpiLSC) and examined ERF levels (*16,*

109   *17*). While RAS$^{lox/lox}$ 2iL-ESC showed high levels of ERF, these cells exhibited negligible levels when

110   differentiated into EpiLSC (Fig. 1D). ERF is quickly phosphorylated upon differentiation and the

111   decrease in protein levels is due to transcriptional repression (fig. S1B) (*18*). We next examined

112   whether this downregulation depended on FGF/MAPK activation. Indeed, RAS-deficient cells

113   (RAS^{KO}) retained elevated levels of ERF after differentiation (Fig. 1D). Together, these results

114   pointed to a role for ERF in the exit from naïve pluripotency in a MAPK-dependent manner.

115

116

4

117 **Downregulation of ERF is necessary for the successful exit from naïve pluripotency**

118 To evaluate the implication of ERF in the exit from naïve pluripotency, we examined its expression

119 along with markers of naïve pluripotency during the formation of embryonic rosettes. In this

120 assay, single cell suspensions of ESC embedded in Matrigel exit from naïve pluripotency and

121 develop into polarized rosettes that undergo lumenogenesis, mimicking the morphogenic events

122 of the epiblast (*19*). This *in vitro* system recapitulates the development of the ICM during peri-

123 implantation. Importantly, ESC rosettes generated under 2iL conditions lack lumen, continue to

124 express naïve markers, and eventually become disorganized over time (*20*). This showed that exit

125 from naïve pluripotency is necessary for successful polarization and lumenogenesis. Upon 2iL

126 removal and resuspension in Matrigel, RAS$^{lox/lox}$ rosettes showed polarization and lumen,

127 expression of the sialomucin protein podocalyxin (PDX) and downregulation of the naïve marker

128 NANOG (Fig. 2A-C and fig. S2A and B). However, RAS^{KO} rosettes developed into a disorganized

129 group of cells, which lacked expression of PDX and retained naïve pluripotency markers (Fig. 2A-

130 C and fig. S2A and B). To test whether ERF regulates the exit from naïve pluripotency, we

131 generated ERF-knockouts (ERF^{KO}) and evaluated rosette formation in all different genotypes

132 (RAS$^{lox/lox}$, ERF^{KO}, RAS^{KO}, and RAS^{KO}; ERF^{KO}). Elimination of ERF is sufficient to induce exit from

133 naïve pluripotency and to rescue the failed morphogenic events (polarization and

134 lumenogenesis) observed in RAS^{KO} rosettes (Fig. 2A-C and fig. S2A and B). To further support

135 these observations, we generated reporter cell lines in our ESC by replacing the endogenous

136 coding sequence of the gene REX1 (also known as ZFP42) with a short half-life form of eGFP

137 (REX1-deGFP). Similar REX1-deGFP reporter ESC have been widely used as a faithful system to

138 monitor exit from naïve pluripotency (*21*). Induction of EpiLSC by FA in RAS$^{lox/lox}$ and ERF^{KO} ESC

139 demonstrated efficient downregulation of the reporter while RAS^{KO} ESC showed no signs of

140 downregulation (Fig. 2D). However, exit from naïve pluripotency was efficiently achieved in

141 RAS^{KO}; ERF^{KO} ESCs (Fig. 2D). Finally, we also examined the clonogenicity potential of RAS^{KO}; ERF^{KO}

142 ESC primed to exit naïve pluripotency. For this, cultures of ESC (RAS$^{lox/lox}$, ERF^{KO}, RAS^{KO}, and RAS^{KO};

143 ERF^{KO}) were withdrawn of 2iL for 2 days and plated back in 2iL conditions for a colony forming

144 assay. Cells that exit naïve pluripotency are irreversibly committed and have lost the ability to

145 generate colonies in 2iL conditions. Indeed, while RAS^{KO} ESC were able to generate alkaline

5

146   phosphatase positive colonies and remained trapped in a naïve pluripotent state, RAS^{KO}; ERF^{KO}

147   ESCs lost this ability (Fig. 2E). Our results demonstrate a predominant role for ERF controlling the

148   exit from naïve pluripotency.

149

150   **ERF controls the progression to primed pluripotency**

151   To obtain mechanistic insights on the rescue mediated by the loss of ERF during the exit from

152   naïve pluripotency, we performed RNA sequencing (RNAseq) analysis of ESC cultured under naïve

153   conditions (2iL) or differentiated into EpiLSC in FA. Principal component plot (PCA) analysis

154   segregated the samples based on their differentiation status alongside PC1, and MAPK activity

155   alongside PC2 (Fig. 3A). Interestingly, FA-RAS^{KO} ESC are localized at an intermediate state

156   between naïve and primed pluripotency (Fig. 3A). FA-RAS^{KO} ESC are characterized by the

157   intermediate to high expression of naïve pluripotent markers as well as of primed associated

158   genes including OTX2 (Fig. 3B, C). Indeed, while naïve pluripotent ESC are characterized by a

159   NANOG$^+$/OTX2$^-$/OCT6$^-$ state, FA-RAS^{KO} ESC showed a NANOG$^+$/OTX2$^+$/OCT6$^-$ state (Fig. 3B, C).

160   Furthermore, although FA-RAS$^{lox/lox}$ ESC have undetectable levels of ERF upon differentiation, FA-

161   RAS^{KO} ESC retain high ERF expression (Fig. 3C). Using available datasets that evaluate temporal

162   transcriptional dynamics during the transition from naïve to primed pluripotency, we determined

163   that FA-RAS^{KO} ESC transcriptionally resembled to ESC primed for 12-24 hours (Fig. 3D and fig.

164   S3A) (*18*). Interestingly, FA-RAS^{KO} ESC are reminiscent of the recently described intermediate

165   pluripotent states, rosette and formative pluripotent states (RSC, FSC and XPSC) (*6-8*). PCA

166   analyses showed that PC1 (differentiation status) placed all intermediate pluripotent states in a

167   similar transcriptional space but PC2 segregated FA-RAS^{KO} ESC from rosette or formative

168   pluripotent states (Fig. 3E). Hierarchical clustering analysis based on the expression of naïve and

169   primed markers revealed that RSC and FA-RAS^{KO} ESC are transcriptionally more comparable,

170   likely to the defective MAPK signaling in both intermediate states (Fig. 3F). Finally, we observed

171   that prolonged culture of FA-RAS^{KO} ESC in FGF2/Activin-A/XAV939 (FAX) conditions

172   demonstrated their trapping in this intermediate state of pluripotency (fig. S3B). Our results

173   showed that loss of ERF is necessary and sufficient to overcome the developmental blockage of

174   FA-RAS^{KO} ESC in its intermediate pluripotent state. Indeed, RNAseq data revealed that loss of ERF

6

175   in RAS^{KO}; ERF^{KO} ESC restored the overall gene expression profile to be indistinguishable from FA-

176   RAS$^{lox/lox}$. In summary, our data point to ERF as the MAPK-dependent switch that triggers full

177   commitment to primed pluripotency.

178

179   **Chromatin-bound ERF ensures an optimal naïve pluripotency state**

180   The expression of ERF correlates with high expression levels of naïve markers *in vivo* and *in vitro*

181   and is quickly downregulated upon induction to primed pluripotency (Fig. 1). Moreover, ERF is

182   enriched in a total of 2074 bona-fide ESC enhancers identified in RAS^{KO} ESC (Supplemental Data

183   S1 and 2) (*9*). Among these, ERF binds to most of the super-enhancers identified in ESC (198/231)

184   (Supplemental Data S3) (*9, 22, 23*). ERF-bound super-enhancers are associated to highly

185   transcribed naïve genes (KLF4, ESRRB, PRDM14, NANOG, TBX3 or ZFP42) (some examples in fig.

186   S4A) as well as to general pluripotent genes (OCT4 or SOX2). Although ERF is considered to be a

187   transcriptional repressor, we hypothesized that ERF might play a different role at ESC enhancers.

188   To explore the relevance of ERF at these enhancers, we first examined the level of occupancy of

189   the pluripotent transcription factors OCT4, SOX2 and NANOG (O, S and N, respectively) in the

190   subset of ERF-bound 2074 ESC enhancers. We observed increased occupancy of bound OSN at

191   these compared to a non-ERF bound randomized set of different 2074 enhancers (Fig. 4A). In

192   addition, ERF-bound ESC enhancers are also characterized by higher H3K27Ac levels, increased

193   p300 binding and chromatin accessibility detected by ATACseq (fig. S4B). These results suggested

194   that ERF could be regulating the activity of these enhancers, and thus, the expression of their

195   associated genes. Therefore, we examined the expression levels of essential naïve pluripotent

196   genes associated to ERF-bound super-enhancers in RAS$^{lox/lox}$ and ERF^{KO} ESC grown in 2iL

197   conditions. Interestingly, NANOG, PRDM14 and ZFP42 showed decreased expression in 2iL-ERF^{KO}

198   ESC, while TBX3, KLF4 and ESRRB do not (Fig. 4B). It has been shown that NANOG promotes

199   chromatin accessibility and binding of pluripotent factors to enhancers (*24*). Thus, we

200   hypothesized that reduced levels of NANOG could affect the expression of additional naïve-

201   associated genes. Indeed, ERF^{KO} cells showed an overall decreased expression of the naïve

202   pluripotent transcriptional network (Fig. 4C). Consistent with these results, unidimensional PCA

203   analysis (PC1) segregating samples based on their differentiation status showed that 2iL-ERF^{KO}

7

204 ESC are biased toward differentiation (Fig. 4D). These combined results revealed that ESC

205 growing under naïve conditions required ERF binding at enhancers to maintain an optimal naïve

206 pluripotent state. To confirm this observation, we employed native Cut&Run sequencing to

207 examine SOX2 and NANOG occupancy in RAS$^{lox/lox}$ and ERF^{KO} ESC grown in 2iL conditions. As

208 predicted, both pluripotent transcription factors showed an overall decreased enrichment in ESC

209 enhancers in naïve ERF^{KO} ESC (Fig. 4E, F and fig. S4C). These data support our findings and

210 revealed an unexpected unique role for ERF in promoting naïve pluripotency.

211

212 **The naïve enhancer landscape is active in FA-RAS^{KO} ESC**

213 Inactivation of the naïve transcriptional network is necessary for the successful progression to

214 primed pluripotency. This is associated with decommission of naïve enhancers by

215 downregulation of naïve transcription factors including ESRRB, KLF4 or NANOG and a widespread

216 OCT4 genomic relocation (*25*). Indeed, OCT4 shifts from enhancers associated with key players

217 in naïve pluripotency and engages in new enhancer elements at genes implicated in post-

218 implantation development. Using datasets for OCT4 occupancy in ESC and EpiLSC and based on

219 the relative occupancy between both cellular states, we defined top OCT4 sites preferentially

220 enriched in ESC (4759 sites, OCT4^{ESC}), EpiLSC (2921 sites, OCT4^{EpiLSC}) and commonly shared

221 between the two (9144 sites, OCT4^{Common}) (fig. S5A and Supplemental Data S4-6). Genes

222 associated with OCT4^{ESC} sites are mostly downregulated during the transition to primed

223 pluripotency whereas genes associated to OCT4^{EpiLSC} sites are upregulated (*25*). Out of all OCT4

224 bound sites, a total of 12.52% (596/4759 ERF/OCT4^{ESC} sites), 13.62% (1246/9144 ERF/OCT4^{Common}

225 sites) and 0.15% (44/2921 ERF/OCT4^{EpiLSC} sites) were co-occupied by ERF in naïve ESC (Fig. 5A).

226 This suggested that ERF does not play a specific role in genes associated with EpiLSC-specific

227 sequences but instead in regulating OCT4^{ESC} and OCT4^{Common} sites. We first focused on OCT4^{ESC}

228 sites as we observed that FA-RAS^{KO} ESC remain trapped in an intermediate state between naïve

229 and primed pluripotency and showed expression of naïve pluripotent markers (Fig. 3A-C).

230 Differentially expressed genes associated with ERF/OCT4^{ESC} sites were mostly downregulated

231 during FA differentiation in RAS$^{lox/lox}$ ESC (Fig. 5B). While many of these genes did not change

232 their expression level in FA-RAS^{KO} ESC, those that changed showed a lower differential expression

233   extent (Fig. 5B). Enhancer decommission of naïve enhancers by OCT4 relocation is followed by

234   changes in enhancer chromatin patterns including decreased levels of H3K27ac (*24*). Thus, we

235   examined whether the naïve enhancer landscape is still fully active in FA-RAS^{KO} ESC by performing

236   native Cut&Run sequencing to evaluate the levels of H3K27ac as a marker for active enhancers

237   and NANOG occupancy in ESC and EpiLSC from all genotypes. Interestingly, 2iL-RAS^{KO} ESC showed

238   increased acetylation and NANOG occupancy at ERF/OCT4^{ESC} sites (Fig. 5C, D), a phenotype that

239   is also observed in all OCT4^{ESC} sites (fig. S5B). This increase in H3K27ac levels does not result in

240   major overall transcriptional changes and the expression level of naïve associated genes is similar

241   between RAS$^{lox/lox}$ and RAS^{KO} ESC (Fig. 4C). This observation supports the idea that levels of

242   H3K27ac does not necessarily determine enhancer activity but rather discriminates between

243   active or poised enhancers (*26*). Exit from naïve pluripotency correlated with an overall decrease

244   in H3K27ac levels, negligible chromatin-bound NANOG and reduced expression of naïve-

245   associated genes in RAS$^{lox/lox}$ ESC (Fig. 5C, D and fig. S5B). However, FA-RAS^{KO} ESC showed

246   elevated gene expression, NANOG occupancy and H3K27ac levels at ERF/OCT4^{ESC} sites to, in some

247   cases, a comparable level as detected in naïve RAS$^{lox/lox}$ ESC (Fig. 5C, D). Together, these data

248   showed that naïve-specific associated OCT4 binding sites remain fully active in the FA-RAS^{KO} ESC

249   intermediate state.

250   We next focused specifically on the regulatory sequences and expression level of naïve

251   transcription factors associated with ERF-bound super-enhancers. Among these, the expression

252   level of ERF-dependent genes (NANOG, PRDM14 and ZFP42) is unaffected in FA-RAS^{KO} ESC and is

253   comparable to that of RAS$^{lox/lox}$ ESCs as ERF is still expressed (Fig. 5E). However, ERF-independent

254   genes (ESRRB, TBX3 and KLF4) showed a significant decrease in expression, suggesting that

255   additional mechanisms ensure optimal expression of these genes in 2iL-ESC (Fig. 5E). Collectively,

256   these results demonstrated that FA-RAS^{KO} ESC retain an active naïve transcriptional network.

257   Although some naïve markers showed a marked decreased expression at this stage, elevated

258   levels of NANOG, PRDM14 or ZFP42 likely sustain the naïve like-state in FA-RAS^{KO} ESC mediated

259   by ERF-dependent mechanisms.

260

261

9

262   **OTX2 co-occupy binding sites with NANOG in FA-RAS[KO] ESC**

263   Global reorganization of OCT4 genomic binding is driven by increased expression of OTX2 during

264   naïve to primed transition (*25*). In fact, ectopic OTX2 expression in 2iL-ESC shows that it behaves

265   as a pioneering factor engaging in previously inaccessible enhancer sites, relocates OCT4 to these

266   sites, and induces the expression of primed-associated genes (*25*). OTX2 expression is

267   independent of MEK signals but is efficiently repressed by the WNT pathway, thus explaining the

268   elevated levels of OTX2 in rosette-like ESC (*6*). High levels of OTX2 have also been found to be

269   associated to formative pluripotency (*7, 8*). In addition, we also detected similar high OTX2 levels

270   in FA-RAS[KO] ESC compared to FA-RAS[lox/lox] (Fig. 6A). We have shown that ERF binding sustains

271   optimal naïve transcription factor expression (Fig. 4) and thus, we asked whether ERF is involved

272   in OTX2 regulation. Indeed, ERF binds to the super-enhancer region associated with OTX2 in

273   RAS[KO] ESC. Furthermore, the low levels of OTX2 expression observed in 2iL-RAS[lox/lox] ESC are

274   further decreased in 2iL-ERF[KO] ESC (Fig. 6A, B). This decrease correlates with lower NANOG and

275   SOX2 binding and suggests that ERF binding might prevent further OTX2 repression in the

276   absence of FGF signaling (Fig. 6B). In agreement, FA-RAS[KO] ESC showed strong NANOG

277   enrichment as well as OTX2 itself, which could potentially sustain its own expression after

278   NANOG downregulation (fig. S6A). We next focused on the ERF/OCT4[Common] sites as they gained

279   OTX2 and H3K27ac enrichment in EpiLSC compared to ESC (*24*). Differentially expressed genes

280   associated with ERF/OCT4[Common] sites were mostly upregulated in FA-RAS[lox/lox] (Fig. 6C). Similar

281   to what we observed with genes associated with ERF/OCT4[ESC] sites, many of these genes did not

282   change their expression levels in FA-RAS[KO] ESC and those that changed showed a lower

283   differential expression extent (Fig. 6C). Based on the elevated levels of NANOG expression in FA-

284   RAS[KO] ESC, we hypothesized that ERF/OCT4[Common] sites also retained chromatin bound NANOG.

285   While FA-RAS[lox/lox] showed negligible enrichment of NANOG at these sites, FA-RAS[KO] ESC retained

286   NANOG bound (Fig. 6D). In addition, we also observed stronger OTX2 enrichment at these sites

287   in FA-RAS[KO] ESC compared to FA-RAS[lox/lox] (Fig. 6D). Interestingly, we also detected OTX2 binding

288   in FA-RAS[KO] ESC at the ERF/OCT4[ESC] sites, naïve associated sequences that are decommissioned

289   during the transition to EpiLSC (Fig. 6D). These observations were also confirmed globally in all

290   OCT4 sites (fig. S6B). Furthermore, OTX2 as well as NANOG were also strongly enriched at the

10

291 OCT4^{EpiLSC} sites in FA-RAS^{KO} ESC, suggesting that, besides OCT4, OTX2 might also relocate

292 additional naïve pluripotent factors (if expressed) to these sites prior to full differentiation (Fig.

293 6D and fig. S6B). Finally, we examined whether OTX2 and NANOG enrichment correlated with

294 increased expression of post-implantation epiblast genes associated with ERF/OCT4^{Common} sites

295 in FA-RAS^{KO} ESC. We found genes showing a similar level of expression in FA-RAS^{KO} ESC compared

296 to FA-RAS$^{lox/lox}$ (Fig. 6E and fig. S6C). Our data demonstrate that FA-RAS^{KO} ESC show promiscuous

297 OTX2 and NANOG binding at OCT4^{ESC} sites and OCT4^{EpiLSC} sites, which correlates with the

298 expression of both naïve and primed markers in preparation for the transition towards primed

299 pluripotency.

300

301 **ERF controls the expression of LIN28 proteins**

302 We next investigated how ERF controls the exit from the developmental blockage of RAS-

303 deficient cells by mining our RNAseq data (Fig. 3). By combining differential gene expression

304 between our different genetic conditions and differentiation status together with nearby ERF

305 binding occupancy, we identified LIN28A and B as putative regulators. Expression of LIN28A/B is

306 low in 2iL-ESC but is upregulated by FGF signaling during the transition to primed pluripotency.

307 Of note, LIN28A and/or LIN28B-deficient ESC showed impaired conversion into EpiLSC, revealing

308 a critical role for these proteins in regulating the exit from naïve pluripotency (*27*). Induction of

309 EpiLSC differentiation in RAS$^{lox/lox}$ and ERF^{KO} ESC demonstrated efficient LIN28s upregulation

310 while FA-RAS^{KO} ESC showed low or negligible levels of LIN28A and LIN28B, respectively (Fig. 7A).

311 ERF deficiency restored the levels of LIN28 proteins in FA-RAS^{KO}; ERF^{KO} (Fig. 7A). Interestingly,

312 2iL-ERF^{KO} ESC showed already elevated levels of LIN28 proteins before differentiation, suggesting

313 that ERF mediates a direct negative regulation on these genes. Indeed, we found ERF bound to

314 an enhancer region in the second intron of LIN28 proteins in RAS^{KO} ESC (Fig. 7B and fig. S7A).

315 Moreover, transition to EpiLSC correlated with increased levels of H3K27ac at the LIN28

316 enhancers, and thus, LIN28 expression (Fig. 7B).

317 LIN28 proteins are RNA binding proteins known for binding to and inactivating the let-7 microRNA

318 family. Functionally, let-7 microRNAs target a number of mRNA transcripts for degradation

319 including MYC, RAS, HMGA2 and the two *de novo* methyltransferases DNMT3A and DNMT3B

11

320   (*28*). Indeed, naïve ESC exhibit low levels of genome wide CpG methylation, which increase during

321   the transition to EpiLSC, correlating with the silencing of the naïve transcriptional program (*29,*

322   *30*). Accordingly, induction of RAS$^{lox/lox}$ and ERF^{KO} ESC into EpiLSC by FA demonstrated efficient

323   upregulation of DNMT3 proteins while FA-RAS^{KO} ESC showed low levels of DNMT3B (Fig. 7A). As

324   expected, expression of DNMT3 proteins was efficiently rescued in FA-RAS^{KO}; ERF^{KO} (Fig. 7A). We

325   next examined the levels of 5-methylcytosine (5mC) by dot blot in all different genotypes,

326   RAS$^{lox/lox}$, ERF^{KO}, RAS^{KO}, and RAS^{KO}; ERF^{KO}, grown under naïve or primed conditions (fig. S7B). The

327   low levels of 5mC detected in 2iL rapidly increased upon FA treatment in RAS$^{lox/lox}$, ERF^{KO} and

328   RAS^{KO}; ERF^{KO} ESC. However, FA-RAS^{KO} ESC remained largely hypomethylated (fig. S7B). To get

329   further insights into methylation dynamics on a genome-wide manner, we performed reduced

330   representation bisulfite sequencing (RRBS) in all genotypes and pluripotent states (Fig. 7C, D)

331   (*31*). 2iL-RAS$^{lox/lox}$ ESC presented low levels of methylation (6.55% of analyzed CpG sites are

332   methylated), which were increased after differentiation (21.75%). Furthermore, 2iL-ERF^{KO} ESC

333   showed higher average methylation levels in 2iL (12.1%) and FA (24.7%) conditions, consistent

334   with their bias toward differentiation and their less optimal naïve transcriptional network (Fig.

335   7C, D). Conversely, RAS^{KO} ESC showed extremely low methylation levels in 2iL conditions (1.55%)

336   reaching similar levels to those found in 2iL-RAS$^{lox/lox}$ ESC when differentiated (6.9%) (Fig. 7C, D).

337   As expected, the defect in methylation observed in RAS^{KO} ESC was rescued upon ERF deletion

338   (Fig. 7C, D). Interestingly, ERF^{KO} ESC showed higher level of methylation around transcription start

339   sites as well as in CpG islands (Fig. 7E), especially after differentiation. Collectively, our data

340   showed that ERF regulates negatively the expression of Lin28 and DNMT3 proteins. The altered

341   expression of these proteins might underlie the global disbalance in the methylation levels at a

342   genome-wide level in ERF-deficient cells.

343

344

345

346

347

348

DISCUSSION

Preceding implantation, the cells residing within the naïve ICM of the blastocyst transition through a distinct phase of primed pluripotency in preparation for lineage commitment. This transition is initiated by fibroblast growth factor 4 (FGF4)-dependent activation of the RAS/MAPK pathway (*32, 33*). Despite the relevance of the MAPK pathway during this critical period, little is known about how RAS proteins instruct this transition. Here, we used RAS-deficient ESC for a careful dissection of the FGF pathway and demonstrated that the transcriptional factor ERF is the sole MAPK-dependent switch that controls the naïve-to-primed transition.

FA-RAS[KO] ESC are characterized by an active naïve transcriptional network, high expression of OTX2 and elevated expression of primed markers. This suggests that 1) activation of the primed transcriptional program does not require full exit from naïve pluripotency, and 2) FGF/MAPK is the main pathway controlling the exit from naïve pluripotency. Indeed, our data revealed that ERF imposes an exquisite level of coordination during the transition from naïve to primed pluripotency. In the absence of MAPK signaling, chromatin bound ERF ensures elevated expression of naïve pluripotent factors, including NANOG (Fig. 4). High levels of NANOG maintain the undifferentiated state of ESC in the absence of LIF and might shield naïve embryonic cells from premature commitment (*24, 34*). Accordingly, ERF[KO] ESC showed decreased levels of NANOG and undergo a partial exit from naïve pluripotency under 2iL conditions (Fig. 4). Thus, our data suggest that ERF supports naïve pluripotency, at least partially, by sustaining elevated levels of NANOG. ERF also controls negatively the expression of LIN28 proteins as 2iL-ERF[KO] ESC showed increased levels of both LIN28A and B. Although LIN28 proteins regulate the expression of DNMT3 proteins by sequestering let-7 microRNAs, we did not detect upregulation of the methyltransferases in naïve ERF[KO] ESC (Fig. 7). This is likely due to high levels of PRDM14, which represses expression of DNMT3 methyltransferases by recruiting Polycomb repressive complex 2 to their promoters, and it is regulated by ERF (Fig. 5E) (*35*). Thus, coordinated PRDM14 downregulation and increased LIN28 expression through ERF-dependent mechanisms could contribute to DNMT3 upregulation and increased methylation during the transition to primed pluripotency. Although DNMT3A and DNMT3B are dispensable for a successful exit from naïve pluripotency, *de novo* DNA methylation facilitates a timely progression towards a primed state

378 (*35*). In fact, DNMT3A/B double knockout ESC exhibited a delayed exit from the naïve state and

379 defective expression of the post-implantation markers OCT6, FGF5 or OTX2 (*36*). This is ultimately

380 due to persistent NANOG expression after 2iL withdrawal as DNMT3A/B deficiency leads to

381 reduced CpG methylation at the proximal NANOG promoter (*36, 37*). Similarly, FA-RAS[KO] ESC

382 showed negligible expression of DNMT3B and, therefore, high levels of NANOG. Interestingly,

383 ERF[KO] ESC showed increased DNA methylation compared to RAS[lox/lox] ESC and is particularly

384 evident in genes and CpG islands after the transition to EpiLSC (Fig. 7E). Naïve ERF[KO] ESC have low

385 levels of DNMT3 expression but elevated levels of DNMT3L, the catalytically inactive regulatory

386 factor of *de novo* DNA methyltransferases (fig. S7C), which can contribute to this bias during

387 differentiation (*38*). Furthermore, DPPA3, also known as Stella and implicated in preventing

388 excessive DNA methylation by sequestering the E3 ubiquitin ligase UHRF1, is also decreased in

389 ERF[KO] ESC (Fig. S7C) (*39, 40*). Importantly, downregulation of DPPA3 mediated by DNMT3

390 methylation has been shown to be a key event in the naïve to primed conversion (*41*).

391 Collectively, elevated levels of DNMT3L and low levels of DPPA3 in naïve conditions could

392 contribute to the aberrant level of methylation detected in FA-ERF[KO]. Nevertheless, these

393 changes in methylation do not result in major transcriptional changes in FA-ERF[KO] compared to

394 their respective control cells.

395 FA-RAS[KO] ESC showed high expression of OTX2, which has been shown a critical transcription

396 factor for thew maintenance of intermediate states of pluripotency (*7*). The consequence of

397 having high levels of OTX2 in this context can be inferred by previous studies (*25*). Indeed, OTX2

398 over-expression has been shown to redirect OCT4 to previously inaccessible sites in EpiLSC,

399 decrease expression of naïve markers (TBX3 or ESRRB), and induce EpiLSC genes, such as FGF5

400 (*25*). Interestingly, these effects take place in the presence of MEKi, a situation that is mirrored

401 in FA-RAS[KO] ESC. The ability of OTX2 to engage new enhancer sequences not only depends on its

402 levels but also on the cooperative help of additional transcription factors (*25*). This cooperation

403 might explain the enrichment of OTX2 in active naïve enhancers in FA-RAS[KO] ESC. Moreover, our

404 data also showed that, in addition to OCT4, NANOG is mobilized to sites that become active in

405 EpiLSCs, which results in partial or full transcriptional activation (Fig. 6E). In this context, NANOG

406 relocation might facilitate OTX2 positioning as it has been shown to promote chromatin

14

407   accessibility and TF recruitment together with BRG1, part of the large remodeling complex

408   SWI/SNF (*24*).

409   Finally, we showed that FA-RAS^{KO} ESC can be maintained in culture under primed conditions and

410   could be a suitable genetic model to study intermediate states of pluripotency. Indeed, FA-RAS^{KO}

411   ESC mirrored transcriptionally cells that have departed 12-24 hours from naïve pluripotency and

412   are reminiscent of the recently described pluripotent intermediate states (*6-8*). Future studies

413   will be necessary to determine the similarity of all these intermediate states with corresponding

414   *in vivo* counterparts. Nevertheless, pluripotency can be considered as a dynamic property

415   associated to different transient stem cell states that can be recapitulated by using different

416   inhibitors and/or growth factors. By using RAS^{KO} ESC we focused specifically on the FGF pathway

417   and the role of RAS proteins during the naïve to primed transition.

418   Future studies will be needed to determine if ERF plays a role in other cell fate decisions during

419   early embryonic development such as in primitive endoderm specification, where FGF4-

420   dependent activation of the MAPK pathway is also necessary. Moreover, the control of ERF over

421   PRDM14 expression suggested that ERF might also play an important role in germline fate (*42*).

422   In conclusion, here we demonstrated the essential role of ERF as regulator of the timely transition

423   to primed pluripotency.

424

425

426

427

428

429

430

431

432

433

434

435

15

MATERIALS AND METHODS

Embryo Culture

All the animal work included here was performed in C57BL/6J mice obtained from the Jackson Laboratory in compliance with the NIH Animal Care & Use Committee (ACUC) Guideline for Breeding and Weaning. 4-weeks old female mice were injected intraperitoneally with 5IU Pregnant Mare Serum Gonadotropin (PMSG, Prospec) followed by 5 IU human Chorionic Gonadotropin (hCG, Sigma-Aldrich) 46-48 hours later. Alternatively, 8-weeks naturally pregnant females were euthanized, and embryos collected in M2 media (MR-015-D, Sigma-Aldrich) at indicated time points after hCG injection: E2.75, E3.5, E4.0, E4.75. The sex of embryos was not determined. Embryos were fixed in 4% Paraformaldehyde (Electron Microscopy Sciences) for 10 min, permeabilized for 30 min in 0.3% Triton X-100 and 0.1M Glycine in PBS 1X and blocked for 1 hour (1% BSA, 0.1% Tween in PBS 1X). Embryos were incubated overnight with primary antibodies (see STAR methods), washed in 0.1% Tween in PBS 1X and incubated with the secondary antibody accordingly for 1 hour at room temperature. Embryos were imaged using a Nikon Ti2-E microscope (Nikon Instruments) equipped with a CSU-W1 spinning disk (Yokogawa), Photometrics Prime BSI sCMOS (Photometrics), and 60x Nikon Apochromat TIRF objective (NA = 1.49). Z-stacks were acquired with a x-y pixel size of 0.11 mm and z-step of 0.9 mm. For quantification, embryo z-stack images were quantified using Imaris Bitplane (Oxford Instruments). 3D surfaces were rendered based on nuclear DAPI-staining and the corresponding regions were used to quantify the fluorescence intensity of ERF, NANOG, and KLF4. For experiment shown in Fig. 1A shows one representative experiment with the following embryos and cells used for quantification: Exp 1: E2.75: 3 Embryos, 24 cells in total. E3.5: 3 Embryos, 117 cells total, E4.0: 4 Embryos, 168 cells total, E4.75: 1 Embryo: 75 cells. Exp 2: E3.5: 8 embryos, E3.75: 4 embryos, E4.0: 6 embryos.

Cell culture and differentiation

N-$Ras^{-/-}$; H-$Ras^{-/-}$; K-$Ras^{f/f}$; *Ubiq-Cre^{ERT2}* (RAS$^{lox/lox}$) (*15*) ESC were grown in N2B27 media supplemented with 2i/LIF (1 mM PD0325901, 3 mM CHIR99021, both from Tocris and 1:500 LIF,

465   made in house). N2B27 media consisted of a 1:1 mix of DMEM/F12 and Neurobasal Medium, 1X

466   N2 supplement, 1X B27 supplement, 0.1 mM nonessential amino acids, 55 μM β-

467   mercaptoethanol and 1% penicillin/streptomycin (all from Life Technologies). Cells were

468   routinely cultured in 0.1% gelatinized plates and passaged with Accutase (Gibco) unless

469   otherwise indicated. To induce EpiLSC differentiation, cells were grown for few passages in plates

470   pretreated with 10 μg/ml polyL-ornithine and 5 μg/ml laminin (Corning). A total of 200,000–

471   300,000 cells per 10 cm^2 were plated on plates pretreated with 5 μg/ml Fibronectin (Millipore)

472   in N2B27 media supplemented with 1% KOSR, 12 ng/ml FGF2 (R&D systems) and 20 ng/ml Activin

473   A (PeproTech) for 48 hours including daily media changes. For inducing CRE-mediated deletion

474   of K-RAS allele, we incubated $RAS^{lox/lox}$ ESC with 1μM 4-hydroxytamoxifen (4-OHT, Sigma-Aldrich)

475   for 6 days before performing any experiment. To maintain EpiLSC in culture, ESCs were plated in

476   N2B27 media with 12 ng/ml FGF2, 20 ng/ml Activin A and 1 mM XAV939 on Fibronectin-coated

477   plates at a density of 10,000 cells per cm^2. EpiLSC were passaged the first time with Accutase

478   including 1 mM Y27632 to enhance plating efficiency. Media was changed every other day and

479   passaged every 2-3 days. HEK293T (American Type Culture Collection) cells were grown in

480   DMEM, 10% FBS, and 1% penicillin/streptomycin.

481 To target a short half-life form of eGFP (deGFP) in the endogenous REX1 gene, we generated a

482   targeting vector by inserting deGFP in pCR®-Blunt II TOPO® (Zero Blunt TOPO PCR cloning kit,

483   Invitrogen). Homology arms PCR-amplified from endogenous sequences upstream of the start

484   codon and downstream of the stop codon from REX1 were cloned to flank pCR®-Blunt II TOPO®-

485   deGFP. In addition, specific small guide RNA (sgRNA) sequences targeting the surroundings of the

486   REX1 start codon were cloned into the plasmid pX330-U6-Chimeric_BB-CBh-hSpCas9 (Addgene,

487   42230) (*44*). The sequences of the sgRNAs were designed with the Genetic Perturbation Platform

488   sgRNA designer tool (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design).

489   Both plasmids were transfected in ESC using Jetprime (Polyplus transfection) to induce the

490   targeting and cell sorted based on GFP intensity to isolate individual clonal ESC lines. See Table

491   S1 for primer information.

492

493

17

Clonogenicity Assay

ESC were withdrawn of 2i/LIF for 48 hours and plated at single cell density (50 cells/cm^2) in N2B27 media with 2i/LIF on plates coated with 0.1% gelatin (Sigma). At day 5, alkaline phosphatase staining was performed using the Alkaline Phosphatase Detection Kit (Millipore). Colonies were counted manually. At least three independent experiments with three replicates per experiment were performed.

Generation of self-organizing embryonic spheres

ESC growing on gelatinized plates in N2B27 media with 2i/LIF were dissociated with Accutase and washed with PBS before their resuspension in growth factor reduced Matrigel (Corning) at a concentration of 10,000 cells/20µl of matrigel. The suspension was deposited in drops in 8-µwell Ibidi microplates and incubated at 37°C until the Matrigel solidified. Wells were then filled with N2B27 media without 2i/LIF and cultured for 48-72 hours at 37°C and 5% CO2. ESC-derived spheres were fixed in 4% PBS–paraformaldehyde (PFA) for 10 min at room temperature. Permeabilization was performed in PBS containing 0.3% Triton X-100 (Sigma) and 0.1 M glycine (Sigma) for 30 minutes at room temperature. Spheres were incubated with primary antibodies at 4°C overnight, followed by incubation with corresponding fluorescently conjugated Alexa Fluor secondary antibodies for 2 hours at room temperature. Both primary and secondary antibodies were diluted in PBS containing 1% BSA (Sigma) and 0.1% Tween20 (Sigma). See Table S2 for antibody information.

Western blot

Cells were lysed in 50 mM Tris pH 8, 8 M Urea (Sigma) and 1% Chaps (Millipore) followed by 30 min of shaking at 4°C. 20 µg of supernatants were run on 4%-12% NuPage Bis-Tris Gel (Invitrogen) and transferred onto Nitrocellulose Blotting Membrane (GE Healthcare). Membranes were blocked in 5% skim milk (Millipore) and 0.1% Tween 20 (Sigma) in PBS. Membranes were incubated with the primary antibody overnight at 4°C (see STAR methods), followed by incubation with HRP-conjugated secondary antibodies (1:5000) for 1 h at room

18

522 temperature. Membranes were developed using SuperSignal West Pico PLUS (Thermo Scientific).

523 See Table S2 for antibody information.

524

525 **Dot blot**

526 Trypsinized ESC were lysed for 2-3 hours at 55°C with lysis buffer (100 mM Tris-HCl pH 8, 5 mM

527 EDTA, 0.2% SDS, 20 mM NaCl and 100 μg/ml Proteinase K). DNA was isolated by adding an equal

528 volume of Phenol/Chloroform/Isoamyl to the samples and using phase-lock tubes (5PRIME Phase

529 Lock Gel™ Heavy, Quantabio), followed by an extraction with identical volume of Chloroform.

530 DNA was precipitated with 2 volumes of 100% Ethanol plus 0.3 M NaAc, washed with 70% Ethanol

531 and resuspended in water. A total of 500 ng of DNA was diluted in 0.3 M NaOH and denatured at

532 42°C for 12 minutes. After incubation, the samples were rapidly transferred by spotting each

533 sample into a nitrocellulose membrane (Nitrocellulose Blotting Membrane, GE Healthcare). After

534 the transfer, DNA was crosslinked with a Stratalinker® UV crosslinker (Stratagene) using the

535 Autocrosslink setting. The membrane was blocked in 5% skim milk (Millipore) and 0.1% Tween

536 20 (Sigma) in PBS, incubated with 1:500 dilution of the anti-5mC antibody (see STAR methods)

537 overnight at 4°C, followed by incubation with HRP-conjugated secondary antibody (1:5000) for 1

538 h at room temperature. Membrane was developed using SuperSignal West Pico PLUS (Thermo

539 Scientific). See Table S2 for antibody information.

540

541 **Flow cytometry**

542 For flow cytometry experiments, cells were dissociated into single cell suspensions and analyzed

543 for GFP gene expression using a FACS Fortessa (BD Biosciences). DAPI was added to detect cells

544 with compromised membrane integrity. Data was analyzed using FlowJo. At least two

545 independent experiments were performed.

546

547 **Immunofluorescence**

548 Cells were fixed in 4 % Paraformaldehyde (Electron Microscopy Sciences) for 10 min at room

549 temperature, permeabilized in 100 mM Tris-HCl pH 7.4, 50 mM EDTA pH 8.0, 0.5% Triton X-100

550 and incubated with the corresponding primary antibodies overnight (see STAR methods). This

19

551 was followed by incubation with corresponding fluorescently conjugated Alexa Fluor secondary

552 antibodies for 2 h at room temperature. Both primary and secondary antibodies were diluted in

553 PBS containing 1% BSA (Sigma) and 0.1% Tween20 (Sigma). Images were acquired using either a

554 Nikon spinning disk confocal microscope (CSU-W1) or a Zeiss LSM880 Airyscan microscope. See

555 Table S2 for antibody information.

556

557 **RNAseq and data analysis**

558 RNA was isolated using the ISOLATE II RNA Mini Kit (Bioline) following manufacturer's

559 recommendations. DNA libraries for RNAseq analysis were prepared using NEBNext Ultra II

560 Directional RNA Library Prep Kit for Illumina (New England Biolabs, NEB) and NEBNext rRNA

561 Depletion Kit (Human/Mouse/Rat) (NEB) according to the manufacturer's protocol. Sequencing

562 was performed on the Illumina NextSeq550 (75bp pair-end reads). RNAseq reads were adapter

563 trimmed using fastp v.0.20.0 (*45*). Transcript expression was quantified via mapping to mouse

564 gencode v25 transcripts using salmon (*46*). Identification of differentially expressed genes

565 between samples was performed using DESEQ2 (PMID: 25516281). RNAseq coverage tracks

566 were generated by aligning RNAseq reads to UCSC version mm10 of the mouse genome using the

567 STAR v2.6.1a aligner (*47*) followed by application of the 'bamCoverage' utility from deeptools

568 (*48*) to generate signal track files with the following parameters: normalization=RPKM,

569 bin_size=50, smooth_length=1. For comparison to other published RNAseq data sets, gene

570 counts across samples were quantile-normalized using the limma package (*49*). Batch correction

571 was then performed on quantile-normalized counts using COMBAT (*50*). See Table S3 for further

572 software and algorithm information.

573

574 **CUT&RUN protocol**

575 The CUT&RUN protocol was slightly modified from (*51, 52*). In brief, cells were washed thrice

576 with Wash Buffer (20 mM HEPES-KOH pH 7.5, 150 mM NaCl, 0.5 mM spermidine, Roche complete

577 Protease Inhibitor tablet EDTA free) and bound to activated Concanavalin A beads (Polysciences)

578 for 10 minutes at room temperature. Cells were then permeabilized in Digitonin Buffer (0.05 %

579 Digitonin and 0.1% BSA in Wash Buffer) and incubated with corresponding antibodies at 4°C for

20

580   2 hours. For negative controls, Guinea Pig anti-Rabbit IgG (Antibodies-online) was used. Following

581   antibody incubation, cells were washed with Digitonin Buffer and incubated with a hybrid protein

582   A-protein G-Micrococcal nuclease (pAG-MNase) at 4°C for 1 hour. Samples were washed in

583   Digitonin Buffer, resuspended in 150 µl Digitonin Buffer and equilibrated to 0°C on ice water for

584   5 minutes. To initiate MNase cleavage, 3 µl 100 mM $CaCl_2$ was added to cells and after 1 hour of

585   digestion, reactions were stopped with the addition of 150 µl 2x Stop Buffer (340 mM NaCl, 20

586   mM EDTA, 4 mM EGTA, 0.02 % Digitonin, 50 µg/ml RNase A, 50 µg/ml Glycogen). Samples were

587   incubated at 37°C for 10 minutes to release DNA fragments and centrifuged at 16,000 g for

588   5 minutes. Supernatants were collected and a mix of 1.5 µl 20% SDS / 2.25 µl 20 mg/ml Proteinase

589   K was added to each sample and incubated at 65°C for 35 minutes. DNA was precipitated with

590   ethanol and sodium acetate and pelleted by high-speed centrifugation at 4°C, washed, air-dried

591   and resuspended in 10 µ 0.1x TE. See Table S2 for antibody information.

592

593   **Library preparation and sequencing**

594   The entire precipitated DNA obtained from CUT&RUN was used to prepare Illumina compatible

595   sequencing libraries. In brief, end-repair was performed in 50 µl of T4 ligase reaction buffer,

596   0.4 mM dNTPs, 3 U of T4 DNA polymerase (NEB), 9 U of T4 Polynucleotide Kinase (NEB) and 1 U

597   of Klenow fragment (NEB) at 20°C for 30 minutes. End-repair reaction was cleaned using AMPure

598   XP beads (Beckman Coulter) and eluted in 16.5 µl of Elution Buffer (10 mM Tris-HCl pH 8.5)

599   followed by A-tailing reaction in 20 µl of dA-Tailing reaction buffer (NEB) with 2.5 U of Klenow

600   fragment exo- (NEB) at 37°C for 30 minutes. The 20 µl of the A-tailing reaction were mixed with

601   Quick Ligase buffer 2X (NEB), 3000 U of Quick Ligase (NEB) and 10 nM of annealed adaptor

602   (Illumina truncated adaptor) in a volume of 50 µl and incubated at room temperature for 20 min.

603   The adaptor was prepared by annealing the following HPLC-purified oligos: 5'-

604   Phos/GATCGGAAGAGCACACGTCT-3' and 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATC*T-3'

605   (*phosphorothioate bond). Ligation was stopped by adding 50 mM of EDTA, cleaned with AMPure

606   XP beads and eluted in 14 µl of Elution Buffer. All volume was used for PCR amplification in a

607   50 µl reaction with 1 µM primers TruSeq barcoded primer p7, 5'-

608   CAAGCAGAAGACGGCATACGAGATXXXXXXXXGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T-3'

21

609 and   TruSeq   barcoded   primer   p5   5'-

610 AATGATACGGCGACCACCGAGATCTACACXXXXXXXXXACACTCTTTCCCTACACGACGCTCTTCCGATC*T-

611 3' (* represents a phosphothiorate bond and XXXXXXXX a barcode index sequence), and 2X Kapa

612 HiFi HotStart Ready mix (Kapa Biosciences). The temperature settings during the PCR

613 amplification were 45 s at 98°C followed by 15 cycles of 15 s at 98°C, 30 s at 63°C, 30 s at 72°C

614 and a final 5 min extension at 72°C. PCR reactions were cleaned with AMPure XP beads (Beckman

615 Coulter), run on a 2% agarose gel and a band of 300bp approximately was cut and gel purified

616 using QIAquick Gel Extraction Kit (QIAGEN). Library concentration was determined with KAPA

617 Library Quantification Kit for Illumina Platforms (Kapa Biosystems). Sequencing was performed

618 on the Illumina NextSeq550 (75bp pair-end reads).

619

620 **Cut&Run data processing**

621 Data were processed using a modified version of Cut&RunTools (*53*). Reads were adapter

622 trimmed using fastp v.0.20.0 (*45*). An additional trimming step was performed to remove up to

623 6bp adapter from each read. Next, reads were aligned to the mm10 genome using bowtie2 (*54*)

624 with the 'dovetail' and 'sensitive' settings enabled. macs2 (*55*) was used to call peaks with q-

625 value cutoff < 0.01. Normalized (RPKM) signal tracks were generated using the 'bamCoverage'

626 utility from deepTools with parameters bin-size=25, smooth length=75, and 'center_reads' and

627 'extend_reads' options enabled (*48*). See Table S3 for further software and algorithm

628 information.

629

630 **Processing for published ChIP datasets**

631 External next generation sequencing data were downloaded from the Sequence Read Archive

632 (SRA) and analyzed as follows. These analyses include a re-analysis of our original datasets (*9*).

633 Reads were aligned to the mm10 genome using bowtie2 (*54*). Duplicate reads were removed

634 using MarkDuplicates from the Picard toolkit ("Picard Toolkit." 2019. Broad Institute, GitHub

635 Repository. http://broadinstitute.github.io/picard/). Normalized (RPKM) signal tracks were

636 generated bamCoverage utility from deepTools (*48*), using the parameters bin-size=25, smooth

637 length=75, 'center_reads' and 'extend_reads'. For paired-end data, read mates were extended

22

638   to the fragment size defined by the two read mates. For single-end ChIP-seq data, reads were

639   extended to the estimated fragment length estimated by phantompeakqualtools (*56*). See Table

640   S3 for further software and algorithm information.

641

642   **Identification of OCT4 binding sites in ESC and EpiLSC**

643   Fastq files from published OCT4 ChIP-seq data for ESC (SRR1202455, SRR1202456), EpiLSC plus

644   and minus Activin-A (SRR1202468, SRR1202469), and associated input controls (SRR1202465,

645   SRR1202464, SRR1202477 SRR1202478) (*25*) were downloaded from the Sequence Read Archive

646   (SRA). Single-end reads were aligned to the mm10 genome using bwa (*57*). Duplicate reads were

647   removed using MarkDuplicates from the Picard toolkit ("Picard Toolkit." 2019. Broad Institute,

648   GitHub Repository. http://broadinstitute.github.io/picard/), and peaks for each sample were

649   called using macs2 (*55*) with q-value cutoff < 0.01 and extension length determined using

650   phantompeakqualtools (*56*). Diffbind (*58*) using the DeSeq2 method was used to determine

651   differentially bound peaks, treating EpiLSC plus and minus Activin-A samples as replicate

652   experiments as was done in the original study (*25*). Peaks were determined to be ESC- or EpiLC-

653   specific if they differed by 2-fold read concentration with p-val<0.01 and FDR <0.03. A subset of

654   peaks with > mean read concentration for EpiLC and ESC with <0.5 fold difference were selected

655   as "common" or shared peaks. See Table S3 for further software and algorithm information.

656

657   **RRBS data processing**

658   DNA libraries for RRBS analysis were prepared using the Premium RRBS kit (Diagenode) following

659   manufacturer's recommendations. Sequencing was performed on the Illumina NextSeq550

660   (75bp single-end reads). Single-end RRBS reads were adapter and quality trimmed (phred33

661   score>=20) using trimgalore v0.6.5

662   (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore) with the RRBS option

663   invoked. Bismark v0.22.1 (*59*) was used to align reads to UCSC version mm10 of the mouse

664   genome. CpG methylation was extracted using bismark ignoring the first 4 bases of the read after

665   inspection of m-bias plots. Differentially methylated regions were called using the methylKit

666   package (*60*). See Table S3 for further software and algorithm information.

667

668 **QUANTIFICATION AND STATISTICAL ANALYSIS.**

669 All information regarding statistical details of the experiments, number of experiments, statistical

670 tests used, number of cells/experiments can be found in the corresponding figure legends o

671 methods section.

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697   **REFERENCES**

698

699   1.   J. Wray, T. Kalkan, A.G. Smith, The ground state of pluripotency. *Biochem. Soc. Trans.* **38**,
700   1027-32 (2010).

701   2.   J. Nichols, A. Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487-492 (2009).

702   3.   L. Weinberger, M. Ayyash, N. Novershtern, J.H. Hanna, Dynamic stem cell states: naive to
703   primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* **17**, 155-69 (2016).

704   4.   Q.L. Ying, J. Wray, J. Nichols, L. Batlle-Morera, B. Doble, J. Woodgett, P. Cohen, A. Smith, The
705   ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-23 (2008).

706   5.   P.J. Tesar, J.G. Chenoweth, F.A. Brook, T.J. Davies, E.P. Evans, D.L. Mack, R.L., Gardner, R.D.
707   McKay, New cell lines from mouse epiblast share defining features with human embryonic
708   stem cells. *Nature* **448**, 196-199 (2007).

709   6.   A. Neagu, E. van Genderen, I. Escudero, L. Verwegen, D. Kurek, J. Lehmann, J. Stel, R.A.M.
710   Dirks, G. van Mierlo, A. Maas, C. Eleveld, Y. Ge, A.T. den Dekker, R.W.W. Brouwer, W.F.J. van
711   IJcken, M. Modic, M., Drukker, J.H. Jansen, N.C. Rivron, E.B. Baart, H. Marks, D. Ten Berge,
712   In vitro capture and characterization of embryonic rosette-stage pluripotency between naive
713   and primed states. *Nat. Cell Biol.* **22**, 534-545 (2020).

714   7.   M. Kinoshita, M., Barber, W. Mansfield, Y. Cui, D. Spindlow, G.G. Stirparo, S. Dietmann, J.
715   Nichols, A. Smith, Capture of Mouse and Human Stem Cells with Features of Formative
716   Pluripotency. *Cell Stem Cell* **28***: DOI:https://doi.org/10.1016/j.stem.2020.11.005, (2020).

717   8.   L. Yu, Y. Wei, H. Sun, A.K. Mahdi, C.P. Arteaga, M. Sakurai, D.A. Schmitz, C. Zheng, E.D.
718   Ballard, J. Li, N. Tanaka, A. Kohara, D. Okamura, A.A. Mutto, Y. Gu, P.J. Ross, J. Wu, Derivation
719   of Intermediate Pluripotent Stem Cells Amenable to Primordial Germ Cell Specification. *Cell*
720   *Stem Cell* **28**: DOI:https://doi.org/10.1016/j.stem.2020.11.003 (2020).

721   9.   C. Mayor-Ruiz, T. Olbrich, M. Drosten, E. Lecona, M. Vega-Sendino, S. Ortega, O. Dominguez,
722   M. Barbacid, S. Ruiz O. Fernández-Capetillo, ERF deletion rescues RAS deficiency in mouse
723   embryonic stem cells. *Genes&Dev* **32**, 568-576 (2018).

724   10.   D.N. Sgouras, M.A. Athanasiou, G.J. Jr. Beal, R.J. Fisher, D.G. Blair, G.J. Mavrothalassitis, ERF:
725   an ETS domain protein with strong transcriptional repressor activity, can suppress ets-

726   associated tumorigenesis and is regulated by phosphorylation during cell cycle and
727   mitogenic stimulation. *EMBO J.* **14**, 4781–4793 (1995).

728 11. L. Le Gallic, D. Sgouras, G. Jr. Beal, G. Mavrothalassitis, Transcriptional repressor ERF is a
729   Ras/mitogen-activated protein kinase target that regulates cellular proliferation. *Mol. Cell.*
730   *Biol.* **19**, 4121-4133 (1999).

731 12. L. Le Gallic, L., Virgilio, P. Cohen, B. Biteau, G. Mavrothalassitis, ERF nuclear shuttling, a
732   continuous monitor of Erk activity that links it to cell cycle progression. *Mol. Cell. Biol.* **24**,
733   1206-18 (2004).

734 13. E. Vorgia, A. Zaragkoulias, I. Peraki, G. Mavrothalassitis, Suppression of Fgf2 by ETS2
735   repressor factor (ERF) is required for chorionic trophoblast differentiation. *Mol. Reprod. Dev.*
736   **84**, 286-295 (2017).

737 14. S. Nowotschin, M. Setty, Y.Y. Kuo, V. Liu, V., Garg, R. Sharma, C.S. Simon, N. Saiz, R. Gardner,
738   S.C. Boutet, D.M. Church, P.A. Hoodless, A.K. Hadjantonakis, D. Pe'er, The emergent
739   landscape of the mouse gut endoderm at single-cell resolution. *Nature* **569**, 361-367 (2019).

740 15. M. Drosten, A. Dhawahir, E.Y. Sum, J. Urosevic, C.G. Lechuga, L.M. Esteban, E. Castellano, C.
741   Guerra, E. Santos, M. Barbacid, Genetic analysis of Ras signalling pathways in cell
742   proliferation, migration and survival. *EMBO J.* **29**, 1091-1104 (2010).

743 16. K. Hayashi, H. Ohta, K. Kurimoto, S. Aramaki, M. Saitou Reconstitution of the mouse germ
744   cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519-532 (2011).

745 17. F. Nakaki, K. Hayashi, H. Ohta, K. Kurimoto, Y. Yabuta, M. Saitou, Induction of mouse germ-
746   cell fate by transcription factors in vitro. *Nature* **501**, 222-226. (2013).

747 18. P. Yang, S.J. Humphrey, S. Cinghu, R. Pathania, A.J. Oldfield, D. Kumar, D. Perera, J.Y.H. Yang,
748   D.E. James, M. Mann, R. Jothi, Multi-omic Profiling Reveals Dynamics of the Phased
749   Progression of Pluripotency. *Cell Syst.* **8**, 427-445 (2019).

750 19. I. Bedzhov, M. Zernicka-Goetz, Self-organizing properties of mouse pluripotent cells initiate
751   morphogenesis upon implantation. *Cell* **156**, 1032-1044 (2014).

752 20. M.N. Shahbazi, A. Scialdone, N. Skorupska, A. Weberling, G. Recher, M. Zhu, A. Jedrusik, L.G.
753   Devito, L. Noli, I.C. Macaulay, C. Buecker, Y. Khalaf, D. Ilic, T. Voet, J.C. Marioni, M. Zernicka-

754   Goetz Pluripotent state transitions coordinate morphogenesis in mouse and human
755   embryos. *Nature* **552**: 239-243 (2017).

756 21. T. Kalkan, N. Olova, M. Roode, C. Mulas, H.J. Lee, I. Nett, H. Marks, R. Walker, H.G.
757   Stunnenberg, K.S. Lilley, J. Nichols, W. Reik, P. Bertone, A. Smith, Tracking the embryonic
758   stem cell transition from ground state pluripotency. *Development* **144**, 1221-1234 (2017).

759 22. D. Hnisz, B.J. Abraham, T.I. Lee, A. Lau, V. Saint-André, A.A. Sigova, H.A. Hoke, R.A. Young,
760   Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934-947 (2013).

761 23. W.A. Whyte, D.A. Orlando, D. Hnisz, B.J. Abraham, C.Y. Lin, M.H. Kagey, P.B. Rahl, T.I. Lee,
762   R.A. Young, Master transcription factors and mediator establish super-enhancers at key cell
763   identity genes. *Cell* **153**, 307-319 (2013).

764 24. V. Heurtier, N. Owens, I. Gonzalez, F. Mueller, C. Proux, D. Mornico, P. Clerc, A. Dubois, P.
765   Navarro, The molecular logic of Nanog-induced self-renewal in mouse embryonic stem cells.
766   *Nat. Commun.* **10**, 1109 (2019).

767 25. C. Buecker, R. Srinivasan, Z. Wu, E. Calo, D. Acampora, T. Faial, A. Simeone, M. Tan, T. Swigut,
768   J. Wysocka, Reorganization of enhancer patterns in transition from naive to primed
769   pluripotency. *Cell Stem Cell* **14**, 838-53 (2014).

770 26. T. Zhang, Z. Zhang, Q. Dong, J. Xiong, B. Zhu, Histone H3K27 acetylation is dispensable for
771   enhancer activity in mouse embryonic stem cells. *Genome Biol.* **21**, 45 (2020).

772 27. J. Zhang, S.Ratanasirintrawoot, S. Chandrasekaran, Z. Wu, S.B. Ficarro, C. Yu, C.A. Ross, D.
773   Cacchiarelli, Q. Xia, M. Seligson, G. Shinoda, W. Xie, P. Cahan, L. Wang, S.C. Ng. S. Tintara, C.
774   Trapnell, C. Onder, Y.H. Loh, T. Mikkelsen, P. Sliz, M.A. Teitell, J.M. Asara, J.A. Marto, H. Li,
775   J.J. Collins, G.Q. Daley, LIN28 Regulates Stem Cell Metabolism and Conversion to Primed
776   Pluripotency. *Cell Stem Cell* **19**, 66-80 (2016).

777 28. J. Balzeau, M.R. Menezes, S. Cao, J.P. Hagan, The LIN28/let-7 Pathway in Cancer. *Front.*
778   *Genet.* **8**, 31 (2017).

779 29. G. Auclair, S. Guibert, A. Bender, M. Weber, Ontogeny of CpG island methylation and
780   specificity of DNMT3 methyltransferases during embryonic development in the mouse.
781   *Genome Biol*. **15**, 545 (2014).

782 30. S. Takahashi, S., Kobayashi, I. Hiratani Epigenetic differences between naïve and primed
783   pluripotent stem cells. *Cell Mol. Life Sci.* **75**, 1191-1203 (2018)

784 31. A. Meissner, A., Gnirke, G.W. Bell, B. Ramsahoye, E.S. Lander, R. Jaenisch, Reduced
785   representation bisulfite sequencing for comparative high-resolution DNA methylation
786   analysis. *Nucleic Acids Res.* **33**, 5868-5877 (2005).

787 32. M. Kang, A. Piliszek, J. Artus, A.K. Hadjantonakis, FGF4 is required for lineage restriction and
788   salt-and-pepper distribution of primitive endoderm factors but not their initial expression in
789   the mouse. *Development* **140**: 267-279 (2013).

790 33. Y. Yamanaka, F. Lanner, J. Rossant, FGF signal-dependent segregation of primitive endoderm
791   and epiblast in the mouse blastocyst. *Development* **137**: 715-724 (2010).

792 34. K. Mitsui, Y. Tokuzawa, H. Itoh, K. Segawa, M. Murakami, K. Takahashi, M. Maruyama, M.,
793   Maeda, S. Yamanaka, The homeoprotein Nanog is required for maintenance of pluripotency
794   in mouse epiblast and ES cells. *Cell* **113**, 631 -642 (2003).

795 35. M. Yamaji, J. Ueda, K. Hayashi, H. Ohta, Y. Yabuta, K. Kurimoto, R. Nakato, Y. Yamada, K.
796   Shirahige, M. Saitou, PRDM14 ensures naive pluripotency through dual regulation of
797   signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* **12**, 368-382
798   (2013).

799 36. M.A. Li, P.P. Amaral, P. Cheung, J.H. Bergmann, M. Kinoshita, T. Kalkan, M. Ralser, S. Robson,
800   F. von Meyenn, M. Paramor, F. Yang, C. Chen, J. Nichols, D.L. Spector, T. Kouzarides, L. He,
801   A. Smith, A lncRNA fine tunes the dynamics of a cell state transition involving Lin28, let-7 and
802   de novo DNA methylation. *Elife* **6**, e23468 (2017).

803 37. C.R. Farthing, G. Ficz, R.K. Ng, C.F. Chan, S. Andrews, W. Dean, M. Hemberger, W. Reik, Global
804   mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of
805   pluripotency genes. *PLoS Genet.* **4**, e1000116 (2008).

806 38. N. Veland, Y. Lu, S. Hardikar, S., Gaddis, Y. Zeng, B. Liu, M.R. Estecio, Y. Takata, K. Lin, M.W.
807   Tomida, J. Shen, D. Saha, H. Gowher, H., Zhao, T. Chen, DNMT3L facilitates DNA methylation
808   partly by maintaining DNMT3A stability in mouse embryonic stem cells. *Nucleic Acids Res.*
809   **47**, 152-167. (2019).

810   39. W. Du, Q. Dong, Z. Zhang, B. Liu, T. Zhou, R.M. Xu, H. Wang, B. Zhu, Y. Li, Stella protein
811   facilitates DNA demethylation by disrupting the chromatin association of the RING finger-
812   type E3 ubiquitin ligase UHRF1. *J Biol Chem.* **294**, 8907-8917 (2019).

813   40. T. Nakamura, Y. Arai, H. Umehara, M. Masuhara, T. Kimura, H. Taniguchi, T. Sekimoto, M.
814   Ikawa, Y. Yoneda, M. Okabe, S. Tanaka, K. Shiota, T. Nakano, PGC7/Stella protects against
815   DNA demethylation in early embryogenesis. *Nat. Cell Biol.* **9**, 64–71 (2007).

816   41. H. Sang, D. Wang, S. Zhao, J. Zhang, Y. Zhang, J. Xu, X. Chen, Y. Nie, K. Zhang, S. Zhang, Y.
817   Wang, N. Wang, F. Ma, L. Shuai, Z. Li, N. Liu, Dppa3 is critical for Lin28a-regulated ES cells
818   naïve–primed state conversion. *J. Mol. Cell Biol.* **11**, 474-488 (2019).

819   42. N. Grabole, J. Tischler, J.A. Hackett, S. Kim, F. Tang, H.G. Leitch, E. Magnúsdóttir, M.A. Surani,
820   Prdm14 promotes germline fate and naive pluripotency by repressing FGF signaling and DNA
821   methylation. *EMBO Rep.* **14**, 629-637. (2013).

822   43. C. Galonska, M.J. Ziller, R. Karnik, A. Meissner, Ground State Conditions Induce Rapid
823   Reorganization of Core Pluripotency Factor Binding before Global Epigenetic
824   Reprogramming. *Cell Stem Cell* **17**, 462-470 (2015).

825   44. L. Cong, F.A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P.D. Hsu, X. Wu, W. Jiang, L.A.
826   Marraffini, F. Zhang, Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**:
827   819-23 (2013)

828   45. S. Chen, Y. Zhou, Y., Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor.
829   *Bioinformatics* **34**: 884-i890 (2018)

830   46. R. Patro, G. Duggal, M.I. Love, R.A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware
831   quantification of transcript expression. *Nat. Methods* **14**, 417-419 (2017).

832   47. A. Dobin, C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T.R.
833   Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21. (2013).

834   48. F. Ramírez, D.P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A.S. Richter, S. Heyne, F. Dündar,
835   T. Manke, deepTools2: a next generation web server for deep-sequencing data analysis.
836   *Nucleic Acids Res* **44**: 160-165 (2016)

49. M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. **43**: e47 (2015).

50. W.E. Johnson, C. Li, A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**: 118-127 (2007).

51. P.J. Skene, S. Henikoff, An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6**, e21856 (2017).

52. M.P. Meers, T.D. Bryson, J.G. Henikoff, S. Henikoff, Improved CUT&RUN chromatin profiling tools. *Elife* **8**, e46314 (2019).

53. Q. Zhu, N. Liu, S.H. Orkin, G.C. Yuan, CUT&RUNTools: a flexible pipeline for CUT&RUN processing and footprint analysis. *Genome Biol*. **20**: 192 (2019).

54. B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**: 357-359. (2012).

55. Y. Zhang, T. Liu, C.A. Meyer, J. Eeckhoute, D.S. Johnson, B.E. Bernstein, C. Nusbaum, R.M.Myers, M. Brown, W. Li, X.S. Liu, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**: R137. (2008).

56. P.V. Kharchenko, M.Y. Tolstorukov, P.J. Park, Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat. Biotechnol.* **26**: 1351-1359 (2008).

57. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).

58. R. Stark G.D. Brown DiffBind: Differential Binding Analysis of ChIP-Seq Peak Data. Bioconductor. Available online at: http://bioconductor.org/packages/release/bioc/html/DiffBind.html. (2011).

59. F. Krueger, S.R. Andrews, Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**: 1571-1572 (2011).

60. A. Akalin, M. Kormaksson, S. Li, F.E. Garrett-Bakelman, M.E. Figueroa, A. Melnick, C.E. Mason, methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol*. **13**, R87 (2012).

865   61. D. van Dijk, R. Sharma, J. Nainys, K. Yim, P. Kathail, A.J. Carr, C. Burdziak, K.R. Moon, C.L.
866   Chaffer, D. Pattabiraman, B. Bierie, L. Mazutis, G. Wolf, S. Krishnaswamy, D. Pe'er,
867   Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **174**: 716-
868   729.e27 (2018).

869   62. J. Wu, B. Huang, H. Chen, Q. Yin, Y. Liu, Y. Xiang, B. Zhang, B., Liu, Q. Wang, W. Xia, W. Li, Y.
870   Li, J. Ma, X. Peng, H. Zheng, J. Ming, W. Zhang, J. Zhang, G. Tian, F. Xu, Z. Chang, J. Na, X.
871   Yang, W. Xie, The landscape of accessible chromatin in mammalian preimplantation
872   embryos. *Nature* **534**, 652-657 (2016).

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

908

909

910

911

912

913

914

915

916

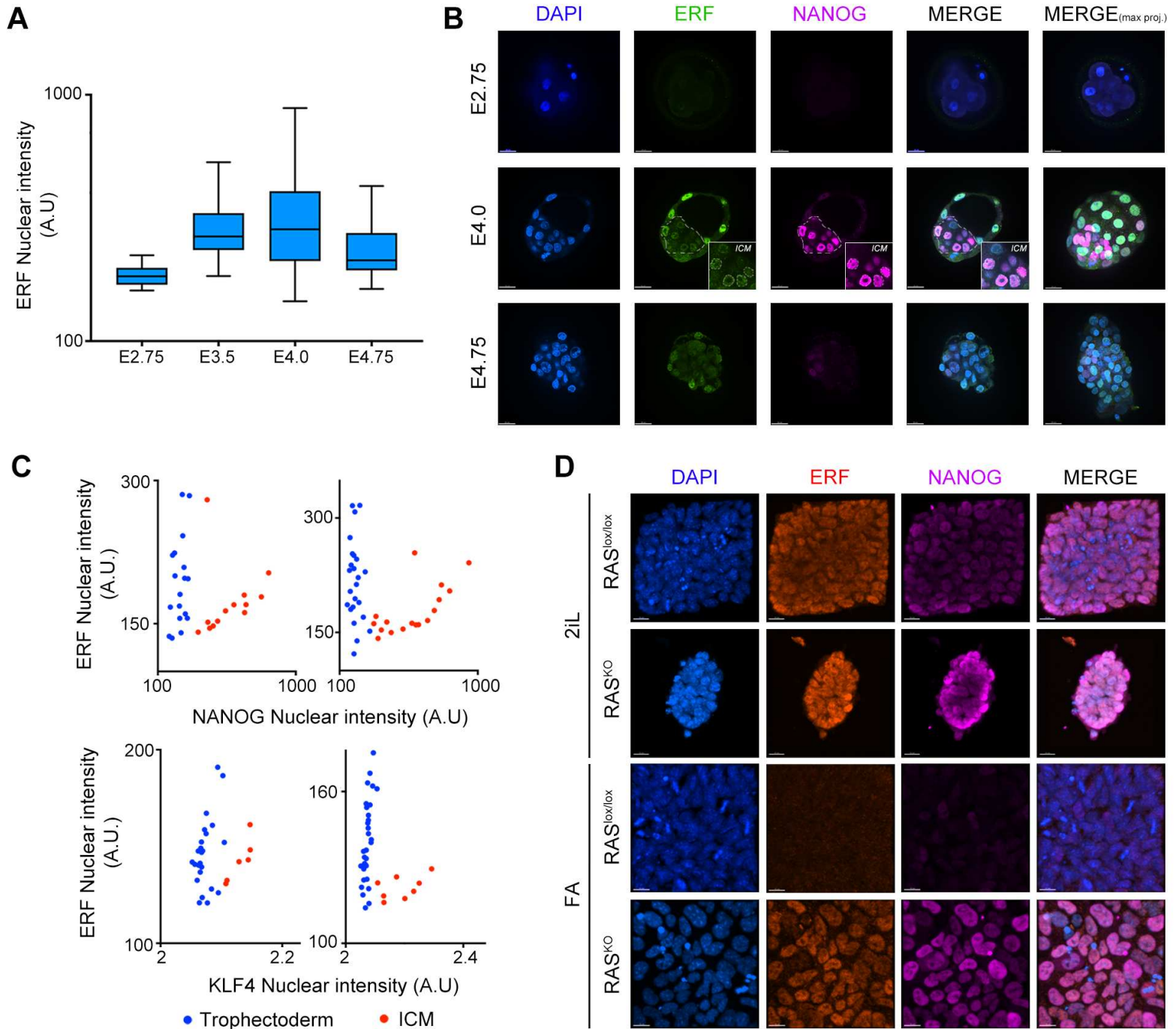917

918

919

920

921

922

Figure 1

923 **Fig. 1. ERF expression correlates with naïve pluripotency markers.** (**A**) Graph showing mean

924 nuclear fluorescent intensity for ERF in mouse embryos at different days of embryonic

925 development (E). (**B**) Immunofluorescence analysis of NANOG and ERF in mouse embryos at

926 E2.75, E4.0 and E4.75. Note that ERF is expressed in both ICM and TE. However, at E4.75 ERF and

927 NANOG are mostly downregulated in epiblast cells. Dashed line highlights the ICM. DAPI was

928 used to visualize nuclei. Scale bars, 20μm. (**C**) Graphs showing relative nuclear fluorescence

929 intensity of ERF and NANOG (upper plots) and ERF and KLF4 (lower plots). Every dot represents

930 one single nucleus, and each plot corresponds to an individual E3.5 embryo. Two representative

931 examples are shown but at least 10 embryos were analyzed. (**D**) Immunofluorescence analysis of

932 2iL and FA treated RAS$^{lox/lox}$ and RAS^{KO} ESC and stained for ERF (red) and NANOG (purple). DAPI

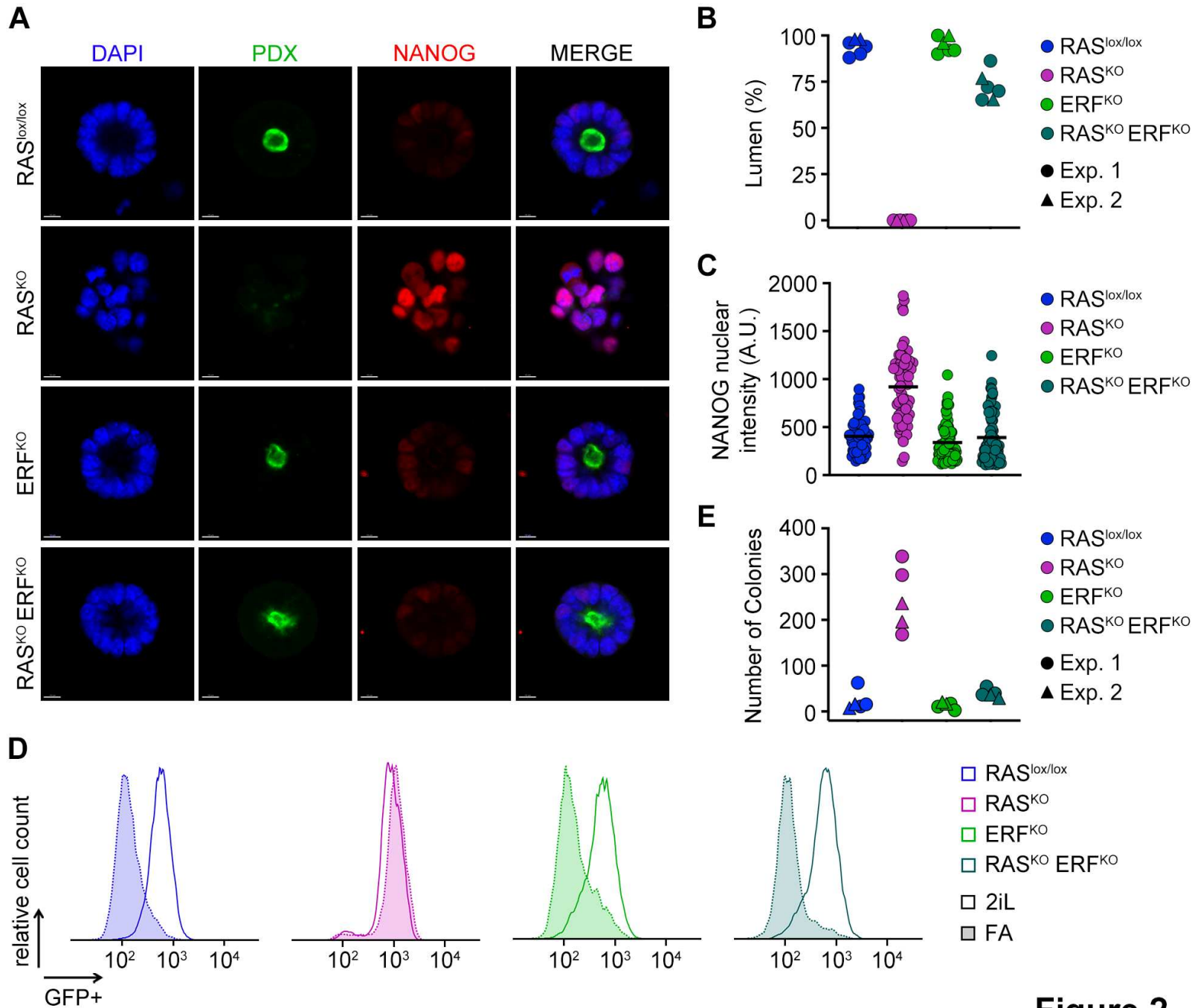933 was used to visualize nuclei. Scale bars, 15μm.

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

33

Figure 2

Fig. 2. Successful exit from naïve pluripotency requires downregulation of ERF. (**A**) Central confocal optical sections of RAS$^{lox/lox}$, ERF^{KO}, RAS^{KO}, and RAS^{KO}; ERF^{KO} embryonic cell rosettes embedded in matrigel 48 hours after seeding and stained for NANOG (red) and podocalyxin (PDX; green). DAPI was used to visualize nuclei. Scale bars, 10μm. (**B**) Graph showing the percentage of embryonic rosettes generating a lumen (PDX+) in all genotypes. Two independent experiments are shown and at least, a total of 50 rosettes were counted per sample. (**C**) Graph showing mean NANOG nuclear fluorescence intensity per nuclei in embryonic rosettes 48 hours after seeding. One representative experiment is shown and a total of 70 nuclei from different rosettes were counted per sample. (**D**) Flow cytometry analysis of REX1-deGFP reporter ESC from all genotypes in 2iL-ESC and EpiLSC after 48 hours of induction with FA. Three independent experiments were performed but one representative experiment is shown. (**E**) Graph showing the number of alkaline phosphatase positive colonies in a colony forming assay using ESC from all genotypes. Two independent experiments are shown with at least two technical replicates.
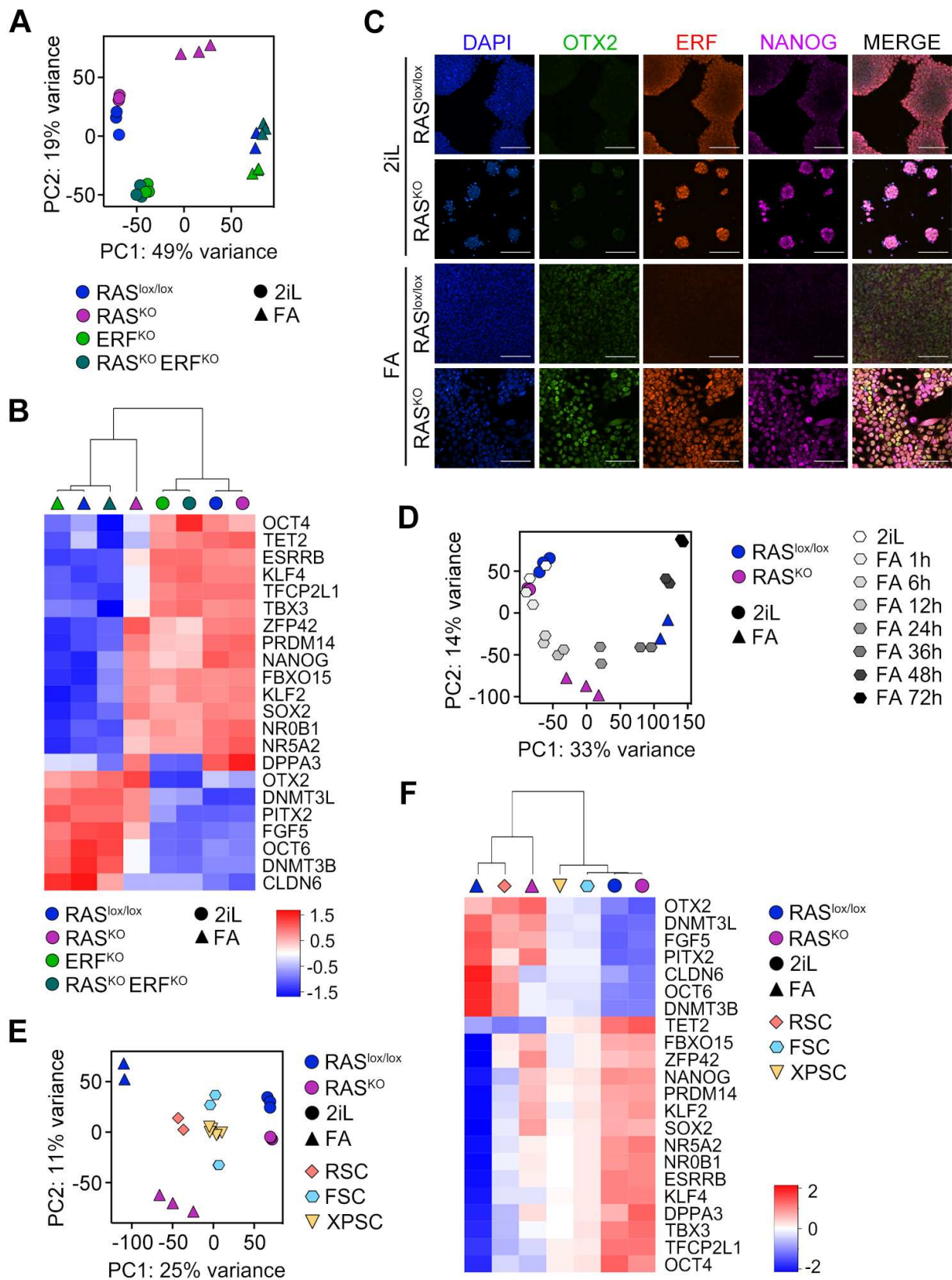
34

Figure 3

Fig. 3. ERF controls the transition to primed pluripotency. (**A**) PCA plot of RNAseq data of RAS$^{lox/lox}$, ERF^{KO}, RAS^{KO}, and RAS^{KO}; ERF^{KO} ESC cultured in naïve conditions (2iL) or induced to differentiate to EpiLSC (FA) for 48 hours. Three replicates per condition are shown. (**B**) Heatmap generated from RNAseq data from samples described in A) showing the average from three replicates. (**C**) Immunofluorescence analysis of 2iL and FA treated RAS$^{lox/lox}$ and RAS^{KO} ESC stained for OTX2 (green), ERF (red) and NANOG (purple). DAPI was used to visualize nuclei. Scale bars, 100μm. (**D**) PCA plot of RNAseq datasets (three replicates) showing 2iL and FA treated RAS$^{lox/lox}$ and RAS^{KO} ESC along with RNAseq datasets (two replicates) from a time course experiment during EpiLSC induction (0, 1, 6, 12, 24, 36, 48 and 72 hours) (*18*). (**E**) PCA plot of RNAseq datasets showing 2iL and FA treated RAS$^{lox/lox}$ and RAS^{KO} ESC along with RNAseq datasets from RSC (ESC lines grown in LIM (LIF, IWP2 (WNT inhibitor) and MEKi), FSC and XPSC (*6-8*). At least two replicates are shown. (**F**) Heatmap generated from RNAseq data from samples described in (E) showing the average from two or three replicates as applicable to each sample.
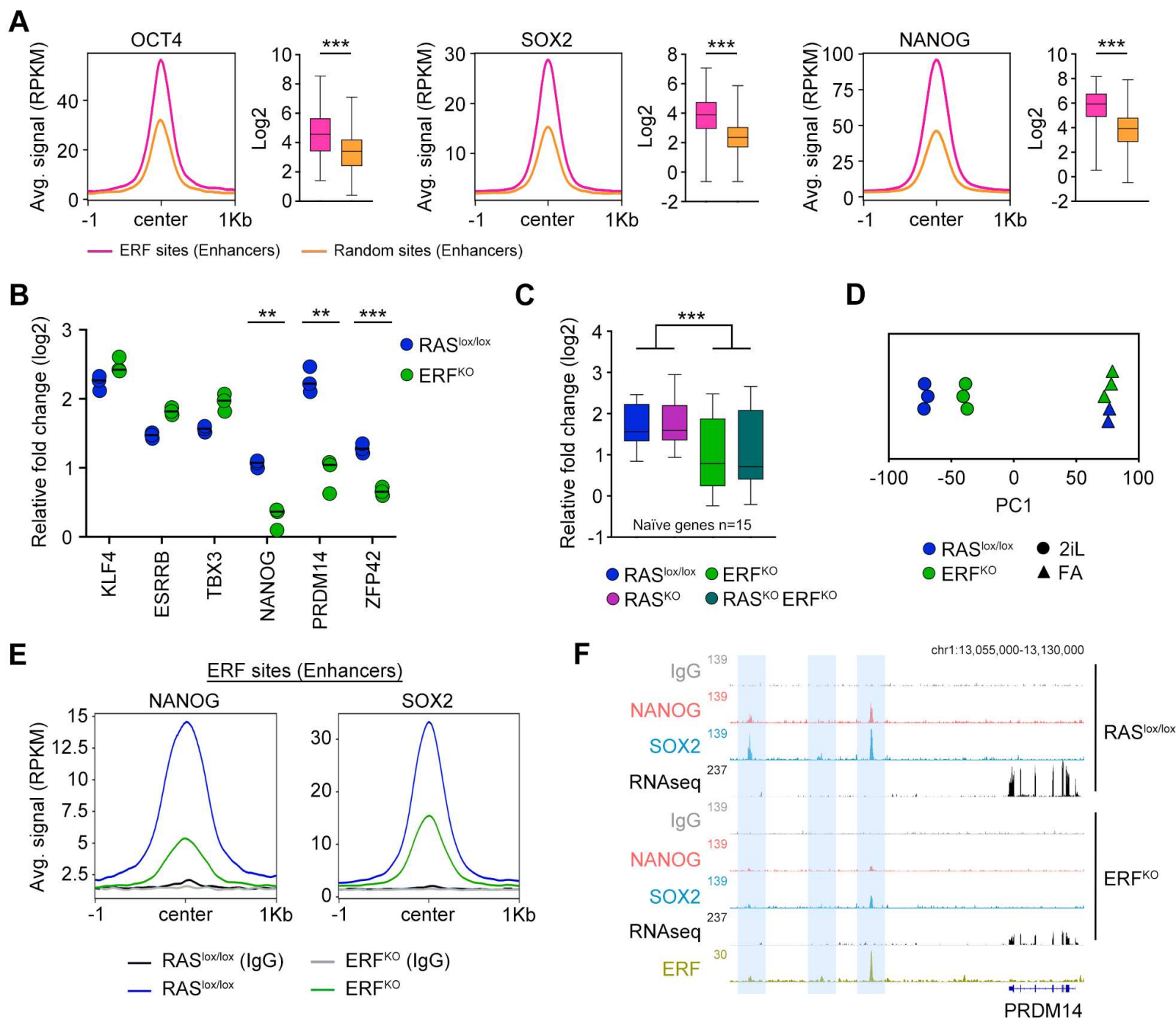
35

Figure 4

1009   **Fig. 4. ERF ensures optimal naïve pluripotent transcription factor expression in ESC. (A)** ChIPseq

1010   read density plot (RPKM) showing OCT4, SOX2 and NANOG occupancy at 2074 ERF-binding sites

1011   at enhancers (pink) or 2074 randomly selected non-ERF bound enhancers (orange). Graphs show

1012   quantifications of the TF enrichment in each set of sites. *** = p<0.001, T-student. Data was

1013   obtained from (*43*). **(B)** Graph showing relative fold change (log2) expression of the indicated

1014   genes in RAS$^{lox/lox}$ and ERF^{KO} ESC grown in 2iL. Genes showing at least a 50% reduction are

1015   highlighted. For each gene, data was normalized to the average across all samples. *** = p<0.001;

1016   ** = p<0.01. T-student. Data is shown as triplicates. **(C)** Graph showing the relative fold change

1017   (log2) expression of a subset of 15 naïve associated genes (*Klf4, Esrrb, Tfcp2l1, Tbx3, Klf2, Elf3,*

1018   *Klf8, Nanog, Prdm14, Tcl1, Zfp42, Nrob1, Prmd16, Dazl and Crxos*) in the different genotypes in

1019   2iL conditions. For each gene, data was normalized to the average across all samples. *** =

1020   p<0.001. T-student. **(D)** Unidimensional PCA plot of RNAseq data of RAS$^{lox/lox}$ and ERF^{KO} ESC

1021   cultured in naïve conditions (2iL) or induced to differentiate (FA) during 48 hours to EpiLSC. Three

1022   replicates per condition are shown. PC1 separates samples by differentiation status. **(E)** Cut&Run

1023   read density plot (RPKM) showing NANOG and SOX2 occupancy in the set of 2074 ERF-binding

1024   sites at enhancers in RAS$^{lox/lox}$ (blue) and ERF^{KO} (green) ESC cultured in 2iL. Corresponding inputs

1025   (IgG) are also shown as reference control. **(F)** Genome browser tracks showing NANOG and SOX2

1026   occupancy and RNAseq RPKM read count at the PRDM14 gene in the indicated genotypes. ERF

1027   binding profile in RAS^{KO} ESC is also shown. Inputs (IgG) are also shown as a reference control. ERF

1028   binding sites are highlighted.

1029

1030

1031
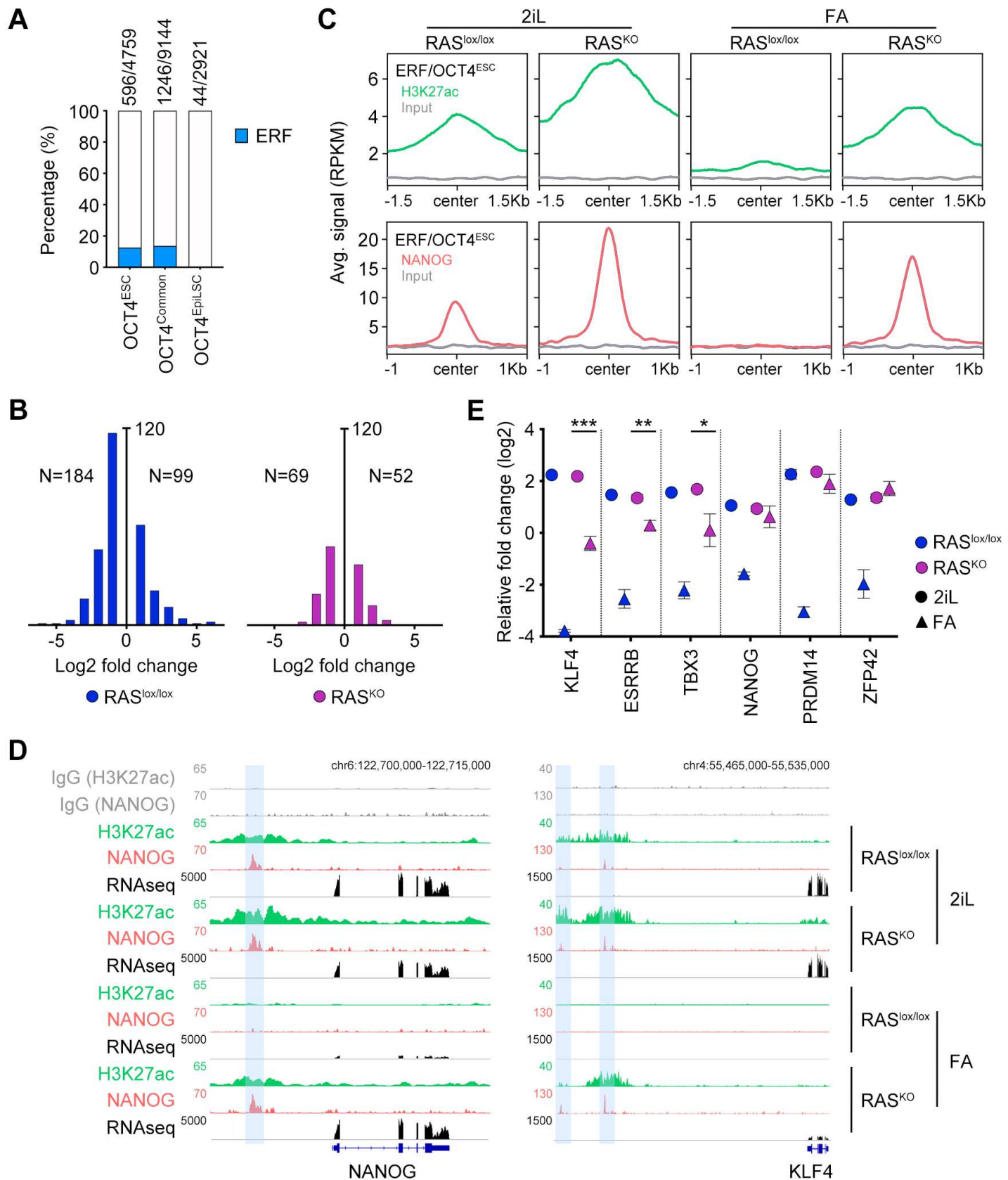
1032

1033

1034

1035

1036

Figure 5

Fig. 5: The naïve pluripotent transcription factor network is active in FA-RAS^KO ESC. (A) Plot showing the percentage of OCT4 binding sites co-occupied by ERF in ESC, EpiLSC and common between ESC and EpiLSC. (**B**) Histogram plots showing fold expression changes (log2) for genes associated to ERF/OCT4^{ESC} that are characterized by differential expression between ESC and EpiLSC in RAS$^{lox/lox}$ and RAS^{KO} cells. N indicates the total number of genes that are up or downregulated. Genes were associated at every binding site by proximity using GREAT (ref PMID: 20436461). (**C**) Cut&Run read density plot (RPKM) showing H3K27ac (plots above, green) and NANOG (plots below, red) occupancy in ERF/OCT4^{ESC} sites in RAS$^{lox/lox}$ and RAS^{KO} cultured in 2iL or differentiated to EpiLSC (FA). Corresponding inputs (IgG) are also shown as reference control. (**D**) Genome browser tracks showing H3K27ac deposition, NANOG occupancy and RNAseq RPKM read count at the KLF4 and NANOG genes in the indicated genotypes. Inputs (IgG) are also shown as a reference control. (**E**) Graph showing the relative fold change (log2) expression of the indicated ERF-bound super-enhancer associated genes in RAS$^{lox/lox}$ and RAS^{KO} cultured in 2iL or differentiated to EpiLSC (FA). Genes showing at least a 50% reduction are highlighted. For each gene, data was normalized to the average across all samples. *** = p<0.001; ** = p<0.01; *=<0.05 T-student. Data shown are averages from triplicates.
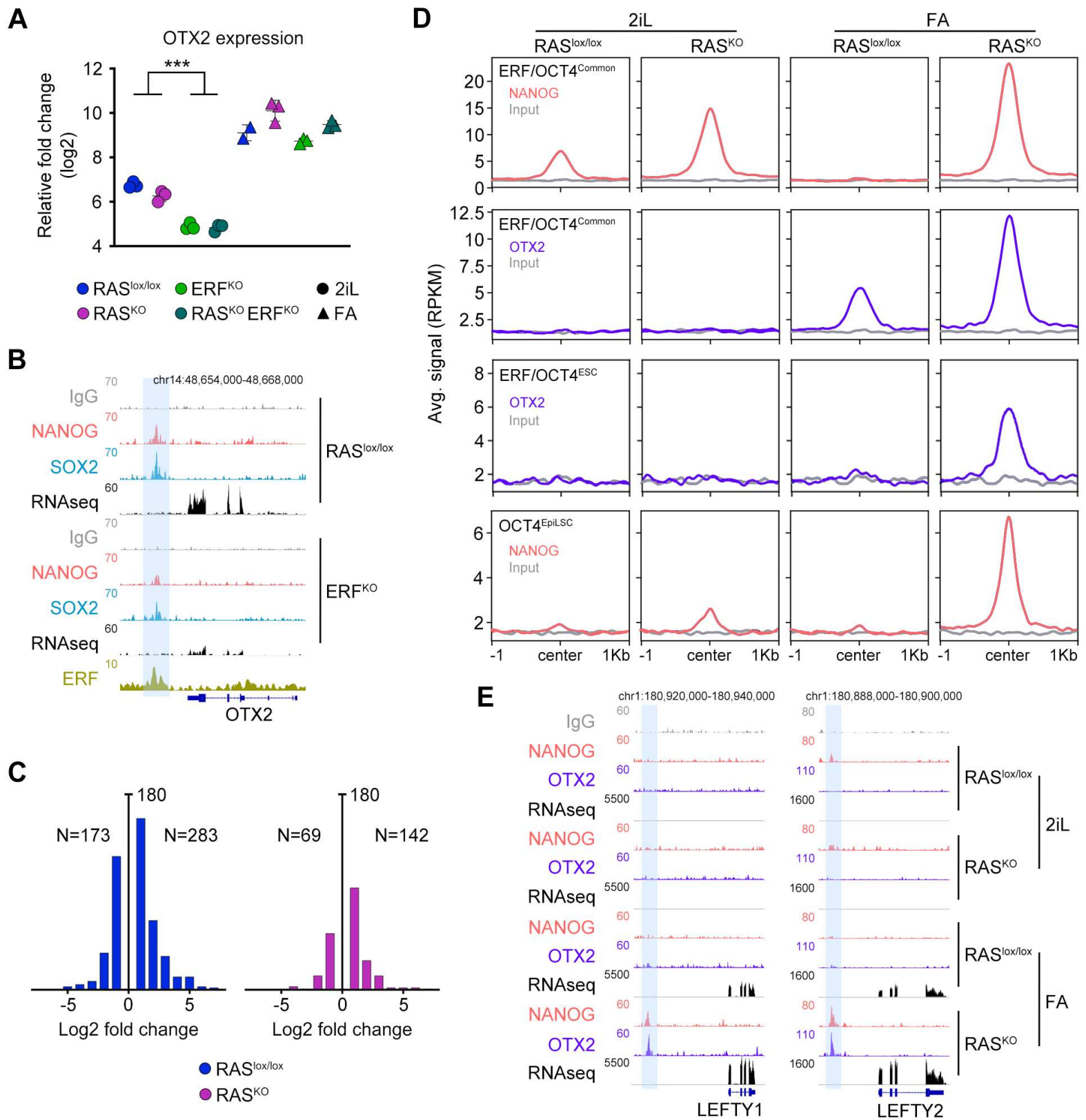
Figure 6

1066   **Fig. 6. OTX2 shows promiscuous binding in naïve and primed genes in FA-RAS^{KO} ESC. (A)** Plot

1067   showing the relative fold change (log2) expression for OTX2 in ESC from all genotypes cultured in

1068   2iL or differentiated to EpiLSC (FA). Data was normalized to the average across all samples. ***=

1069   p<0.001; T-student. Data shown from triplicates. (**B**) Genome browser tracks showing SOX2 and

1070   NANOG occupancy and RNAseq RPKM read count at the OTX2 gene in RAS$^{lox/lox}$ and ERF^{KO} ESC

1071   cultured in 2iL. ERF binding profile in RAS^{KO} ESC is also shown. Inputs (IgG) are also shown as a

1072   reference control. ERF binding sites are highlighted. (**C**) Histogram plots showing fold expression

1073   changes (log2) for genes associated to ERF/OCT4^{Common} that showed differential expression

1074   between ESC and EpiLSC in RAS$^{lox/lox}$ and RAS^{KO} cells. N indicates the total number of genes that

1075   are up or downregulated. Genes were associated at every binding site by proximity using GREAT

1076   (ref PMID: 20436461). (**D**) Cut&Run read density plot (RPKM) showing NANOG (red) and OTX2

1077   (purple) occupancy in the indicated ERF/OCT4^{Common}, ERF/OCT4^{ESC} and ERF/OCT4^{EpiLSC} sites in

1078   RAS$^{lox/lox}$ and RAS^{KO} ESC cultured in 2iL or differentiated to EpiLSC (FA). Corresponding inputs (IgG)

1079   are also shown as reference control. (**E**) Genome browser tracks showing NANOG and OTX2

1080   occupancy and RNAseq RPKM read count at the LEFTY1 and LEFTY2 genes in RAS$^{lox/lox}$ and RAS^{KO}

1081   ESC cultured in 2iL or differentiated to EpiLSC (FA). Inputs (IgG) are also shown as a reference

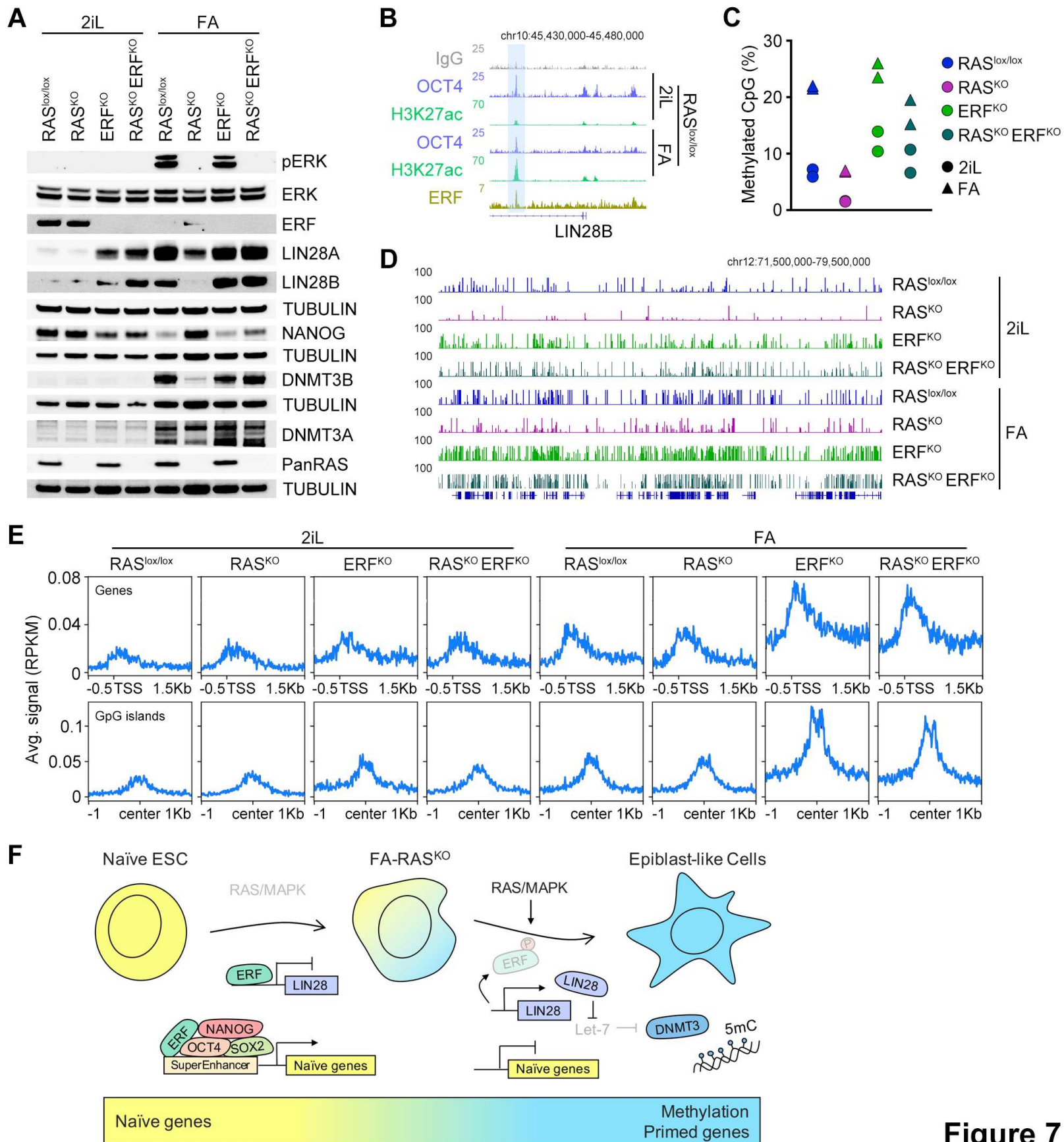1082   control. Blue squares showed ERF binding sites.

1083

1084

1085

1086

1087

1088

1089

1090

1091

1092

1093

Figure 7

1094   **Fig. 7. ERF controls the levels of *de novo* methylation during transition to EpiLSC. (A)** Western

1095   blot analysis of the indicated proteins performed in ESC from all genotypes cultured in 2iL or

1096   differentiated to EpiLSC (FA). Tubulin levels for the corresponding blots are shown as a loading

1097   control. One representative experiment is shown but at least two independent clones per

1098   condition and genotype were used. **(B)** Genome browser tracks showing H3K27ac deposition and

1099   OCT4 occupancy at the LIN28B gene in RAS[lox/lox] ESC cultured in 2iL or differentiated to EpiLSC

1100   (FA). ERF binding profile in RAS[KO] ESC is also shown. Inputs (IgG) are also shown as a reference

1101   control. ERF binding sites are highlighted. **(C)** Graph showing the percentage of methylated CpG

1102   sites identified by RRBS in ESC from all genotypes cultured in 2iL or differentiated to EpiLSC (FA).

1103   **(D)** Genome browser tracks showing as a representative example the level of methylation at CpG

1104   sites in a region of chromosome 12 in ESC from all genotypes cultured in 2iL or differentiated to

1105   EpiLSC (FA). **(E)** %CpG methylation averaged across the TSS of all protein-coding mouse genes

1106   (upper panels) or centered at CpG islands (lower panels) in ESC from all genotypes cultured in 2iL

1107   or differentiated to EpiLSC (FA). **(F)** Schematic model showing the dual role of ERF during the

1108   naïve to primed pluripotent transition. In the absence of FGF signaling, ERF ensures an optimal

1109   level of expression for naïve transcription factors. Following ERF phosphorylation and gene

1110   silencing, ESC are able to exit and transition into EpiLSC along with global CpG methylation and

1111   silencing of the naïve transcriptional network.

1112

1113

1114

1115

1116

1117

1118

1119

1120

1121

1122