ERH elicits cell lineage restriction in mammalian preimplantation development and differentiation from pluripotency via H3K9me3-heterochromatin

Andrew Katznelson*[1], Blake Hernandez[1], Kylea Tapia[2], Holly Fahning[2], Adam Burton[3], Jingchao Zhang[1], Maria-Elena Torres-Padilla[3,4], Nicolas Plachta[1], Kenneth S. Zaret[1]**, Ryan L. McCarthy[2,5,6]**

[1]Institute for Regenerative Medicine and Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[2]Center for Developmental Biology and Regenerative Medicine, Seattle Children's Research Institute, Seattle, WA 98101, USA

[3]Institute of Epigenetics and Stem Cells (IES), Helmholtz Zentrum München D-81377 München, Germany

[4]Faculty of Biology, Ludwig-Maximilians Universität, München, Germany

[5]Department of Pediatrics, University of Washington School of Medicine, Seattle, WA 98195, USA

[6]Lead contact

**co-corresponding authors

R. McCarthy, Ryan.McCarthy@seattlechildrens.org

K. Zaret, zaret@pennmedicine.upenn.edu

Key words: ERH, H3K9me3, heterochromatin, embryo, pre-implantation, differentiation, pluripotency, naïve

Abstract

Enhancer of Rudimentary Homolog (ERH) is an evolutionarily conserved protein originally characterized as promoting fission yeast heterochromatin[1] and recently shown to maintain H3K9me3 heterochromatin in human fibroblasts[2]. Here, we find that ERH depletion in fibroblasts reverts the somatic cell H3K9me3 landscape of broad megabase size domains to an embryonic stem cell (ESC) state composed of mainly H3K9me3 peaks and enables activation of naïve and pluripotency genes and transposable elements during induced pluripotent stem cell (iPSC) reprogramming. Concordantly, we find that ERH represses totipotent and alternative lineage programs during mouse preimplantation development and is required for proper segregation of the inner cell mass and trophectoderm cell lineages. During human ESC differentiation into germ layer lineages, ERH silences naïve and pluripotency genes, transposable elements, and alternative lineage somatic genes. As in fission yeast, we find that mammalian ERH interacts with RNA-binding proteins to engage and repress its chromatin targets. Our findings reveal a conserved, fundamental role for ERH in cell fate specification via the initiation and maintenance of early developmental gene repression.

38 **Main text**

39 Heterochromatin marked by histone H3 lysine 9 trimethylation (H3K9me3) represses the
40 expression of many protein-coding genes and DNA repeat elements[3–5]. Yet despite its stability,
41 H3K9me3-heterochromatin is dynamically established, maintained, and rearranged during
42 embryogenesis[3,6,7]. Zygotic genome activation (ZGA), marked by the initial transcription of key
43 genes and transposable elements[8], occurs at the 1-2 cell stage in mice and the 4-8 cell stage in
44 humans[8] and requires demethylation of maternal H3K9me3 marks created in oocytes[9,10]. As the
45 embryo loses totipotency and differentiates into the first two developmentally-specified lineages,
46 the trophectoderm of the blastocyst and the inner cell mass (ICM), many ZGA-specific protein-
47 coding genes and repeats gain H3K9me3 and are subsequently silenced [6,11,12]. As development
48 proceeds, H3K9me3 is further established on lineage-inappropriate genes and repeat elements
49 [6,11,13], with the highest number of genes marked by H3K9me3 at the germ-layer stage[3].
50 Throughout germ layer specification to mature somatic lineages, H3K9me3 is removed from
51 genes that become active and is deposited onto early developmental genes that become
52 repressed[3]. The H3K9me3 landscape changes substantially from pluripotency to early and
53 terminally differentiated states[14,15]. The mechanisms operating in the initial targeting of histone
54 methyltransferases to deposit H3K9me3 and the maintenance of the initially established
55 H3K9me3 domains remain poorly understood. In this study, we compared how *enhancer of*
56 *rudimentary homolog* (ERH), a highly conserved heterochromatin protein distinct from histone
57 methyltransferases, contributes to H3K9me3 dynamics and gene repression in differentiated
58 somatic cells, pluripotent cell lines, and early mammalian development.

59 ERH is an evolutionarily conserved protein originally discovered in fission yeast to regulate
60 sporulation, the primary form of yeast cell differentiation, via H3K9 methylation [1,16]. In *S. pombe,*
61 Erh1 forms a complex with the RNA binding protein Mmi1, which is recruited by binding of nascent
62 meiotic transcripts through its YTH domain[1,15]. Erh1-Mmi1 then causes the degradation of its
63 bound RNA transcripts, the establishment of H3K9 methylation, and consequent repression of

64 meiotic and developmental genes, transposons, and rDNA[1,16]. We recently found that in
65 terminally-differentiated primary human fibroblasts, ERH is highly enriched in H3K9me3-
66 heterochromatin, where it represses meiotic and germ cell genes, similar to that observed in
67 fission yeast[1,2,5]. We further found that ERH in mammals has evolutionarily expanded its function
68 to include the majority of regions that exhibit H3K9me3 heterochromatin in somatic cells, including
69 alternative-lineage protein coding genes and repeat elements[2]. While ERH is highly conserved
70 among eukaryotes, its fission yeast partner Mmi1 lacks a metazoan homolog[17]. Thus, how cell-
71 type specific cofactors recruit ERH to chromatin to regulate cell identity throughout mammalian
72 development is unknown.

73 Here, we find that upon ERH loss in fibroblasts, there is reversion to a pluripotency-like
74 H3K9me3 state which is permissive to binding of pluripotency transcription factors, which in turn
75 activate pluripotency and alternative lineage genes and repeat elements. The derepression of
76 pluripotency and alternative lineage programs led us to investigate the role of ERH in the earliest
77 developmental stages, where repression by H3K9me3 has been implicated during trophectoderm
78 and inner cell mass segregation in the blastocyst[6,11,12] and the differentiation from pluripotency to
79 the germ layer lineage[3]. By genetic perturbation, chromatin and gene expression profiling, and
80 imaging, we identify a role for ERH across mouse and human systems in repressing early naïve
81 pluripotency and alternative lineage genes. We find that targeting of mammalian ERH to
82 chromatin occurs through cell-type specific interactions with RNA binding proteins, enabling
83 context-dependent functions**.** Together, our findings demonstrate a fundamental role for ERH
84 during developmentally acquired H3K9me3-mediated gene repression and maintenance of such
85 repression in somatic cells.

86

87 **Results**

88

89 **ERH depletion in primary human fibroblasts causes H3K9me3 to revert to an ESC-like state**

90 We assessed how H3K9me3 is disrupted after ERH loss by CUT&RUN for H3K9me3 after siRNA
91 knockdown in human primary fibroblasts (Fig. 1a). Consistent with our previous study[2], we
92 confirmed that H3K9me3 is globally reduced in human fibroblasts after ERH knockdown (Fig. 1b).
93 Yet while large megabase-scale domains were diminished (Fig. 1c top, purple domains in top two
94 tracks), as seen previously[2],10,526 H3K9me3 residual peaks were retained in the ERH
95 knockdown (Fig. 1c, lower, second track) and Supplementary Table 1). Remarkably, the residual
96 H3K9me3 peaks in siERH fibroblasts resemble the H3K9me3 profile in H1 ESCs (Fig. 1c lower,
97 third track). Across the genome, 85% of the residual fibroblast siERH H3K9me3 sites overlap with
98 H3K9me3 peaks in H1 hESCs (Extended Data Fig. 1a). Pearson correlation of H3K9me3 peaks
99 from fibroblasts after siControl or siERH, as well as in hESCs, supports a shift in the H3K9me3
100 heterochromatin landscape to an embryonic state (Fig. 1d,e). Thus, while loss of ERH in primary
101 fibroblasts causes a global loss of H3K9me3 domains, embryonic H3K9me3-marked targets are
102 retained.

103 We hypothesized that since ERH depleted fibroblasts retain hESC-like H3K9me3 peaks
104 (Fig. 1c-e), ERH may be dispensable for maintaining H3K9me3 in hESCs. To test this hypothesis
105 we depleted ERH in H1 hESCs cultured in pluripotency-maintaining conditions, which caused no
106 observable changes in colony morphology or loss of pluripotency (Fig. 1f), even after sustained
107 knockdown (Extended Data Fig. 1d). H3K9me3 level and nuclear organization by
108 immunofluorescence was not changed in siERH-treated hESCs (Fig. 1g; Extended Data Fig. 1e).
109 We also observed no change in H3K27me3 (Extended Data Fig. 1f). Despite no apparent
110 disruption of the pluripotent state or global H3K9me3 levels, ERH-depleted hESCs show an
111 increased nuclear area and nuclear blebbing (Extended Data Fig. 1g), which are changes
112 associated with chromatin decompaction[18]. We previously found that ERH depletion did not
113 change the expression or protein level of H3K9me3 HMTs[2], we similarly observe no significant
114 decrease in H3K9me3 HMT expression upon ERH depletion in pluripotent hESCs (Extended Data
115 Fig. 1h). To investigate potential locus-specific changes in hESC H3K9me3 upon ERH depletion,

116 we conducted H3K9me3 CUT&RUN in siControl and siERH treated hESCs. Consistent with our
117 immunofluorescence results, we observed minimal change in H3K9me3 (Extended Data Fig. 1i)
118 and high correlation between H3K9me3 peaks in siControl and siERH treated hESCs (Extended
119 Data Fig. 1j).

120 To identify chromatin-associated proteins and pathways regulating the residual H3K9me3
121 peak sites in human fibroblasts, we analyzed the regions using the cistromeDB Toolkit, which
122 uses a compendium of published ChIP-seq data sets[16,17]. We find that most proteins enriched at
123 the residual peak sites in ES cells (i.e., prior to ERH knockdown) are either KRAB-ZNFs or
124 involved in H3K9me3 repression via KAP1, which is recruited to chromatin through binding the
125 KRAB domain (Fig. 1h)[19]. Likewise, de novo motif analysis finds that enriched sequence motifs in
126 the residual peak sites are for KRAB-ZNFs factors (Extended Data Fig. 1k). Consistent with the
127 enrichment analysis, the residual siERH H3K9me3 peak sites are highly enriched for H3K9me3,
128 KAP1, and HP1 in ES cells (Fig. 1e,i; Extended Data Fig. 1l). As KRAB-ZNFs repress transposons
129 via H3K9me3, we hypothesized the residual peaks would be enriched over repeats. Repeat
130 enrichment analysis found that the residual peaks are enriched over canonical KRAB-ZNF
131 targets: LINEs, LTRs, and satellite repeats (Extended Data Fig. 1m). We conclude that ERH is
132 required to maintain the broad H3K9me3 domains that are established over the course of
133 development and are present in terminally differentiated cells, but does not control H3K9me3
134 peaks attributed to the DNA-dependent association of KAP1 that persists from pluripotency over
135 constitutively repressed repeats.

136

137 **ERH-dependent H3K9me3 suppresses SOX2 binding and activation of pluripotency genes**
138 **and repeats during OSKM reprogramming**

139 H3K9me3 impairs induced pluripotent stem cell (iPSC) reprogramming by blocking transcription
140 factor binding, and pluripotency genes marked by H3K9me3 show delayed activation compared
141 to genes in heterochromatin marked with H3K27me3 or neither mark (Extended Data Fig. 2a)[20].

142 We therefore assessed the impact of ERH depletion in the context of OSKM transcription factor
143 binding and gene activation during iPSC reprogramming. We focused on a 48 hour time point
144 after initiation of iPSC reprogramming (Fig. 1a and Supplementary Tables 2-4), as H3K9me3 is
145 inhibitory to transcription factor activity at this stage[20].

146 CUT&RUN of SOX2 after 48 hours of OSKM reprogramming showed that while most
147 binding sites were shared between siControl and siERH fibroblasts, 2,538 binding sites were
148 gained (Extended Data Fig. 3a and Supplementary Table 2). We observed that ERH knockdown
149 enabled SOX2 to target loci that lose H3K9me3, including within heterochromatin domains
150 previously annotated as resistant to reprogramming factor binding in fibroblasts[20] (Fig. 2a, and
151 Extended Data Fig. 3b-d). By assessing the genomic locations of siERH-gained SOX2 peaks in
152 the cistromeDB toolkit[21,22], we found that the newly gained SOX2 sites had the most similarity to
153 ChIP-seq peak sets of SOX2 and its pluripotency partners NANOG and OCT4 in embryonic stem
154 cells (Extended Data Fig. 3e). Inspection of genome browser tracks identified specific loci where
155 loss of ERH-controlled H3K9me3 enabled SOX2 binding over normal SOX2 and NANOG targets
156 in ES cells (Fig. 2b, Extended Data Fig. 3f). We next assessed if loss of ERH-dependent
157 H3K9me3 enables other transcription factors to bind previously refractory sites by performing
158 CUT&RUN footprinting of SOX2's primary binding cofactors during reprogramming, KLF4 and
159 OCT4, and find enhanced footprints for each factor at SOX2 sites gained by siERH (Extended
160 Data Fig. 3g). Thus, ERH loss during pluripotency reprogramming enables transcription factor
161 binding in what were heterochromatic H3K9me3 regions.

162 Protein coding genes embedded in H3K9me3 domains, normally refractory to transcription
163 factor activation, were activated after siERH (136 up, 13 down; Fig. 2c). Significantly, genes
164 associated with the human 8-cell like cell (8CLC) state[23] were activated, including *ZSCAN4,*
165 *DUXA, DPPA4, LIN28A, TPRXL*[24]*, KLF5,* and the histone variant *H3.Y*[25] (Fig. 2d and Extended
166 Data Fig. 4a). We compared the activation levels of these genes in siERH+OSKM with existing
167 RNA-seq datasets[26] and found that *H3.Y, DUXA* and *TPRXL* do not normally get activated at

168 similar levels to siERH+OSKM at any stage during iPSC reprogramming (Extended Data Fig. 4b).
169 However, activation of 8CLC genes in siERH+OSKM is incomplete, as many genes activated in
170 naïve ESCs[27] are not expressed or expressed at lower levels in siERH+OSKM (Extended Data
171 Fig. 4c), suggesting low activation of naïve programs but not full conversion to a naïve cell state.
172 While the activation of marker genes such as DUXA and ZSCAN4 is lower than in 8C embryos[28],
173 we conclude that ERH-dependent H3K9me3 safeguards early totipotent and pluripotent genes
174 and that ERH loss creates a permissive environment for transcription factor binding and
175 activation.

176 We observed that *LTR7* and *HERVH* repeats, which are expressed in and required to
177 maintain pluripotency[29,30], lose H3K9me3 after ERH depletion and are activated in siERH+OSKM
178 (Fig. 2e). *LTR7Y* repeats, which promote naïve pluripotency[31], exhibited similar ERH-dependent
179 H3K9me3 and repression (Fig. 2f,g and Extended Data Fig. 4d). Notably, residual H3K9me3
180 peaks retained at *LTR7Y* repeats were insufficient to maintain repression (Fig. 2f,g). Furthermore,
181 we observed siERH+OSKM activated repeat elements that function specifically during the human
182 8 cell-like cell stage, and are later marked by H3K9me3 and silenced in the human blastocyst[6]
183 (Extended Data Fig. 4e). These results demonstrate that ERH maintains H3K9me3 repression of
184 a subset of 8-cell stage and pluripotency repeats in terminally differentiated human cells.

185 Depletion of H3K9 HMTs has been shown to increase iPSC reprogramming[20]. However,
186 despite increased activation of pluripotency and 8-cell genes at 48 hours after initiation of OSKM-
187 based iPSC reprogramming in siERH treated fibroblasts, by day 2 we observed substantial cell
188 death that persisted 7 days into reprogramming, past the initial apoptotic wave (Extended Data
189 Fig. 4f,g) and resulting in no iPSC colonies forming in siERH+OSKM conditions, compared to the
190 abundant colonies in siControl+OSKM conditions (Extended Data Fig. 4h,i). The failure to
191 generate iPSC colonies was not rescued by simultaneous knockdown of p53 (Extended Data Fig.
192 4i). Lower ERH siRNA decreased cell death, but did not change the emerging iPSC colony
193 number (Extended Data Fig. 4g,h). Using XDeathDB[32] on our RNA-seq data identified increased

194 expression of genes activated in several cell death pathways, including RIPK2 signaling
195 (Extended Data Fig. 4j). We investigated if simultaneous inhibition of RIPK2 along with inhibition
196 of RIPK3 and Caspases 3 and 7, which has been shown to improve survival of cells activating 8
197 cell like genes[23], could enable survival of siERH+OSKM iPSCs. Notably, we found that adding
198 small molecule inhibitors GSK2983559, Ac-DEVD-CHO, and GSK872 to inhibit RIPK2, Caspases
199 3 and 7, and RIPK3, respectively, during reprogramming enabled the generation of iPSC colonies
200 with ERH depletion at a significantly higher efficiency compared to siControl (Extended Data Fig.
201 4k). RIPK2 inhibition alone was necessary and sufficient to successfully generate iPSC colonies
202 expressing pluripotency factors during ERH depletion, while inhibition of RIPK and Caspases 3
203 and 7 was not sufficient but did increase colony size (Extended Data Fig. 4l). These findings
204 demonstrate that the increased accessibility for transcription factor binding in H3K9me3 domains
205 (Fig. 2a,b) and improved activation of the pluripotency network (Fig. 2c,d) observed upon ERH
206 depletion can increase iPSC reprogramming efficiency.
207

208 **ERH is necessary for trophectoderm development, the first embryonic differentiation**

209 Our finding that ERH represses naïve genes that are normally expressed in 2-8 cell embryos, in
210 fully differentiated fibroblasts, led us to investigate if ERH establishes silencing of 2-8 cell genes
211 during pre-implantation development. We first analyzed existing RNA-seq data of mouse
212 embryogenesis timepoints from zygote to blastocyst[33] and found that ERH mRNA is present
213 across pre-implantation development and is significantly increased from the 8-cell stage through
214 the blastocyst stage (Extended Data Fig. 5a). Immunostaining mouse embryos for ERH protein
215 at the zygote, 2-cell, 8-cell and blastocyst stages revealed that, consistent with its mRNA levels,
216 ERH is expressed and nuclear-localized at all stages analyzed, with increasing abundance in the
217 early and late blastocyst stages (Fig. 3a,b). At the blastocyst stage, ERH levels are similar in ICM
218 and trophectoderm cells (Extended Data Fig. 5b).

219 To investigate if ERH is involved in pre-implantation lineage specification, we depleted
220 ERH levels by siRNA micro-injection at the 1-cell stage of mouse embryos, compared to a
221 nontargeting siRNA control, and assessed impacts on the blastocysts (Extended Data Fig. 5c).
222 ERH depletion did not cause major morphological changes in blastocyst structure, as both
223 siControl and siERH blastocysts form a cavity and segregate ICM and trophectoderm cells
224 (Extended Data Fig. 5c,d). Staining blastocysts for key transcription factors involved in lineage
225 specification revealed similar levels of the pluripotency marker OCT4 in ICM and TE cells between
226 siControl and siERH blastocysts (Extended Data Fig. 5e). However, the levels of the
227 trophectoderm-determining transcription factor CDX2[34] (Fig. 3b,c) and the trophectoderm marker
228 Keratin 8[35] (Fig. 3d,e) were significantly decreased in the ERH depleted blastocysts. Thus, ERH
229 is required for proper trophectoderm differentiation.

230 The role of ERH in H3K9me heterochromatin regulation in *S. pombe*[1,16] and human
231 fibroblasts[2] suggests that defects in trophectoderm differentiation may arise from alterations in
232 gene silencing mechanisms. Indeed, we observed a 32% decrease (p=0.0025) in H3K9me3 in
233 blastocysts of ERH-depleted embryos (Fig. 3f,g and Supplementary Table 5). We also observed
234 a significant increase in nuclear DNA shedding in the siERH blastocytes (Fig. 3h,i; Extended Data
235 Fig. 5f), similar to the observed nuclear DNA shedding in siERH hESCs (Extended Data Fig. 1g).
236 Decreased keratin levels as observed in siERH blastocytes (Fig. 3d,e) have been shown to make
237 trophectoderm cells of blastocyts more susceptible to DNA shedding[36], as has decreased
238 H3K9me3[18].

239 To characterize how these changes impact gene expression, we performed RNA-seq on
240 mouse blastocysts 4 days after injecting siControl or siERH at the 1 cell stage (Supplementary
241 Table 6), which confirmed an ~80% knockdown of ERH (Extended Data Fig. 5g). We classified
242 genes upregulated in ERH depleted blastocysts into 4 groups, based upon their expression during
243 normal mouse preimplantation development[33] (Fig. 3j; Supplementary Table 7). Group 1 consists
244 of transcripts present at the zygote stage, including maternally inherited transcripts, that become

245 silenced or degraded as development progresses. Group 2 consists of transcripts that are highly
246 expressed in the 2 to 8 cell stages and low or absent in the zygote and blastocyst stage. Group 3
247 consists of transcripts that are absent or lowly expressed at all stages measured. Group 4 consists
248 of transcripts with the highest expression in the blastocyst stage. Of the ERH-repressed genes
249 (i.e., genes whose expression goes up upon ERH knockdown), 25 normally have peak expression
250 during the 2 to 8 cell stage (group 2) and 54 are normally not expressed during preimplantation
251 development (group 3). ERH-repressed group 2 genes included the 2-cell specific *Zscan4d*[37] and
252 the 4-cell specific *D1Pas1*[38] (Fig. 3k). ERH-repressed group 3 alternative lineage genes included
253 genes specific to the brain[39], liver, and intestine [40] (Fig. 3k).

254 Comparing our RNA-seq data with existing mouse preimplantation H3K9me3 ChIP-seq
255 data[11], we found that the genes repressed by ERH at the 2-cell stage and later, alternative lineage
256 genes were initially marked by H3K9me3 (Fig. 3k). We further found that upon ERH depletion,
257 key markers of primitive endoderm, including *Gata4* and *Sox17*[41], were significantly repressed
258 (Extended Data Fig. 5h), while epiblast genes were unchanged (Extended Data Fig. 5i). We
259 conclude that ERH is necessary to silence 2-8 cell stage genes and prevent the spontaneous
260 activation of later developmental genes during the blastocyst stage, and that failure to silence
261 results in improper lineage specification in the blastocyst.

262

263 **ERH is necessary for proper lineage commitment and restriction during human ESC**
264 **differentiation**

265 Our finding that ERH is critical for blastocyst development led us to ask if ERH depletion impacted
266 the ability of hESCs to differentiate. When we used Activin A and the GSK-3 Inhibitor CHIR99021
267 together to differentiate hESCs to an endoderm identity[42], as expected within 24 hours we
268 observed repression of pluripotency factors and initial expression of endodermal markers in a
269 subset of cells (Extended Data Fig. 6a). Yet depletion of ERH in H1 hESCs followed by 24 hour
270 endoderm differentiation showed a failure to decrease NANOG protein levels (Fig. 4a and

271 Extended Data Fig. 6b) and a reduced emergence of SOX17+ cells (Fig. 4a). To assess if this
272 observation was specific to endoderm, we also performed differentiation to ectoderm, with control
273 and ERH depleted cells, for 2 days with retinoic acid treatment. We observed sustained
274 expression of the pluripotency factor OCT4 and decreased abundance of the ectodermal
275 transcription factor PAX6 (Fig. 4b).

276 To investigate transcriptional changes caused by ERH depletion, we performed RNA-seq
277 of siRNA control and ERH depleted H1 hESCs (Extended Data Fig. 6c and Supplementary Table
278 8) during pluripotent culture, after 1 day of endoderm or 2 days of ectoderm differentiation. While
279 ERH depletion in fibroblasts+OSKM resulted in more gene activation than repression (Fig. 2c;
280 Extended Table 3), depletion in pluripotent hESCs exhibited less derepression (Extended Data
281 Fig. 4d; Extended Table 8), consistent with the difference in the importance of ERH in H3K9me3
282 maintenance between fibroblasts and hESC state (Fig. 1). Genes normally repressed in the
283 siControl endoderm and ectoderm differentiation were preferentially upregulated by ERH
284 depletion (endoderm, Extended Data Fig. 6f; ectoderm Extended Data Fig. 6g).

285 During endoderm differentiation, we observed that ERH depletion preferentially
286 upregulated genes identified as higher in naïve vs primed hESCs[43] (Extended Data Fig. 6h),
287 beyond what was observed in the pluripotent state (Extended Data Fig. 6e). For example, the
288 gene for the pluripotency factor NANOG, previously demonstrated to be repressed by H3K9me3
289 during pluripotency exit[44], was significantly upregulated by ERH depletion (Fig. 4c), consistent
290 with our immunostaining (Fig. 4a and Extended Data Fig. 6b). Additionally, genes found to be
291 expressed in human pluripotent stem cells were upregulated during endoderm and ectoderm
292 differentiation after ERH depletion, including *DUSP5/6, NODAL, LEFTY1/2*, and *DACT1*
293 (Extended Data Fig. 6i and Supplementary Table 8). Genes associated with the 2 cell stage and
294 naïve pluripotency, including *ZSCAN4*[37], *KLF5*[31], *ESRRB*[45], *NR5A2*[46] and *NANOGP1*[47], were
295 upregulated during siERH-treated hESC differentiation (Fig. 4d and Extended Data Fig. 6j). The
296 expression of naïve genes after ERH depletion was at a much lower level than in true naïve

297 hESCs (Extended Data Fig. 6k), indicating permissive expression but not spontaneous primed to
298 naïve conversion. We conclude that ERH is required to repress pluripotency genes during hESC
299 pluripotency exit.

300 During differentiation, a key function of H3K9me3 heterochromatin is to repress alternative
301 lineage genes[3,11]. Our finding that ERH depletion in blastocysts upregulates lineage-incompatible
302 genes (Fig. 3j,i) led us to investigate if a similar consequence occurs during hESC differentiation.
303 GO analysis of genes upregulated by ERH depletion during endoderm and ectoderm
304 differentiation confirmed activation of transcriptional networks typically not induced by the specific
305 differentiation treatment (Fig. 4e, Extended Data Fig. 7a and Supplementary Table 10). ERH
306 depletion upregulated meiotic and gametogenic genes in all tested developmental stages
307 (Extended Data Fig. 7b), consistent with previous findings in human fibroblasts[2] and the role of
308 ERH in S. pombe sporulation[1]. To further assess the cell lineages in which the upregulated genes
309 function, we generated composite expression profiles of the genes upregulated by ERH depletion
310 during endoderm (Fig. 4f) and ectoderm differentiation (Fig. 4g) from expression data downloaded
311 from the BBI Descartes human single cell atlas[48], and found that heart, muscle and liver genes
312 are preferentially activated. Directed differentiation of hESCs to endoderm[42], instead induced
313 both early and late mesodermal genes[49] when ERH was depleted (Fig. 4h,i). Notable
314 inappropriate gene activation in siERH definitive endoderm included *NXK2-5, NPPB* and *FLNC*,
315 typically expressed in heart muscle, as well as *TAGLN,* a canonical smooth muscle marker (Fig.
316 4i). Increased levels of BRACHYURY protein, a key determinant of early mesoderm[50], was also
317 confirmed by immunofluorescence (Fig. 4j). Inappropriate activation during ectoderm
318 differentiation with ERH knockdown included genes typically expressed in liver development
319 (Extended Data Fig. 7c). Thus, while ERH appears dispensable to maintain hESC pluripotency,
320 it is essential to restrict cell identity for proper germ layer differentiation.

321

322 **ERH represses transposable elements and proximal genes**

323 Having observed that ERH represses pluripotency and alternative lineages during hESC
324 differentiation, we proceeded to investigate how this repression was elicited. During pluripotency-
325 to-endoderm differentiation at protein-coding genes that are upregulated in siERH DE vs siControl
326 DE, H3K9me3 levels only slightly increase in and around those genes (i.e., without siERH) during
327 endoderm differentiation (Extended Data Fig. 8a). However, repeat elements including L1 LINEs
328 do normally exhibit a substantial H3K9me3 gain during endoderm differentiation (Fig. 5a).
329 Repression of transposable elements is a canonical function of H3K9me3 heterochromatin[51] and
330 transposable elements can regulate developmental gene networks by acting as enhancers for
331 transcription factor binding[31,52–55]. We therefore assessed transposable elements in our RNA-seq
332 data from siControl and siERH-treated H1 hESCs during self-renewal and differentiation
333 (Supplementary Table 9), with the hypothesis that dysregulation of specific transposable elements
334 might underlie the observed differentiation defects in siERH treated hESCs.

335 We found that transposable elements were upregulated by ERH knockdown in H1 hESCs,
336 definitive endoderm, and ectoderm, including classes of L1 elements and the hominid-specific[56]
337 SINE VNTR and Alu (SVA) elements (Fig. 5b; Extended Data Fig. 8b). Satellite repeats were also
338 derepressed (Fig. 5b), as were morula-blastocyst specific *LTR7Ys*[31] and human 8 cell stage
339 expressed *HERVKs*[57] (Fig. 5b; Extended Data Fig. 8c,d), which act as transcription factor binding
340 sites for pluripotency[52,58] and lineage specific genes[55]. After stratifying L1 elements by their
341 evolutionary age[59], we found that ERH preferentially represses evolutionarily young L1 repeats in
342 pluripotent and differentiating hESCs, but broadly represses L1 elements in terminally
343 differentiated fibroblasts (Fig. 5c). By filtering our data for genes expressed during pluripotency
344 that exhibit ERH-dependent repression during differentiation, we uncovered evolutionarily young
345 transposable elements (Fig. 5d and Extended Data Fig. 8e,f) and L1 repeats within 25 kb
346 (Extended Data Fig. 8g) in regions normally marked by H3K9me3 in hESC-derived endoderm
347 (Fig. 5e note different scales on subset to right; Extended Data Fig. 9a).

348 To assess if repeat expression upon ERH knockdown during endoderm differentiation was
349 due to H3K9me3 loss, we performed H3K9me3 CUT&RUN in siControl and siERH hESCs
350 differentiated to endoderm. We found the increased L1 element RNA expression upon ERH
351 depletion (Fig. 5c) did not globally correspond to a loss of H3K9me3 (Extended Data Fig. 9b).
352 However, when we restricted the analysis to L1 TEs that gained H3K9me3 with at least a 2-fold
353 increase in RNA in siERH endoderm, the L1 loci showed decreased H3K9me3 upon ERH
354 depletion (Fig. 5f). We found repeat element loci with increased RNA expression with (Fig. 5g left)
355 and without H3K9me3 loss (Fig. 5g right) indicating that RNA expression from TEs upon ERH
356 depletion during endoderm differentiation was not necessarily dependent upon H3K9me3 loss.
357 We observe instances of endoderm specific upregulation of genes proximal to evolutionarily
358 young repeats, such as L1 and LTR7Ys, with reduced H3K9me3 in ERH depleted endoderm
359 (Extended Data Fig. 9c). ERH therefore represses evolutionarily young and lineage-specific
360 repeats, thus preventing improper activation of pluripotency and alternative lineage genes during
361 differentiation to maintain lineage specification integrity. The role of H3K9me3 and ERH at this
362 cell fate change likely occurs predominantly at transposable element regions and not gene bodies.
363

364 **ERH is recruited to chromatin by cell-specific cofactors**

365 Our observation that ERH is required for maintenance of H3K9me3 in fibroblasts (Fig. 1b) but
366 dispensable for hESC H3K9me3 (Fig. 4b) led us to investigate potential ERH protein interactions
367 that may drive the difference in regulatory mechanism. We focus the experiments on the
368 pluripotent and terminally differentiated states as they exhibit contrasting requirements for ERH
369 in H3K9me3 maintenance (Fig. 1b-g) and very different H3K9me3 distribution patterns (Fig. 1d,
370 i; Extended Data Fig 1g). In *S. pombe* Erh is recruited by the RNA binding factor Mmi1[1]. Mmi1
371 itself is not seen in mammalian cells[60,61], but since it is a YTH domain-containing RNA binding
372 factor, we investigated potential interactions of ERH with human YTH domain proteins YTHDC1
373 and YTHDC2. YTHDC1 regulates H3K9me3 through SETDB1 recruitment and repression of

374 repeat elements in mouse ESCs[62,63] and YTHDC2 has been shown to positively regulate
375 transposable elements and the non-coding RNA LINC-ROR in human hESCs[64]. To specifically
376 interrogate on chromatin ERH interacting proteins, we isolated soluble nucleoplasm (NPL) and
377 insoluble chromatin-associated fractions (CAF) and then performed immunoprecipitation of ERH
378 (Fig. 6a). In fibroblasts but not hESCs, we observed that ERH interacts with YTHDC1 in the
379 chromatin fraction (Fig. 6b,c). ERH did not interact with YTHDC2 in the chromatin fraction in either
380 hESCs or fibroblasts, despite being present in both cell types (Fig. 6b,c). We also probed for
381 SAFB1, a known ERH interacting protein[65], and observed consistent ERH-SAFB1 interactions in
382 both fibroblasts and pluripotent hESCs (Fig. 6b,c), demonstrating that YTHDC1 but not SAFB1
383 interactions with ERH are cell type-dependent. To investigate the cause of this differential
384 interaction we first analyzed ERH levels, finding that while ERH RNA is not significantly different
385 between pluripotency and fibroblasts (Extended Data Fig. 10a) or between pluripotent and
386 differentiated hESCs (Extended Data Fig. 6c), fibroblasts stain more strongly for ERH protein
387 (Extended Data Fig. 10b). We observe that ERH protein levels are higher in cells on the edge of
388 hESC colonies (Extended Data Fig. 10c) and become higher in all cells after 1 day of endoderm
389 differentiation (Extended Data Fig. 10c). These results indicate that ERH involvement in
390 H3K9me3 regulation and RNA binding protein regulation may be regulated by ERH protein level.

391 As S. pombe Erh1 chromatin recruitment is RNA and Mmi1 dependent, we next tested if
392 chromatin association of ERH was RNA dependent and if either YTHDC1 or YTHDC2 facilitated
393 the recruitment. We employed an immunostaining approach[66] in which permeabilized nuclei were
394 treated with RNase A and then stained for ERH, and found that ERH nuclear retention was
395 significantly decreased (Fig. 6d). In contrast, RNase H treatment, which digests DNA:RNA
396 hybrids, did not release ERH from the nucleus (Fig. 6d). In an orthogonal approach, RNase A
397 treatment of fibroblasts during soluble nuclear protein isolation released a portion of chromatin-
398 bound ERH into the nucleoplasm, indicating that its chromatin retention is at least partially RNA-

399 dependent (Extended Data Fig. 10d). To determine if YTHDC1 or YTHDC2 facilitate ERH
400 chromatin recruitment, we targeted each separately with siRNAs in human fibroblasts and
401 observed that YTHDC1 but not YTHDC2 depletion led to decreased levels of ERH in the
402 chromatin fraction and increased ERH in the soluble nuclear fraction (Fig. 7e). We confirmed the
403 role of YTHDC1 in ERH retention utilizing our cell permeabilization and immunostaining assay
404 (Extended Data Fig. 10e). From these experiments we concluded that ERH associates with
405 chromatin with YTHDC1 in fibroblasts, but not in hESCs, and that a subset of ERH protein requires
406 RNA and YTHDC1 for chromatin localization.

407 We next assessed for co-regulation of common targets by ERH, YTHDC1, and YTHDC2,
408 first by evaluating expression changes of published YTHDC1 and YTHDC2 targets, ERVK
409 elements repressed by YTHDC1[67], and HERVH elements activated by YTHDC2[64]. ERH
410 represses HERVK and HERVH elements in fibroblasts treated with OSKM factors but not in
411 pluripotent hESCs (Fig. 6f), matching the interaction pattern of ERH with YTHDC1 (Fig. 6b,c).
412 Major satellite repeats, which are regulated by SAFB1[68], are repressed by ERH in both fibroblasts
413 and pluripotent hESCs (Fig. 6f).

414 As we found that ERH interacts with YTHDC1 in differentiated cells but not pluripotent
415 cells, we investigated if YTHDC1 depletion recapitulates aspects of ERH depletion during hESC
416 differentiation. Depletion of YTHDC1 but not YTHDC2 during endoderm differentiation of hESCs
417 resulted in an increase of Brachyury staining indicative of a mesodermal fate (Fig. 6g), matching
418 our findings from ERH depletion during endoderm differentiation (Fig. 5f). We found that ERH
419 represses many genes not regulated by its interacting partners YTHDC1 and SAFB1, which
420 predominately repress different subsets of ERH targets (Fig. 6h). Taken together, the results
421 indicate that various RNA binding proteins function non-redundantly to recruit ERH to different
422 targets. Based upon these findings, we propose that ERH interactions with RNA binding proteins,

423 including YTHDC1 and SAFB1, facilitate cell-specific repression of genes and repeat elements
424 (Fig. 6i).

425

426 **Discussion**

427 We discovered that depletion of ERH reverts the H3K9me3 landscape in terminally differentiated
428 cells to an hESC-like state, permissive to activation of early embryonic genes. This surprising
429 result led us to investigate the role of ERH in establishing gene repression during the earliest
430 stages of mammalian lineage segregation, in maintaining repression of pluripotency, and enabling
431 germ layer specification. We found that ERH controls H3K9me3 for repression of naïve and
432 alternative lineage genes during the cell fate commitment of totipotent cells to ICM and
433 trophectoderm in mouse blastocysts as well as of hESCs to the endoderm and ectoderm cell
434 lineages, to enforce cell fate and prevent anomalous gene activation. Between the three systems
435 we employed, mouse preimplantation development, hESC differentiation, and fibroblast iPSC
436 reprogramming, we observe that ERH commonly repressed a shared set of naïve genes such as
437 *ZSCAN4* across all systems, indicating that ERH is involved in both the initial repression of this
438 early transcriptional program and in maintaining repression in terminally differentiated cells.
439 Remarkably, ERH is the sole apparent mechanism for repressing meiotic genes in *S. pombe,* is
440 conserved in repressing germ cell genes in mammals[2], and has expanded to foundationally
441 control diverse other types of developmental genes and stages.

442 In a developmental context, the failure to repress 2-cell, 8-cell and pluripotency genes
443 after ERH knockdown results in delays of specific-lineage gene expression, including primitive
444 endoderm genes in mouse blastocysts and various endoderm and ectoderm genes during hESC
445 differentiation. Furthermore, we demonstrated that ERH is necessary to maintain repression of
446 alternative lineage genes, as ERH depletion elicited precocious expression of genes typical of
447 later cell differentiation in mouse blastocysts and human hESCs, as well as heterochromatinized

genes during initial iPSC reprogramming. We observed that evolutionarily young transposable elements, including the human-specific *LTR7Ys*[31], are especially upregulated by ERH depletion during hESC differentiation and iPSC reprogramming. Much of the lineage specific expression changes upon ERH depletion during hESC differentiation may be attributed to increased accessibility of transcription factor binding sites among derepressed repeats[31,52,55].

A key finding is that ERH depletion in human fibroblasts reverted the H3K9me3 chromatin landscape to a hESC-like state. The reversion to a hESC-like H3K9me3 landscape permits binding of SOX2 in fibroblasts to endogenous hESC SOX2 targets, activating embryonic genes. That ERH is necessary for maintaining the broad H3K9me3 domains observed in somatic cells, but is dispensable for maintaining hESC H3K9me3 peaks, is consistent with our observation that ERH depletion in hESCs did not change global H3K9me3 levels. We suggest that the residual H3K9me3 peaks in pluripotent cells, enriched for zinc finger motifs and KAP1 binding, may be H3K9me3 nucleation sites and that, during differentiation, ERH is essential for the spread of H3K9me3 from these sites. Importantly, when ERH is knocked down in fibroblasts, we observe increased activation of many genes and transposable elements that harbor a residual H3K9me3 peak, yet lose broad H3K9me3, indicating that residual H3K9me3 peaks are not sufficient to block gene activation. Our knowledge of H3K9 methylation nucleation and spread has been primarily elucidated through studies in *S. pombe*[69–71], leaving further investigation needed into how H3K9me3 nucleation and spread occurs in multicellular organisms.

We identify that ERH interacts with and is partially recruited to the chromatin by the RNA binding protein YTHDC1 in somatic cells, but not in pluripotent hESCs. We find that ERH and YTHDC1 repress similar targets in fibroblasts and differentiating hESCs, but not in homeostatic pluripotent hESCs. In contrast, we observe that ERH interacts with the RNA binding protein SAFB1 in both hESCs and human fibroblasts and likewise both repress the satellite repeat expression in hESCs and fibroblasts. Our knockdowns of YTHDC1 and gene expression analysis of YTHDC1 and SAFB1 fail to fully explain how ERH is recruited to chromatin to repress

474 transcription. We thus speculate that ERH is recruited to cell-type specific genes and repeats
475 through a diverse network of chromatin-associated RNA binding factors that are yet to be fully
476 identified.

477 We observe in mouse development and hESC systems that ERH is functionally required
478 in multiple developmental contexts to repress meiotic and gametogenic genes, as it is in *S.*
479 *pombe*[1], although the gene members targeted by ERH has greatly evolved and expanded in
480 mammals. Our findings on meiotic gene repression are supported by the recent identification of
481 ERH in a screen for proteins silencing germline genes in mESCs[72]. Other aspects of *Erh1* function
482 in *S. pombe,* such as the targeting of RNAs for degradation by the RNA exosome[1], appear to be
483 conserved as we observed upregulation of developmental genes typically downregulated by RNA
484 exosome degradation[73]. ERH was recently found by mass spectrometry to associate with the
485 Microprocessor and NEXT complexes in mouse ESCs[74]. Evolutionarily divergent functions of
486 ERH may be driven by evolutionarily young transposable elements that retain RNA expression,
487 similar to how transposable elements have influenced the evolution of KRAB Zinc finger proteins
488 DNA dependent targeting of H3K9me3[75,76].

489 We found that ERH enacts gene repression at the earliest stages of development and that
490 it is essential for maintaining this repression in terminally differentiated cells. Our findings advance
491 the understanding of how heterochromatin-based repression is established and maintained
492 during development. We propose that understanding mechanisms by which ERH is employed to
493 direct H3K9me3 to specific loci in a developmentally coordinated fashion will be crucial to
494 modulating developmental H3K9me3 rearrangements to achieve greater control of cell identity.

495 **References**

496

497 1. Sugiyama, T. *et al.* Enhancer of Rudimentary Cooperates with Conserved RNA-Processing Factors to
498 Promote Meiotic mRNA Decay and Facultative Heterochromatin Assembly. *Mol Cell* **61**, 747–759
499 (2016).
500 2. McCarthy, R. L. *et al.* Diverse heterochromatin-associated proteins repress distinct classes of genes
501 and repetitive elements. *Nat Cell Biol* **23**, 905–914 (2021).
502 3. Nicetto, D. *et al.* H3K9me3-heterochromatin loss at protein-coding genes enables developmental
503 lineage specification. *Science* **363**, 294–297 (2019).
504 4. Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B. & Young, R. A. SetDB1 contributes to
505 repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes &*
506 *development* **23**, 2484–9 (2009).
507 5. Becker, J. S. *et al.* Genomic and Proteomic Resolution of Heterochromatin and Its Restriction of
508 Alternate Fate Genes. *Mol Cell* **68**, 1023-1037.e15 (2017).
509 6. Xu, R. *et al.* Stage-specific H3K9me3 occupancy ensures retrotransposon silencing in human pre-
510 implantation embryos. *Cell Stem Cell* **29**, 1051-1066.e8 (2022).
511 7. Zhang, J. *et al.* Distinct H3K9me3 heterochromatin maintenance dynamics govern different gene
512 programmes and repeats in pluripotent cells. *Nat Cell Biol* **26**, 2115–2128 (2024).
513 8. Schulz, K. N. & Harrison, M. M. Mechanisms regulating zygotic genome activation. *Nat Rev Genet* **20**,
514 221–234 (2019).
515 9. Xu, R. *et al.* Unreprogrammed H3K9me3 prevents minor zygotic genome activation and lineage
516 commitment in SCNT embryos. *Nat Commun* **14**, 4807 (2023).
517 10. Sankar, A. *et al.* KDM4A regulates the maternal-to-zygotic transition by protecting broad H3K4me3
518 domains from H3K9me3 invasion in oocytes. *Nat Cell Biol* **22**, 380–388 (2020).

519 11. Wang, C. *et al.* Reprogramming of H3K9me3-dependent heterochromatin during mammalian
520 embryo development. *Nature cell biology* **20**, 620–631 (2018).
521 12. Yu, H. *et al.* Dynamic reprogramming of H3K9me3 at hominoid-specific retrotransposons during
522 human preimplantation development. *Cell Stem Cell* **29**, 1031-1050.e12 (2022).
523 13. Burton, A. *et al.* Heterochromatin establishment during early mammalian development is regulated
524 by pericentromeric RNA and characterized by non-repressive H3K9me3. *Nat Cell Biol* **22**, 767–778
525 (2020).
526 14. Bhanu, N. V., Sidoli, S. & Garcia, B. A. Histone modification profiling reveals differential signatures
527 associated with human embryonic stem cell self-renewal and differentiation. *Proteomics* **16**, 448–
528 458 (2016).
529 15. Ugarte, F. *et al.* Progressive Chromatin Condensation and H3K9 Methylation Regulate the
530 Differentiation of Embryonic and Hematopoietic Stem Cells. *Stem Cell Reports* **5**, 728–740 (2015).
531 16. Xie, G. *et al.* A conserved dimer interface connects ERH and YTH family proteins to promote gene
532 silencing. *Nat Commun* **10**, 251 (2019).
533 17. Kuznetsov, D. *et al.* OrthoDB v11: annotation of orthologs in the widest sampling of organismal
534 diversity. *Nucleic Acids Res* **51**, D445–D451 (2023).
535 18. Stephens, A. D. *et al.* Chromatin histone modifications and rigidity affect nuclear morphology
536 independent of lamins. *Mol Biol Cell* **29**, 220–233 (2018).
537 19. Stoll, G. A., Pandiloski, N., Douse, C. H. & Modis, Y. Structure and functional mapping of the KRAB-
538 KAP1 repressor complex. *EMBO J* **41**, e111179 (2022).
539 20. Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency
540 reprogramming factors' initial engagement with the genome. *Cell* **151**, 994–1004 (2012).
541 21. Taing, L. *et al.* Cistrome Data Browser: integrated search, analysis and visualization of chromatin
542 data. *Nucleic Acids Res* **52**, D61–D66 (2024).

543 22. Zheng, R. *et al.* Cistrome Data Browser: expanded datasets and new tools for gene regulatory
544 analysis. *Nucleic Acids Res* **47**, D729–D735 (2019).
545 23. Yu, X. *et al.* Recapitulating early human development with 8C-like cells. *Cell Rep* **39**, 110994 (2022).
546 24. Zou, Z. *et al.* Translatome and transcriptome co-profiling reveals a role of TPRXs in human zygotic
547 genome activation. *Science* **378**, abo7923 (2022).
548 25. Taubenschmid-Stowers, J. *et al.* 8C-like cells capture the human zygotic genome activation program
549 in vitro. *Cell Stem Cell* **29**, 449-459.e6 (2022).
550 26. Cacchiarelli, D. *et al.* Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of
551 Induced Pluripotency. *Cell* **162**, 412–424 (2015).
552 27. Bi, Y. *et al.* Cell fate roadmap of human primed-to-naive transition reveals preimplantation cell
553 lineage signatures. *Nat Commun* **13**, 3147 (2022).
554 28. Xue, Z. *et al.* Genetic programs in human and mouse early embryos revealed by single-cell RNA
555 sequencing. *Nature* **500**, 593–597 (2013).
556 29. Wang, J. *et al.* Primate-specific endogenous retrovirus-driven transcription defines naive-like stem
557 cells. *Nature* **516**, 405–409 (2014).
558 30. Lu, X. *et al.* The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell
559 identity. *Nat Struct Mol Biol* **21**, 423–425 (2014).
560 31. Ai, Z. *et al.* Krüppel-like factor 5 rewires NANOG regulatory network to activate human naive
561 pluripotency specific LTR7Ys and promote naive pluripotency. *Cell Rep* **40**, 111240 (2022).
562 32. Gadepalli, V. S., Kim, H., Liu, Y., Han, T. & Cheng, L. XDeathDB: a visualization platform for cell death
563 molecular interactions. *Cell Death Dis* **12**, 1156 (2021).
564 33. Gao, Y. *et al.* Protein Expression Landscape of Mouse Embryos during Pre-implantation
565 Development. *Cell reports* **21**, 3957–3969 (2017).

566 34. Ralston, A. & Rossant, J. Cdx2 acts downstream of cell polarization to cell-autonomously promote
567 trophectoderm fate in the early mouse embryo. *Dev Biol* **313**, 614–629 (2008).
568 35. Lim, H. Y. G. *et al.* Keratins are asymmetrically inherited fate determinants in the mammalian
569 embryo. *Nature* **585**, 404–409 (2020).
570 36. Domingo-Muelas, A. *et al.* Human embryo live imaging reveals nuclear DNA shedding during
571 blastocyst expansion and biopsy. *Cell* **186**, 3166-3181.e18 (2023).
572 37. Falco, G. *et al.* Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic
573 stem cells. *Dev Biol* **307**, 539–550 (2007).
574 38. Hamatani, T., Carter, M. G., Sharov, A. A. & Ko, M. S. H. Dynamics of global gene expression changes
575 during mouse preimplantation development. *Dev Cell* **6**, 117–131 (2004).
576 39. Boussicault, L. *et al.* CYP46A1, the rate-limiting enzyme for cholesterol degradation, is
577 neuroprotective in Huntington's disease. *Brain* **139**, 953–970 (2016).
578 40. Chalhoub, G. *et al.* Carboxylesterase 2 proteins are efficient diglyceride and monoglyceride lipases
579 possibly implicated in metabolic disease. *J Lipid Res* **62**, 100075 (2021).
580 41. Artus, J., Piliszek, A. & Hadjantonakis, A.-K. The primitive endoderm lineage of the mouse blastocyst:
581 sequential transcription factor activation and regulation of differentiation by Sox17. *Dev Biol* **350**,
582 393–404 (2011).
583 42. Teo, A. K. K., Valdez, I. A., Dirice, E. & Kulkarni, R. N. Comparable generation of activin-induced
584 definitive endoderm via additive Wnt or BMP signaling in absence of serum. *Stem Cell Reports* **3**, 5–
585 14 (2014).
586 43. Messmer, T. *et al.* Transcriptional Heterogeneity in Naive and Primed Human Pluripotent Stem Cells
587 at Single-Cell Resolution. *Cell Rep* **26**, 815-824.e4 (2019).
588 44. Dubois, A. *et al.* H3K9 tri-methylation at Nanog times differentiation commitment and enables the
589 acquisition of primitive endoderm fate. *Development* **149**, dev201074 (2022).

590 45. Martello, G. *et al.* Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-
591 renewal. *Cell Stem Cell* **11**, 491–504 (2012).
592 46. Gassler, J. *et al.* Zygotic genome activation by the totipotency pioneer factor Nr5a2. *Science* **378**,
593 1305–1315 (2022).
594 47. Maskalenka, K. *et al.* NANOGP1, a tandem duplicate of NANOG, exhibits partial functional
595 conservation in human naïve pluripotent stem cells. *Development* **150**, dev201155 (2023).
596 48. Domcke, S. *et al.* A human cell atlas of fetal chromatin accessibility. *Science* **370**, eaba7612 (2020).
597 49. Evseenko, D. *et al.* Mapping the first stages of mesoderm commitment during differentiation of
598 human embryonic stem cells. *Proc Natl Acad Sci U S A* **107**, 13742–13747 (2010).
599 50. Bernardo, A. S. *et al.* BRACHYURY and CDX2 mediate BMP-induced differentiation of human and
600 mouse pluripotent stem cells into embryonic and extraembryonic lineages. *Cell Stem Cell* **9**, 144–
601 155 (2011).
602 51. Allshire, R. C. & Madhani, H. D. Ten principles of heterochromatin formation and function. *Nat. Rev.*
603 *Mol. Cell Biol.* **19**, 229–244 (2018).
604 52. Pontis, J. *et al.* Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic
605 Genome Activation and Control Transcription in Naive Human ESCs. *Cell Stem Cell* **24**, 724-735.e5
606 (2019).
607 53. Wu, F. *et al.* Species-specific rewiring of definitive endoderm developmental gene activation via
608 endogenous retroviruses through TET1-mediated demethylation. *Cell Rep* **41**, 111791 (2022).
609 54. Macfarlan, T. S. *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus activity.
610 *Nature* **487**, 57–63 (2012).
611 55. Pontis, J. *et al.* Primate-specific transposable elements shape transcriptional networks during
612 human development. *Nat Commun* **13**, 7178 (2022).

613 56. Wang, H. *et al.* SVA elements: a hominid-specific retroposon family. *J Mol Biol* **354**, 994–1007
614 (2005).
615 57. Grow, E. J. *et al.* Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent
616 cells. *Nature* **522**, 221–5 (2015).
617 58. Theunissen, T. W. *et al.* Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem*
618 *Cell* **19**, 502–515 (2016).
619 59. Khan, H., Smit, A. & Boissinot, S. Molecular evolution and tempo of amplification of human LINE-1
620 retrotransposons since the origin of primates. *Genome Res* **16**, 78–87 (2006).
621 60. Hazra, D., Andrić, V., Palancade, B., Rougemaille, M. & Graille, M. Formation of S. pombe Erh1
622 homodimer mediates gametogenic gene silencing and meiosis progression. *Sci Rep* **10**, 1034 (2020).
623 61. Hazra, D., Chapat, C. & Graille, M. m^6A mRNA Destiny: Chained to the rhYTHm by the YTH-
624 Containing Proteins. *Genes (Basel)* **10**, E49 (2019).
625 62. Liao, S., Sun, H. & Xu, C. YTH Domain: A Family of N6-methyladenosine (m6A) Readers. *Genomics*
626 *Proteomics Bioinformatics* **16**, 99–107 (2018).
627 63. Liu, J. *et al.* The RNA m(6)A reader YTHDC1 silences retrotransposons and guards ES cell identity.
628 *Nature* **591**, 322–326 (2021).
629 64. Sun, T. *et al.* Crosstalk between RNA m6A and DNA methylation regulates transposable element
630 chromatin activation and cell fate in human pluripotent stem cells. *Nat Genet* **55**, 1324–1335 (2023).
631 65. Drakouli, S., Lyberopoulou, A., Papathanassiou, M., Mylonis, I. & Georgatsou, E. Enhancer of
632 rudimentary homologue interacts with scaffold attachment factor B at the nuclear matrix to
633 regulate SR protein phosphorylation. *FEBS J* **284**, 2482–2500 (2017).
634 66. Maison, C. *et al.* Higher-order structure in pericentric heterochromatin involves a distinct pattern of
635 histone modification and an RNA component. *Nature genetics* **30**, 329–34 (2002).

636 67. Xu, W. *et al.* METTL3 regulates heterochromatin in mouse embryonic stem cells. *Nature* **591**, 317–
637 321 (2021).
638 68. Huo, X. *et al.* The Nuclear Matrix Protein SAFB Cooperates with Major Satellite RNAs to Stabilize
639 Heterochromatin Architecture Partially through Phase Separation. *Mol Cell* **77**, 368-383 e7 (2020).
640 69. Obersriebnig, M. J., Pallesen, E. M. H., Sneppen, K., Trusina, A. & Thon, G. Nucleation and spreading
641 of a heterochromatic domain in fission yeast. *Nat Commun* **7**, 11518 (2016).
642 70. Murawska, M. *et al.* The histone chaperone FACT facilitates heterochromatin spreading by
643 regulating histone turnover and H3K9 methylation states. *Cell Rep* **37**, 109944 (2021).
644 71. Greenstein, R. A. *et al.* Noncoding RNA-nucleated heterochromatin spreading is intrinsically labile
645 and requires accessory elements for epigenetic stability. *Elife* **7**, (2018).
646 72. Al Adhami, H. *et al.* Systematic identification of factors involved in the silencing of germline genes in
647 mouse embryonic stem cells. *Nucleic Acids Res* gkad071 (2023) doi:10.1093/nar/gkad071.
648 73. Wu, D. & Dean, J. RNA exosome ribonuclease DIS3 degrades Pou6f1 to promote mouse pre-
649 implantation cell differentiation. *Cell Rep* **42**, 112047 (2023).
650 74. Imamura, K. *et al.* A functional connection between the Microprocessor and a variant NEXT
651 complex. *Mol Cell* **84**, 4158-4174.e6 (2024).
652 75. Yang, P., Wang, Y. & Macfarlan, T. S. The Role of KRAB-ZFPs in Transposable Element Repression and
653 Mammalian Evolution. *Trends Genet* **33**, 871–881 (2017).
654 76. Imbeault, M., Helleboid, P.-Y. & Trono, D. KRAB zinc-finger proteins contribute to the evolution of
655 gene regulatory networks. *Nature* **543**, 550–554 (2017).
656 77. Cao, J. *et al.* A human cell atlas of fetal gene expression. *Science* **370**, eaba7721 (2020).
657 78. Jin, Y., Tam, O. H., Paniagua, E. & Hammell, M. TEtranscripts: a package for including transposable
658 elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **31**, 3593–3599
659 (2015).

660 79. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of
661 DNA binding sites. *Elife* **6**, e21856 (2017).
662 80. Katznelson, A. & Zaret, K. A fluorescence-based protocol to quantitatively titrate CUT&RUN buffer
663 components. *STAR Protoc* **5**, 102866 (2024).
664 81. Liu, N. *et al.* Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch.
665 *Cell* **173**, 430-442.e17 (2018).
666 82. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359
667 (2012).
668 83. Jin, H. *et al.* ChIPseqSpikeInFree: a ChIP-seq normalization approach to reveal global changes in
669 histone modifications without spike-in. *Bioinformatics* **36**, 1270–1272 (2020).
670 84. Meers, M. P., Tenenbaum, D. & Henikoff, S. Peak calling by Sparse Enrichment Analysis for
671 CUT&RUN chromatin profiling. *Epigenetics Chromatin* **12**, 42 (2019).
672 85. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-
673 regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**, 576–89
674 (2010).
675 86. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis.
676 *Nucleic Acids Res* **44**, W160-165 (2016).
677 87. Khan, A. & Mathelier, A. Intervene: a tool for intersection and visualization of multiple gene or
678 genomic region sets. *BMC Bioinformatics* **18**, 287 (2017).
679 88. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
680 89. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning
681 sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
682 90. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq
683 data with DESeq2. *Genome Biol* **15**, 550 (2014).

684 91. Filippi-Chiela, E. C. *et al.* Nuclear morphometric analysis (NMA): screening of senescence, apoptosis
685 and nuclear irregularities. *PLoS One* **7**, e42522 (2012).
686 92. Thomas, P. D. *et al.* PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci* **31**,
687 8–22 (2022).
688 93. Lau, K. X. *et al.* Unique properties of a subset of human pluripotent stem cells with high capacity for
689 self-renewal. *Nat Commun* **11**, 2420 (2020).
690

691 **Figure legends**

692 Figure 1. ERH depletion reverts fibroblast H3K9me3 patterns to an hESC-like state

693 **a.** Schematic of siRNA treatments and timing for H3K9me3 CUT&RUN experiments shown
694 in figure 1 and SOX2 CUT&RUN and RNA-seq experiments shown in figure 2.

695 **b.** Immunofluorescence of ERH (red), H3K9me3 (red) and DAPI (blue) in siControl and
696 siERH human fibroblasts. Scale bars, 50μm.

697 **c.** Browser track of H3K9me3 levels in siControl and siERH treated human fibroblasts and
698 H1 hESCs (GSM1003585) at the megabase-domain level and gene level. Example
699 H3K9me3 peaks common between siERH fibroblasts and H1 hESCs are indicated by red
700 triangles.

701 **d.** Heatmap of fibroblast siControl and siERH H3K9me3, H1 hESC H3K9me3
702 (GSM1003585) and H1 hESC HP1α (GSE123229) levels centered at siERH retained
703 H3K9me3 peaks located within fibroblast H3K9me3 domains +/-10kb.

704 **e.** Correlation of H3K9me3 from BJ fibroblasts after siControl or siERH with H1 hESC
705 H3K9me3.

706 **f.** Representative images showing hESC morphology and SOX2 (green) pluripotency
707 marker staining in siControl and siERH H1 hESCs under pluripotency maintaining
708 conditions. Scale bars, 50μm.

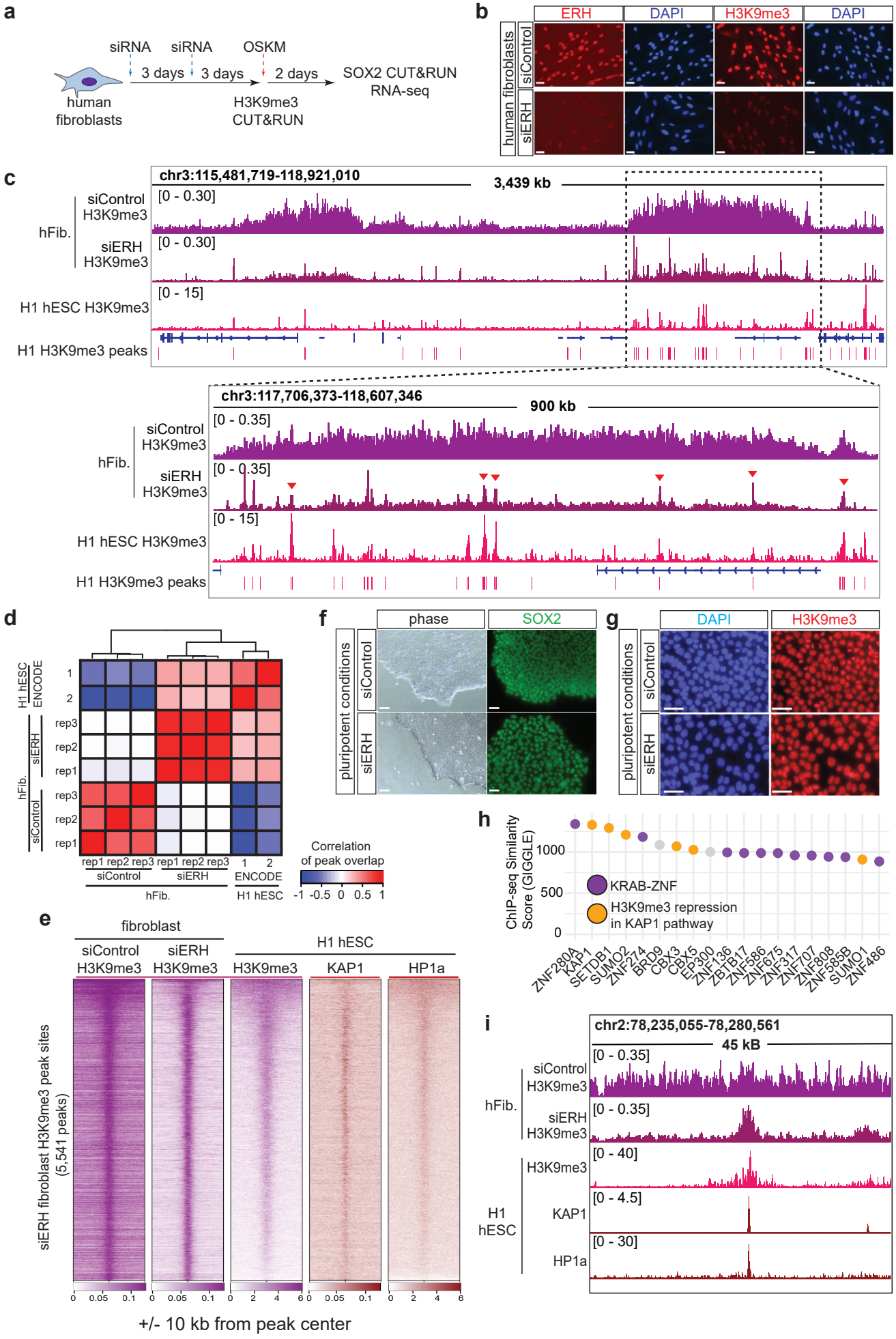
709 **g.** Representative images of H3K9me3 (red) and DAPI (blue) in siControl and siERH H1
710 hESCs under pluripotency maintaining conditions. Scale bars, 50μm.

711 **h.** CistromeDB Toolkit enrichment analysis of fibroblast siERH H3K9me3 peaks. The top 20
712 proteins scored by the analysis are visualized. Proteins colored by characterization as
713 H3K9me3-associated (orange), KRAB zinc-finger transcription factors (purple), and
714 euchromatin-associated (gray).

715 **i.** Browser track comparing siControl and siERH fibroblast H3K9me3 with H1 hESC
716 H3K9me3 (GSM1003585) and HP1a (GSM1701825) patterns. Peak locations called by
717 SEACR analysis on siERH H3K9me3 data are shown.

718 See also Extended Data Fig. 1 and Supplementary Table 1.

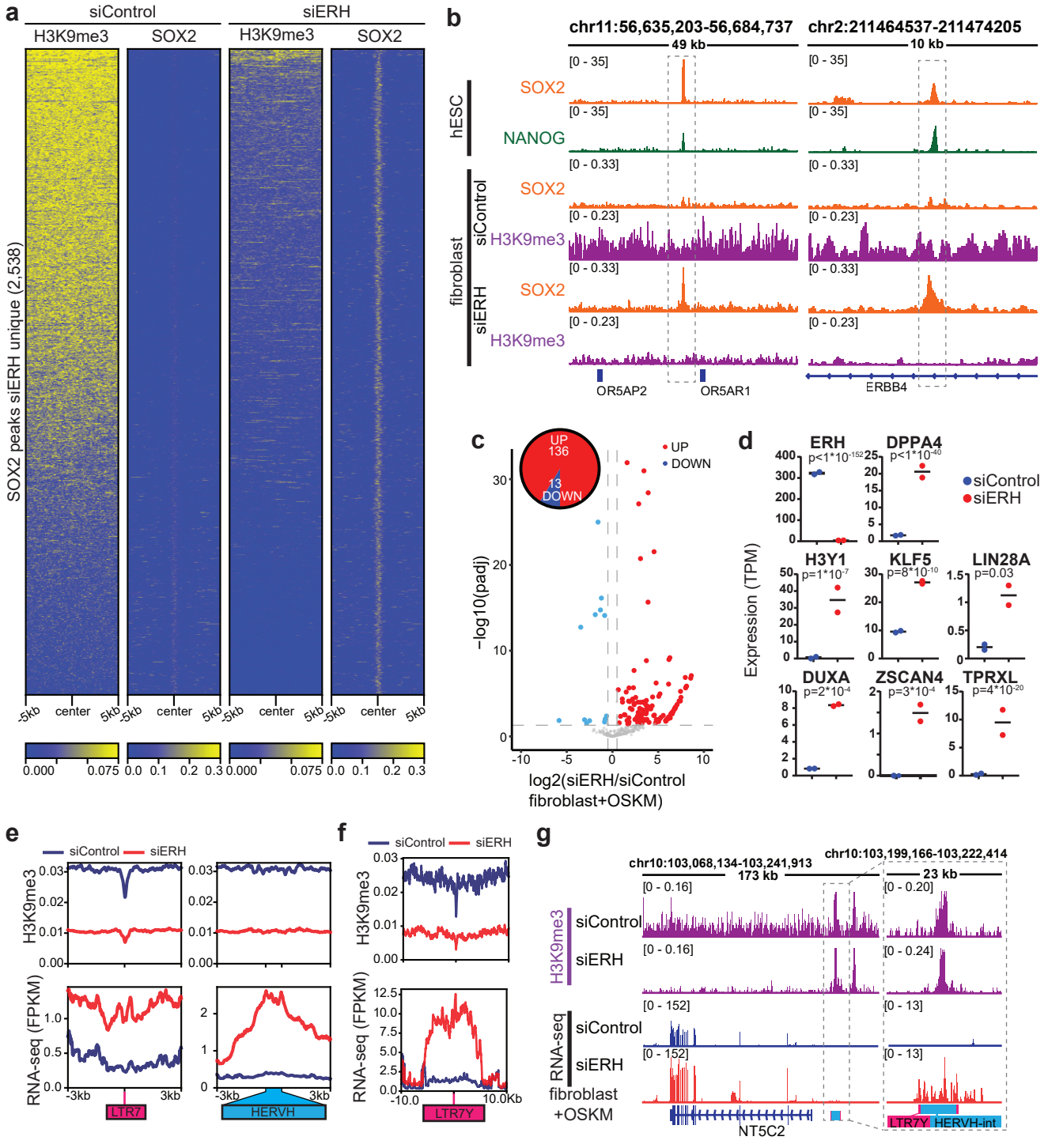
Katznelson et al., Figure 1



719 Figure 2. ERH depletion enhances gene activation of pluripotency and alternative lineage genes
720 during iPSC reprogramming

721 **a.** Heatmap of H3K9me3 CUT&RUN from siControl and siERH human fibroblasts and SOX2
722 CUT&RUN from siControl and siERH fibroblasts + OSKM at siERH-gained SOX2 peaks.
723 **b.** Browser track of example hESC SOX2 and NANOG peak coincident with fibroblast siERH
724 H3K9me3 loss and SOX2 peak gain.
725 **c.** Volcano plot of expression changes (DESeq2) in siERH vs siControl fibroblasts+OSKM of
726 iPSC activated genes located in fibroblast H3K9me3 domains.
727 **d.** Expression of ERH and example naïve pluripotency genes activated in 48 hours OSKM
728 by ERH depletion, transcripts per million (TPM) values from RNA-seq replicates shown
729 (DESeq2 adjusted p-value, n=2).
730 **e.** Profile plots of H3K9me3 and RNA levels from LTR7 and HERVH repeats in siERH vs
731 siControl fibroblasts+OSKM.
732 **f.** Profile plots of H3K9me3 and RNA levels from LTR7Y repeats in siERH vs siControl
733 fibroblasts+OSKM.
734 **g.** Browser track showing H3K9me3 in siControl and siERH fibroblasts and RNA levels in
735 siControl+OSKM and siERH+OSKM for gene and adjacent LTR7Y naïve pluripotency
736 repeat.
737 See also Extended Data Fig. 2-4 and Supplementary Tables 2-4.

Katznelson et al., Figure 2



738 Figure 3. ERH is expressed during pre-implantation mouse development and is critical for normal
739 blastocyst development by repressing 2-cell and alternative lineage genes.

740 **a.** ERH (white) and DAPI (blue) staining during mouse pre-implantation development.
741 Representative confocal z-sections are shown (zygote n=15, N=3, 2-cell stage n=18, N=3,
742 8-cell stage n=6, N=2, blastocyst n=10, N=3). Scale bars, 10μm.

743 **b.** Representative images of trophectoderm marker Cdx2 (green) in siControl and siERH
744 blastocysts with Phalloidin (orange) marking cell boundaries. Scale bars, 10μm.

745 **c.** Quantification of CDX2 intensity in siControl vs siERH blastocysts (unpaired t-test, n=8
746 blastocysts for each siRNA).

747 **d.** Representative image of trophectoderm marker KRT8 (orange) in siControl and siERH
748 blastocysts with DAPI (blue). Scale bars, 10μm.

749 **e.** Quantification of KRT8 intensity in siControl vs siERH blastocysts (unpaired t-test, n=9
750 blastocysts, n=8 blastocysts for each siRNA).

751 **f.** Representative images of H3K9me3 (magenta), DAPI (teal) in siControl vs siERH
752 blastocysts. Scale bars, 10μm.

753 **g.** Quantification of H3K9me3 level in siControl vs siERH blastocysts. (unpaired t-test, n=8
754 for siControl and n=9 for siERH blastocysts).

755 **h.** Percentage of cells per blastocyst with abnormal nuclei. (unpaired t-tes, n=11 for siControl
756 and n=9 for siERH blastocysts).

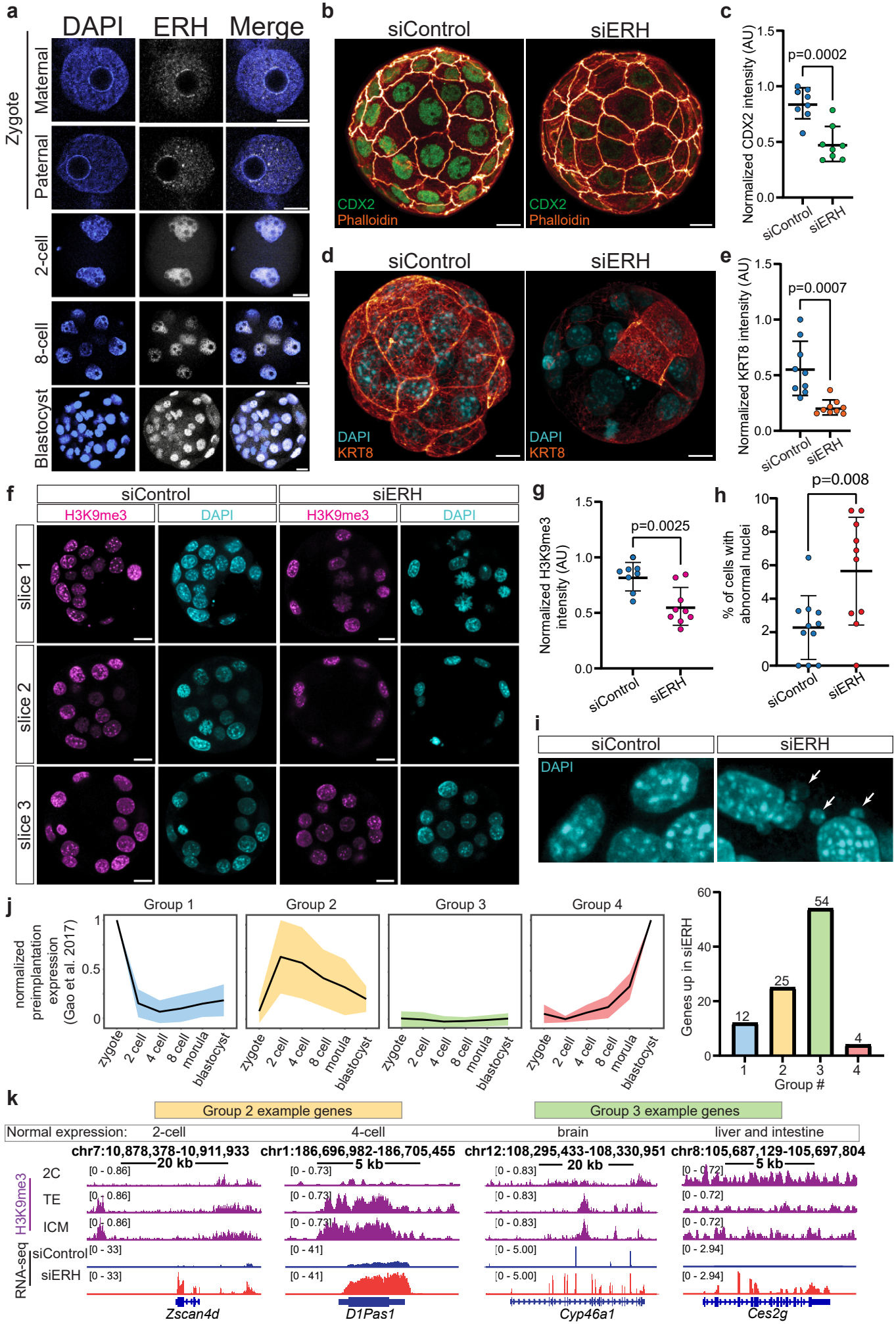
757 **i.** Representative image of DAPI staining siCTRL and siERH blastocysts with white arrows
758 indicating DNA shedding.

759 **j.** ERH repressed blastocyst genes (adjusted p-value<=0.05, log2fc>=0.5) grouped based
760 upon normal expression patterns exhibited during mouse pre-implantation development
761 re-analyzed from Gao et al., 2017[33].

762 **k.** Example gene tracks of genes significantly upregulated in siERH vs siControl blastocysts
763 (adjusted p-value<=0.05, log2fc>=0.5, n=3) with H3K9me3 tracks for 2 cell, trophectoderm
764 and ICM cells from[11].

765 See also Extended Data 5 and Supplementary Tables 5,6.

Katznelson et al., Figure 3



766 Figure 4. ERH represses pluripotency and alternative lineage gene activation during hESC
767 differentiation

768 **a.** Representative images of NANOG (red), SOX17 (red) and DAPI (blue) during 1 day of
769 endoderm differentiation of siControl and siERH H1 hESCs. Scale bars, 50μm.

770 **b.** Representative images of OCT4 (red), PAX6 (green) and DAPI (blue) during 2 days of
771 ectoderm differentiation of siControl and siERH H1 hESCs. Scale bars, 50μm.

772 **c.** Plot of NANOG expression in transcripts per million (TPM) by RNA-seq in siControl and
773 siERH treated H1 hESCs cultured in pluripotency (H1), 1 day endoderm differentiation
774 (DE) and 2 day ectoderm differentiation (EC). (DESeq2 adjusted p-value reported).

775 **d.** Plot of KLF5 and ZSCAN4 expression in transcripts per million (TPM) by RNA-seq in
776 siControl and siERH treated H1 hESCs cultured in pluripotency (H1), 1 day endoderm
777 differentiation (DE) and 2 day ectoderm differentiation (EC). (DESeq2 adjusted p-value
778 reported).

779 **e.** GO analysis (PantherDB, statistical overrepresentation) of genes upregulated
780 (padj<=0.05, log2 fold change >= 0.5) in siERH vs siControl H1 hESCs 1 day endoderm
781 differentiation.

782 **f.** Radar plot showing cumulative expression profile across indicated human tissues, data
783 downloaded from BBI Descartes database[77], for genes upregulated (padj<=0.05, log2 fold
784 change >= 0.5) in siERH vs siControl H1 hESCs 1 day endoderm differentiation.

785 **g.** Radar plot showing cumulative expression profile across indicated human tissues, data
786 downloaded from BBI Descartes database[77], for genes upregulated (padj<=0.05, log2 fold
787 change >= 0.5) in siERH vs siControl H1 hESCs 2 day ectoderm differentiation.

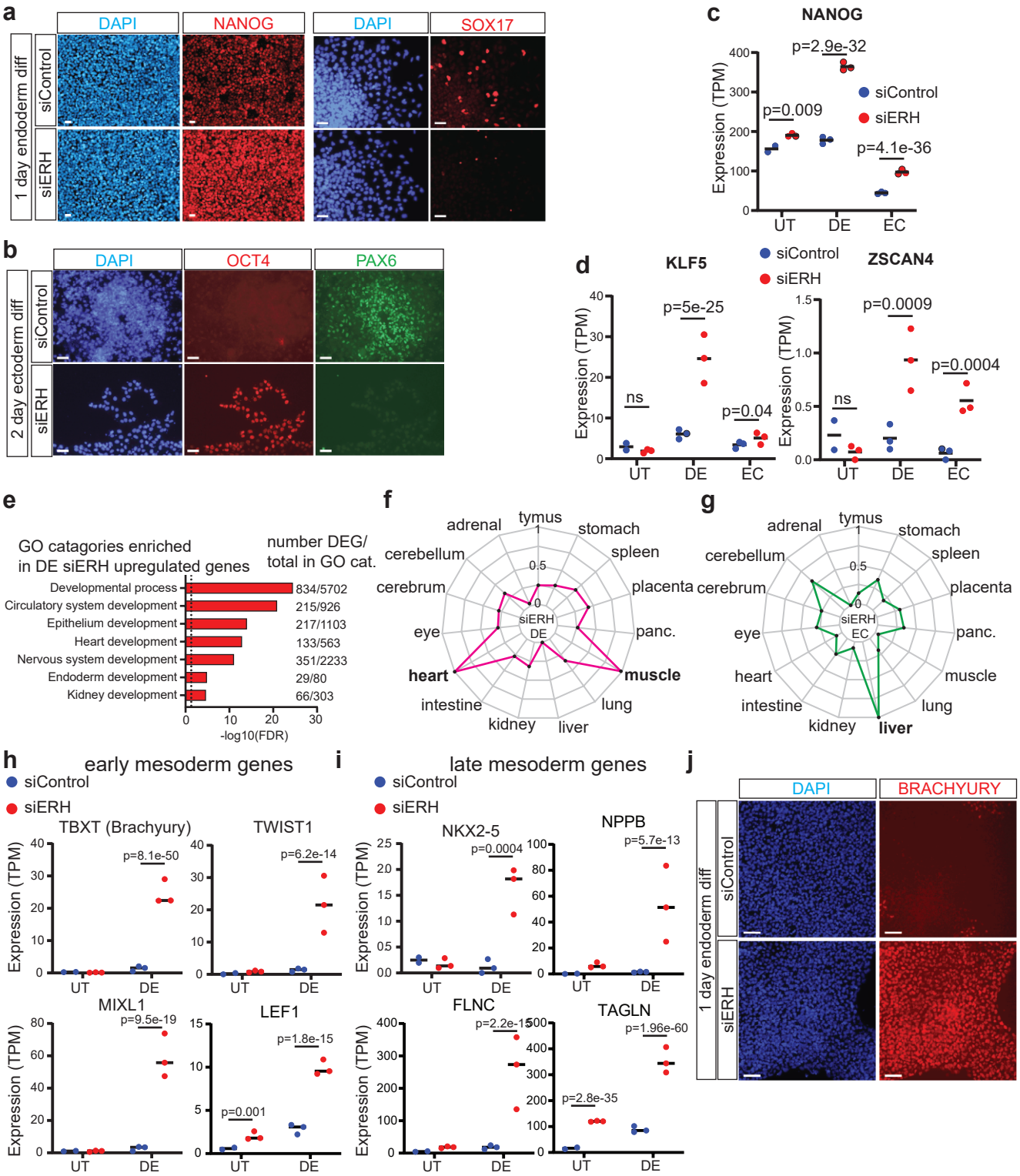
788 **h.** Example early mesodermal genes activated during endoderm differentiation by ERH
789 depletion, transcripts per million (TPM) values from RNA-seq replicates shown (DESeq2
790 adjusted p-value).

791 **i.** Example late mesodermal genes activated during endoderm differentiation by ERH
792 depletion, transcripts per million (TPM) values from RNA-seq replicates shown DESeq2
793 adjusted p-value).

794 **j.** Representative image of DAPI (blue), BRACHYURY (red) of siControl and siERH treated
795 H1 hESCs after 1 day of endoderm differentiation. Scale bars, 50μm.

796 For all RNA-seq TPM values n=3, except for siCTRL UT where n=2. See also Extended
797 Data Fig. 6 and Supplementary Tables 7,8.

Katznelson et al., Figure 4



798 Figure 5. ERH represses evolutionarily young transposable elements with regulatory function

799 **a.** H3K9me3 data from H1 hESCs (GSM1003585) and hESC derived endoderm
800 (GSM916057) at 65,656 L1 element loci +/- 1kb upregulated (adjusted pvalue<=0.05;
801 log2foldchange>=0.5) by ERH depletion during 1 day endoderm differentiation.

802 **b.** Venn diagram showing intersect of repeat classes upregulated (TETranscripts, padj<0.05,
803 log2fc>0.5) by ERH depletion in H1 hESCs untreated (UT), 1 day endoderm (DE) or 2 day
804 ectoderm (EC) conditions.

805 **c.** Expression change and significance, TEtranscripts[78] analysis, of L1 LTR transposable
806 elements in ERH depletion in H1 hESCs untreated (UT), 1 day endoderm (DE) or 2 day
807 ectoderm (EC) listed in order of evolutionary age as determined in[59].

808 **d.** Example browser track showing expression of L1 element and proximal gene in siControl
809 and siERH treated H1 hESCs in the pluripotent and 1 day DE conditions.

810 **e.** Example browser track of gene with promoter proximal repeat element gaining H3K9me3
811 during endoderm differentiation that is upregulated by ERH depletion.

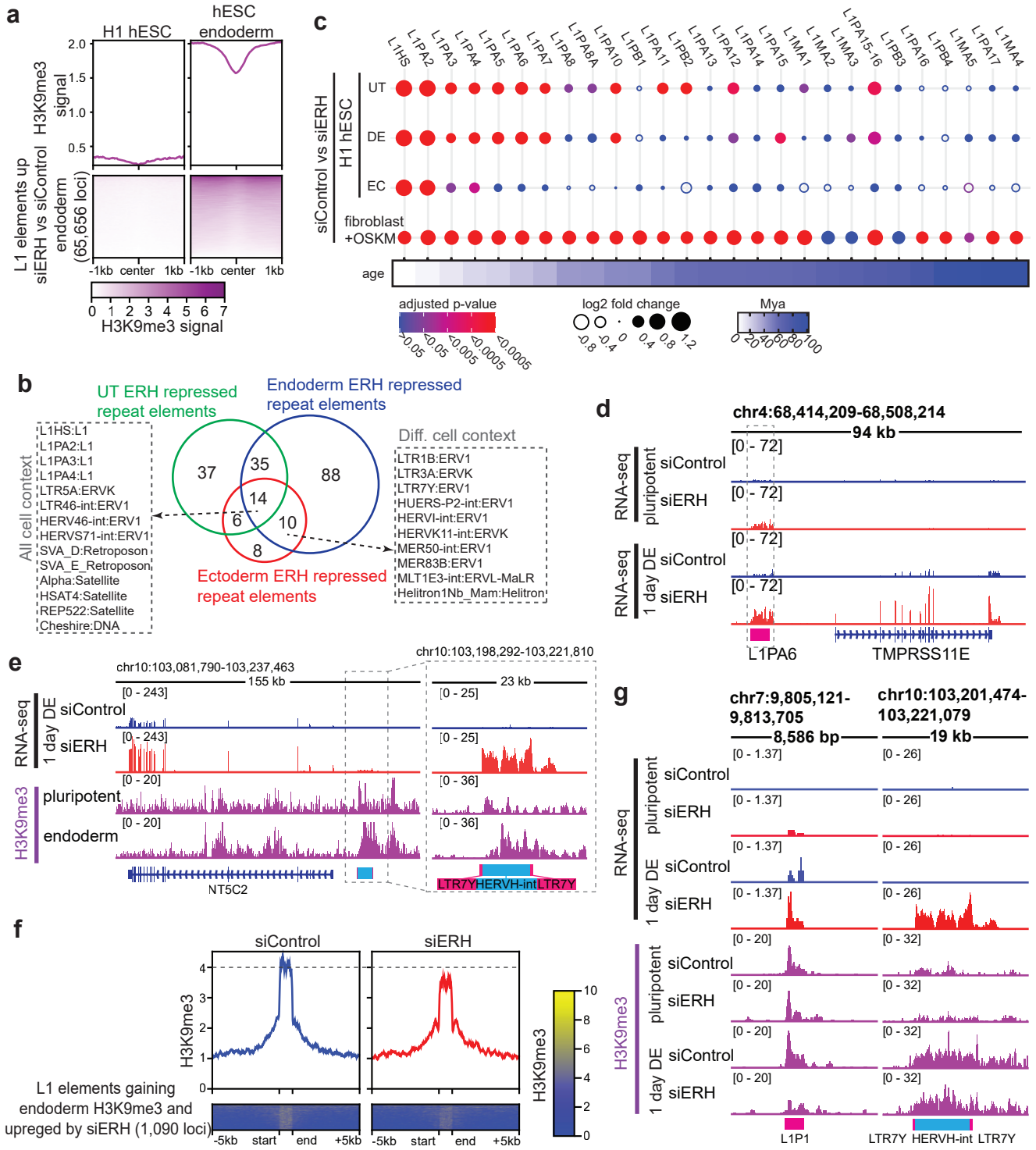
812 **f.** H3K9me3 CUT&RUN signal at L1 elements that exhibited 2-fold RNA increase in siERH
813 vs siControl H1 hESC endoderm and switched from H3K9me3 unmarked in pluripotent to
814 marked in endoderm.

815 **g.** Browser track showing RNA-seq of example repeat element and nearby genes in siControl
816 and siERH treated H1 hESCs after 1 day of endoderm differentiation and H3K9me3 tracks
817 from H1 hESCs (GSM1003585) and hESC derived endoderm (GSM916057).

818 See also Extended Data Fig. 7,8 and Supplementary Table 10.

819

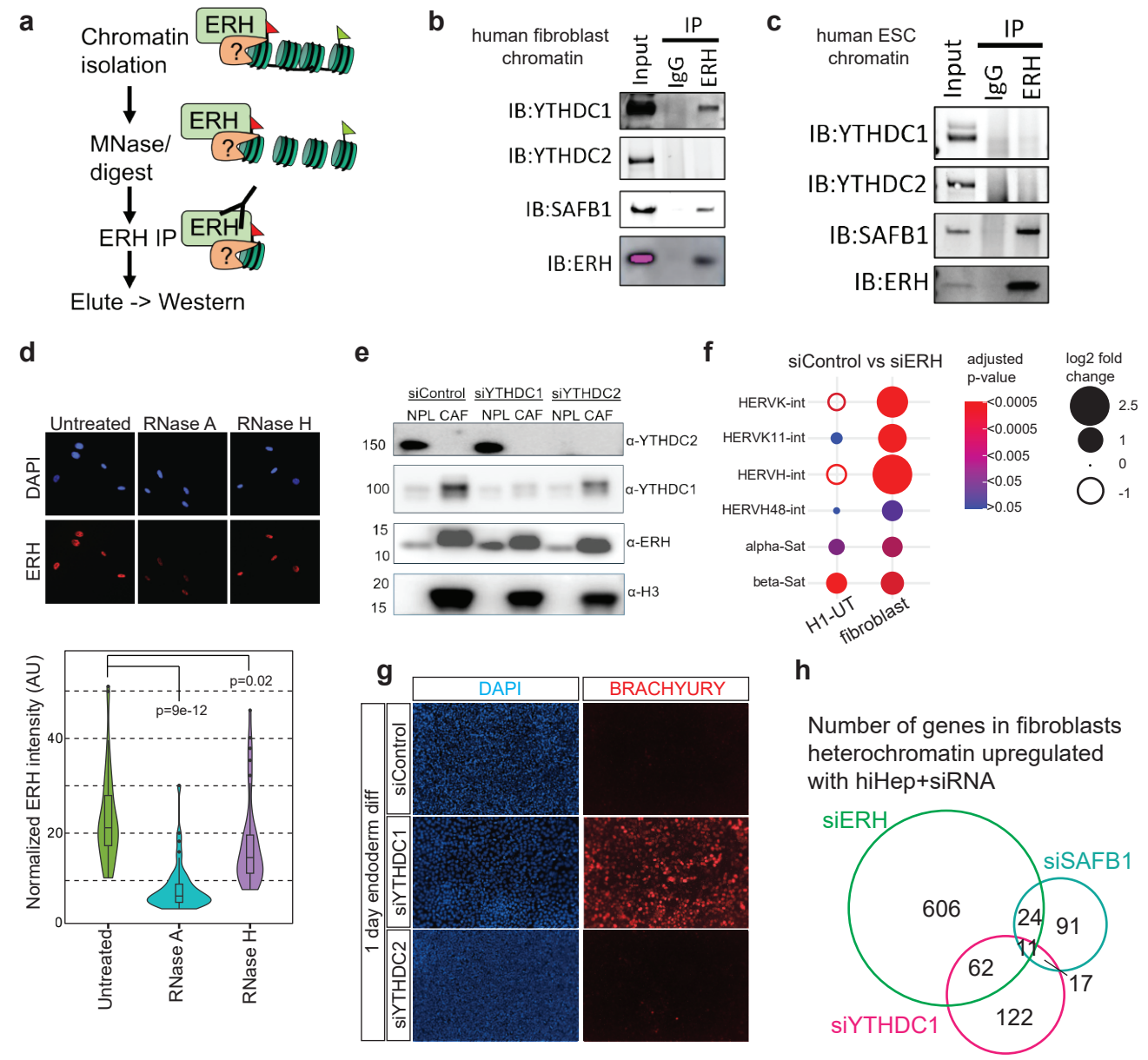
Katznelson et al., Figure 5



820 Figure 6. Cell type specific and independent recruitment of ERH to chromatin by RNA binding
821 proteins regulates gene and repeat repression

822 **a.** Strategy for identifying chromatin-bound ERH interaction partners by isolation of soluble
823 nucleoplasm and chromatin-associated protein fractions prior to immunoprecipitation.
824 **b.** Immunoprecipitation (IP) of ERH from the chromatin fraction of human fibroblasts.
825 Western blot analysis of YTHDC1, YTHDC2, SAFB1, and ERH in the input and IP
826 samples.
827 **c.** Immunoprecipitation (IP) of ERH from the chromatin fraction of human ESCs. Western
828 blot analysis of YTHDC1, YTHDC2, SAFB1, and ERH in the input and IP samples.
829 **d.** Immunofluorescence (IF) of ERH after RNase A and RNase H treatment.
830 **e.** Western blot of ERH, YTHDC1, and YTHDC2 in nucleoplasm and chromatin-associated
831 protein fractions of human fibroblasts after siYTHDC1 and siYTHDC2 treatment.
832 **f.** Expression change and significance, TEtranscripts[78] analysis, of HERVK HERVH and
833 satellite repeats in hESCs and human fibroblasts after siERH.
834 **g.** Immunofluorescence (IF) of BRACHYURY in siCTRL, siYTHDC1 and siYTHDC2 after 1
835 day endoderm differentiation.
836 **h.** Venn diagram of gene coregulation after siERH, siYTHDC1 and siSAFB1 in human
837 fibroblasts after addition of hepatic reprogramming factors, FOXA3, HNF1A, HNF4A.
838 **i.** Proposed model for cell-type specific, and cell-type agnostic gene and repeat regulation
839 by ERH, YTHDC1, and SAFB1.
840 See also Extended Data Fig. 9

Katznelson et al., Figure 6



841 **Methods**

842 **Cell lines culture and condition**

843 Human H1 ES cells were obtained from WiCell (WA01), grown on Matrigel coated wells and
844 cultured in mTeSR1 (Stem Cell Tech, 85850) medium at 37°C and 5% CO_2. Medium was
845 changed daily. Human BJ foreskin fibroblasts were obtained from ATCC (CRL-2522) and
846 cultured in Eagle's Minimum Essential Medium (EMEM, Sigma-Aldrich M2279) supplemented
847 with 10% fetal bovine serum (FBS, Hyclone SH30071) and 2mM Glutamax (ThermoFisher
848 35050061) at 37°C and 5% CO_2.

849

850 **Mouse embryo work**

851 Mouse embryo experimentation was approved by the Biological Resource Center Institutional
852 Animal Care and Use Committee (IACUC), Agency for Science, Technology and Research
853 (IACUC Protocol 806983). C57BL/6 wild-type 3–4-week-old female mice were superovulated
854 using 5 IU of pregnant mare serum (PMS, National Hormone and Peptide Program)
855 gonadotropin given intraperitoneally and 5 IU of recombinant chorionic gonadotrophin (CG,
856 National Hormone and Peptide Program) given 48 h after and immediately before mating,
857 according to animal ethics guidelines of the University of Pennsylvania. Embryos were flushed
858 from oviducts of plugged females using M2 medium (Merck) and cultured in KSOM+AA (Merck)
859 covered by mineral oil (Sigma, M-5310), at 37 °C and 5% CO2.

860

861 **Mouse embryo immunofluorescence and microscopy**

862 Embryos were fixed in 4% paraformaldehyde (Electron Microscopy Sciences, 15710) for 30 min
863 at room temperature washed twice in PBS with 0.1% Triton X-100 (Millipore Sigma, T8787),
864 permeabilized for 20 min in PBS with 0.5% Triton X-100, and blocked in 3% bovine serum
865 albumin (Gemini Bio, 700-100P) for 30 min. Embryos were then incubated at 4 °C overnight in
866 primary antibodies diluted in blocking solution at the following concentrations: ERH (Abcam,

867 ab166620) at 1:200, Keratin-8 (DSHB, THROMA-I) at 1:20, H3K9me3 (Abcam, ab8898) at
868 1:1000, and CDX2 (Invitrogen, ZC007) at 1:200. After primary antibody incubation, embryos
869 were washed 5 times for 20 min in PBS with 0.1% Triton X-100 and incubated for 30 min in
870 Alexa Fluor-conjugated secondary antibodies (Invitrogen, A32731, A32723, A32733, A32728
871 and A11006) diluted in blocking solution to 1:500. Phalloidin-Alexa Fluor 555 (Invitrogen,
872 A34055) diluted to 1:500 and DAPI (Sigma, 10236276001) diluted to 1:500 in blocking solution
873 were also used to label F-actin and DNA, respectively. Embryos were washed three times in
874 PBS with 0.1% Triton X-100 before mounting in PBS covered with mineral oil in an 8-well μ-
875 Slide (Ibidi, #80827). Embryos were imaged using a laser-scanning confocal microscope (Leica
876 SP8) with a water-immersion Apochromat 40× 1.1 NA objective and HyD detectors. Samples
877 were excited with standard laser lines (405-nm, 488-nm, 552-nm, 638-nm).
878

879 **Mouse embryo siRNA injection experiments**

880 Microinjections were performed using a FemtoJet (Eppendorf). For ERH knockdown
881 experiments, mouse embryos were injected with two siRNAs (Qiagen, S104403294 &
882 S104403287) against ERH at a total concentration of 400 nM. For knockdown experiments
883 AllStars negative control siRNA (Qiagen, 1027281) was injected in embryos at a concentration
884 of 400nM.
885

886 **Differentiation of hESCs**

887 Prior to differentiation, for samples to be analyzed by RNA-seq, cells grown to a density of 80%
888 were digested with Accutase (Stem Cell Technologies, 07920) for 7 min. Then, DMEM/F12 was
889 added and gently pipetted to generate a cell suspension. After centrifugation, cells were
890 resuspended in mTeSR1 medium with 5uM ROCK inhibitor Y-27632 (Cayman Chemical,
891 10005583) and seeded at $8X10^5$-$1X10^6$ cells per well in a 6 well plate coated with Matrigel
892 (Millipore Sigma, CLS354277) 24 hours prior to differentiation treatment. For definitive

893 endoderm differentiation of hESCs, mTeSR1 medium was removed and cells were cultured in
894 DMEM/F12 (Thermo Fisher, 11320033) supplemented with 10% (v/v) Fetal Bovine Serum
895 (Hyclone, SH30071) with 100ng/ml Activin A (Peprotech, 120-14P) and 3uM CHIR99021
896 (Cayman Chemical, 13122). For ectoderm differentiation of hESCs, mTeSR medium was
897 removed and cells were cultured in DMEM/F12 with 5uM Retinoic Acid (Cayman Chemical,
898 11017).
899
900 **Cell culture siRNA transfection experiments**
901 All ERH knockdown experiments were performed with two cycles of siRNA transfection
902 (Silencer Select siRNA, Thermo Fisher, s4815), three days apart. Transfections were performed
903 in fibroblasts using 5nM final concentration of siRNA and in hESCs using a 10nM final
904 concentration of siRNA and 2.5ul/ml final concentration of Lipofectamine RNAiMAX (Thermo
905 Fisher #31985070).
906
907 **Reprogramming factor expression experiments**
908 We utilized the CytoTune-iPS 2.0 Sendai reprogramming kit (ThermoFisher, A16517) to express
909 Oct4, SOX2, KLF4 and cMyc in BJ fibroblasts. Cells were treated according to the
910 manufacturer's instructions three days after the second of two cycles of siRNA transfection.
911
912 **Cell culture immunofluorescence and microscopy**
913 For immunofluorescence of human fibroblasts and hESCs cells were washed briefly with PBS,
914 and fixed in 4% paraformaldehyde in PBS for 10 minutes. Fixed cells were washed 3 times with
915 PBS, permeabilized with ice-cold 0.1% Triton-X in PBS for 10 minutes, and washed twice with
916 PBS-T (PBS, 0.05% Tween-20). Samples were blocked with 3% donkey serum (Millipore
917 Sigma, D9663) in PBS for 2 hours at room temperature and incubated overnight at 4°C using
918 the following concentrations: anti-H3K9me3 (1:300, Abcam ab8898), anti-H3K27me3 (1:300,

919 Active Motif, 39155), anti-SOX2 (1:200, Cell Signaling, 23064), anti-OCT4 (1:200, Abcam,
920 ab19857), anti-PAX6 (1:200, BD Biosciences, 561462), anti-Brachyury (1:200, Cell Signaling,
921 81694) in 3% donkey serum in PBS. Cells were washed 3 times with PBS-T and then incubated
922 with AlexaFluor 488- or 594-conjugated secondary antibodies raised in donkey (Thermo Fisher,
923 See Methods) at a 1:500 dilution in 3% donkey serum in PBS.
924

925 **Total RNA isolation and sequencing**

926 BJ fibroblasts grown in a 6-well dish were washed twice with cold PBS, followed by addition of 1
927 mL TRIzol reagent (ThermoFisher 15596026) and transferred to a chilled 1.5 mL tube using a
928 cell scraper. The cells were homogenized by pipetting through a 25G needle attached to a 2 mL
929 syringe 10 times. 200 µl of chloroform was added to the cells and vortexed for 15 seconds,
930 followed by a room temperature incubation of 3 minutes. Phase separation of nucleic acids and
931 protein was performed by centrifugation at 12,000 x g for 15 minutes at 4C. The upper aqueous
932 phase containing RNA was then transferred to a fresh chilled tube. RNA was precipitated from
933 the aqueous phase by adding an equal volume of isopropyl alcohol, followed by a 10 minute
934 incubation at room temperature. The RNA was pelleted by centrifugation at 12,000 x g for 10
935 minutes at 4C. After removal of supernatant, the pellet was washed twice with 1 mL 75%
936 ethanol, and resuspended in nuclease free water. RNA quality was verified using a high-
937 sensitivity RNA screentape (Agilent 5067-5576). 1ug of total RNA was sent to Novogene, who
938 assembled and sequenced bulk RNA-seq libraries. Two to three biological replicates were
939 prepared for each condition.
940

941 **CUT&RUN experiments**

942 CUT&RUN was performed as described in[79], with digitonin and NaCl in wash buffers optimized
943 for permeabilization of BJ fibroblasts and retainment of SOX2 on chromatin as performed in[80].
944 Next, BJ fibroblasts were detached from cell culture plates with Accutase (STEMCELL

945 Technologies 07920), washed, bound to magnetic concanavalin A beads, and permeabilized with
946 a dig-wash buffer (0.1% digitonin, 20 mM HEPES pH 7.5, 75 mM NaCl, 0.5 mM Spermidine,
947 EDTA-free protease inhibitor). Bead-bound cells were incubated with a SOX2 (CST, clone
948 D9B8N, 23064, 1:100) or H3K9me3 antibody (abcam ab8898, 1:100) at 4C overnight. The
949 following morning, the cells were washed twice with dig-wash buffer, incubated with pA/G-MNase
950 for an hour at 4C, and washed twice more. After chilling cells on an ice block for 5 minutes, MNase
951 was activated for 30 minutes by addition of 2mM $CaCl_2$. The reaction was stopped by adding an
952 equal volume of 2X STOP buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 0.1% Digitonin, 0.05
953 mg/mL glycogen, 5 mg/mL RNase A). Digested chromatin was extracted from permeabilized cells
954 at 37C for 30 minutes, and DNA was purified by phenol-chloroform. Libraries were prepared as
955 previously described[81]. Two to three biological replicates were prepared for each condition.
956

957 **CUT&RUN data processing**

958 Paired-end CUT&RUN reads were mapped to the human genome build hg38 using bowtie2[82]
959 (v2.3.4.1) with run parameters --local --very-sensitive --no-unal --no-mixed --no-discordant -I 10 -
960 X 1000. A quality-filtered BAM file was generated with the command samtools (v1.16.1) view -q
961 5 -f 2 -bS. Global scaling factors (SF) for siERH and siControl H3K9me3 data sets were calculated
962 using the default parameters of ChIPSeqSpikeInFree[83] (R v3.6.3). Bigwig files for visualization
963 were generated from indexed BAM files with command bamCoverage --smoothLength 10 --
964 normalizeUsing CPM --scaleFactor 1/SF --ignoreForNormalization chrM (deeptools 3.5.0). BED
965 files were generated from BAMs using the bedtools command bamToBed (bedtools v2.27.1). BED
966 files were converted to bedgraph format using the bedtools command genomecov. Peaks were
967 called on the bedgraph files with SEACR[84] (v1.3), using the stringent setting and selecting the top
968 0.01% of regions by AUC. Peak lists for downstream analyses were derived from individual
969 replicates or the union of overlapping sites from biological triplicates, as indicated in the text and
970 figures. Enrichment analysis of known DNA binding motifs was performed using HOMER[85].

971

972 **Heatmaps**

973 Heatmaps were generated with deeptools[86] version 3.5.0. Counts matrices were generated with
974 command computeMatrix reference-point --missingDataAsZero --referencePoint center. Figures
975 were produced with the plotHeatmap command.

976

977 **Correlation plots**

978 To assess genome wide similarity between hBJ siERH H3K9me3, hBJ siControl H3K9me3, and
979 H1-hESC H3K9me3 data sets, the Pearson correlation of peak overlap was calculated using the
980 Intervene tool[87]. In brief, peak files for individual H3K9me3 replicates were processed with the
981 command intervene pairwise --I *H3K9me3.bed --compute fract --htype color. Visualization of
982 Pearson correlation of fractional overlap between peak sets were generated on
983 https://intervene.shinyapps.io/intervene.

984

985 **RNA-seq data analysis**

986 RNA-seq reads (paired-end) were mapped to the mm10 or hg38 genome for mouse and human
987 experiments respectively using STAR[88] (2.7.10b). The mapped reads to genes were counted
988 using featureCounts[89]. For counting repeat mapping reads the TEcount function from
989 TEtranscripts[78] was run on STAR aligned bam files using the hg38 repeatmasker gtf file
990 annotation. The output count files were analyzed by DESeq2[90] to detect differentially expressed
991 genes (Benjamini adjusted p-value < 0.05, log2 fold change > 1 or log2 fold change > 1.5).
992 Genes and repeats with fewer than 10 total mapped reads were excluded from DESeq2 and
993 TETranscripts analysis. TPM, FPKM and RPM were calculated using an in-house R script.

994

995 **Mouse blastocyst image analysis**

3D visualizations and analysis of embryos were performed using Imaris 8.2 or Imaris 9.7 software (Bitplane). The manual surface rendering module was used for cellular and nuclear segmentation, using phalloidin and DAPI signals to define cellular and nuclear boundaries, respectively. Quantifications of immunofluorescence intensity were performed by measuring the mean fluorescence intensity of voxels within the segmented object, standardized DAPI to account for light scattering.

Cell culture image analysis

Quantification of nuclear area and signal intensity was performed using the Nuclear Morphometric Analysis (NMA) plugin for ImageJ[91]. In ImageJ saturated pixels was set to 0.35% prior to processing. NMA plugin was run using the following settings: show ellipse and use watershed options selected, minimum nuclei area=150, maximum nuclei area=6000, nuclei intensity segmentation threshold=95, nuclei insertion intensity=93%.

Gene ontology enrichment analysis

Gene ontology enrichment analysis was performed using PANTHER[92] (v18.0) with settings, statistical overrepresentation test: GO biological process complete, all genes in the database were used as the background reference list, and statistical significance was determined using Fisher's Exact test with false discovery rate.

Statistical analysis

Statistical analyses were performed in GraphPad Prism, the R software for statistical computing and with DESeq2. Pairwise comparisons were analyzed using an unpaired, two-tailed Student's t-test. Correlation was analyzed with a two-tailed Pearson's Correlation test. Significant overlap between gene sets was analyzed with the hypergeometric test. Differential gene expression analysis was performed using DESeq2 which calculates p-values by the Wald test and adjusted

1022 p-values corrected for multiple testing using the Banjamini and Hochberg method. Reproducibility
1023 was confirmed by independent experiments.
1024

1025 **Acknowledgements**

1026 R.L.M. was supported by an NIH NIDDK K01 postdoctoral fellowship DK117970-01. N.P. was
1027 supported by NIGMS (GM139970-01) and NICHD (HD102013-01A1). R35GM153180 supported
1028 work in the lab of K.S.Z.
1029

1030 **Author contributions**

1031 Conceptualization, R.L.M. and K.Z.; Formal analysis, A.K., B.H., H.F., K.T. and R.L.M.; Funding
1032 acquisition, K.Z; Investigation, A.K., B.H., H.F., K.T., J.Z., A.B. and R.L.M.; Methodology, A.K.,
1033 B.H., R.L.M. and K.Z.; Project administration, R.L.M and K.Z.; Resources, K.Z,; Supervision, M.T-
1034 P., N.P., R.L.M. and K.Z.; Writing – Original Draft, A.K., R.L.M. and K.Z.
1035

1036 **Declaration of Interests**

1037 The authors declare no competing interests.
1038

1039 **Corresponding author**

1040 Correspondence to Ryan McCarthy (ryan.mccarthy@seatttlechildrens.org) and Kenneth Zaret
1041 (zaret@pennmedicine.upenn.edu).
1042

1043 **Data availability**

1044 • RNA-seq and CUT&RUN datasets collected in this study have been deposited in the
1045 GEO of NCBI and are publicly available as of the date of publication. Accession
1046 numbers are listed in the Methods. GEO reviewer tokens for accessing the data

while it remains in private status are for RNA-seq GSE268901 (wbsvkukqhzobpan), and for CUT&RUN GSE268961 (qngpsiywdvyfjyj).

Code availability

- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Supplementary Tables

Supplementary Table 1. H3K9me3 peak coordinates in siERH human fibroblasts.

Supplementary Table 2. Fibroblast +OSKM 48 hours, SOX2 CUT&RUN peaks in siControl and siERH.

Supplementary Table 3. Fibroblast +OSKM 48 hours, siERH/siControl RNA-seq DESeq2 differential gene expression analysis results.

Supplementary Table 4. Fibroblast +OSKM 48 hours, siERH/siControl TETranscripts differential repeat expression analysis results.

Supplementary Table 5. Blastocyst immunofluorescence quantification values.

Supplementary Table 6. Blastocyst RNA-seq DESeq2 siERH/siControl differential gene expression analysis results.

Supplementary Table 7. List of genes categorized as group 1-4 by preimplantation expression from Gao et al., 2017[33] and TPM values from this study's siControl and siERH mouse blastocysts RNA-seq.

Supplementary Table 8. H1 hESC untreated, endoderm and ectoderm RNA-seq DESeq2 siERH/siControl differential gene expression analysis results.

1072 Supplementary Table 9. H1 hESC untreated, endoderm, and ectoderm TETranscripts
1073 siERH/siControl differential repeat expression analysis

1074 Supplementary Table 10. H1 hESC untreated, endoderm and ectoderm GO analysis results.
1075

1076 **Extended data**

1077 Extended Data Fig. 1. ERH depletion reverts fibroblast H3K9me3 patterns to an hESC-like state

1078 **a.** Browser tracks comparing H3K9me3 distribution patterns in human fibroblasts compared
1079 to hESC.

1080 **b.** Venn diagram of H3K9me3 peaks called using SEACR from fibroblast siControl and
1081 siERH H3K9me3 CUT&RUN and H1 hESC ChIP-seq (GSM1003585).

1082 **c.** Browser track showing SEACR peaks called in siControl H3K9me3 domain.

1083 **d.** Nanog staining in hESCs treated for 5 weeks with siRNAs to sustain ERH depletion or
1084 control.

1085 **e.** Immunofluorescence for SOX2 (green) and H3K9me3 (red) in siControl and siERH H1
1086 hESCs in pluripotent conditions at higher magnification to show nuclear organization.
1087 Scale bars, 50μm

1088 **f.** Immunofluorescence for DAPI (blue) and H3K27me3 (red) in siControl and siERH H1
1089 hESCs in pluripotent conditions. Scale bars, 50μm.

1090 **g.** Nuclear area, white arrows indicate nuclear blebs. $p=5.6*10^{-16}$, welch's two sample t-test.
1091 Scale bars, 50μm.

1092 **h.** Expression (TPM) of H3K9me3 HMTs in siControl and siERH H1 hESCs (n=2 for siControl
1093 H1 hESCs and n=3 for siERH H1 hESCs).

1094 **i.** Browser tracks showing H3K9me3 in human fibroblasts or hESCs treated with siControl
1095 or siERH.

1096 **j.** Correlation of H3K9me3 from hESCs after siControl or siERH.

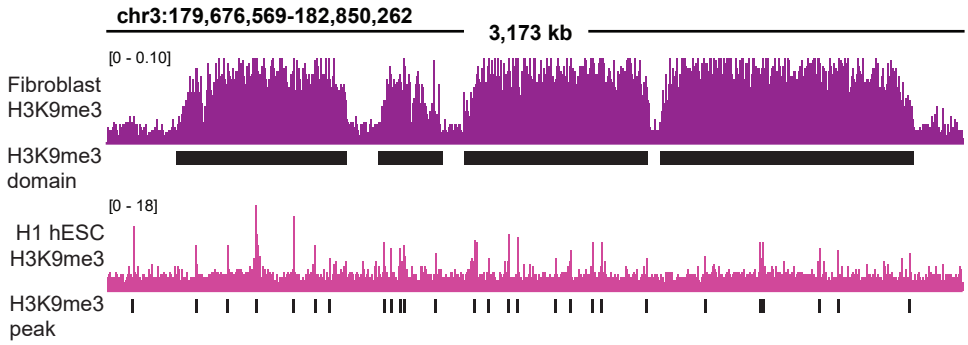
1097 **k.** Browser tracks showing H3K9me3 (GSM1003585) and HP1a (GSM1701825) in H1
1098 hESCs and H3K9me3 in siControl and siERH human fibroblasts.

1099 **l.** Top 10 statistically enriched transcription factor binding motifs at H3K9me3 peaks in
1100 siERH treated human fibroblasts (homer findMotifs).

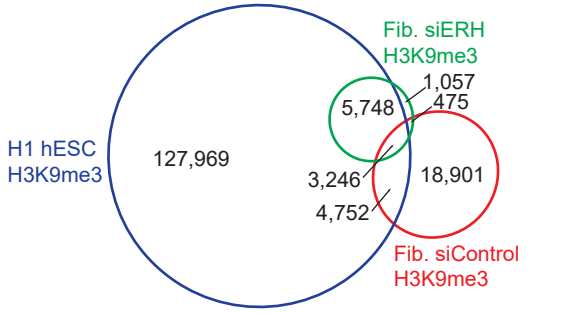
1101 **m.** Occurrence of repeats from RepEnrich database and repeat occurrence at sites of siERH
1102 fibroblasts H3K9me3 peaks.

Katznelson et al., Extended Data Figure 1

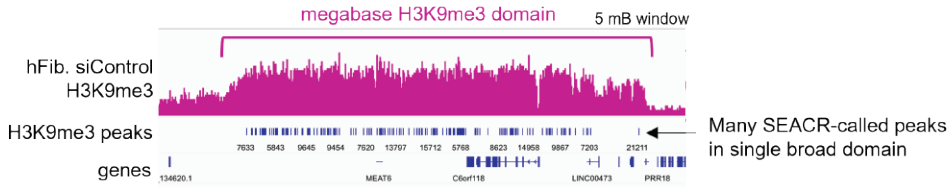
a H3K9me3 distribution in fibroblasts and hESCs



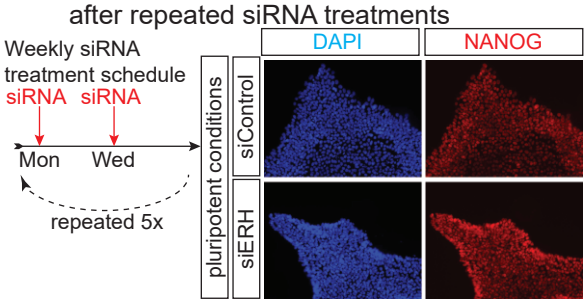
b Venn diagram of overlap of H3K9me3 peaks called by SEACR in H1 hESCs, and siControl and siERH treated fibroblasts



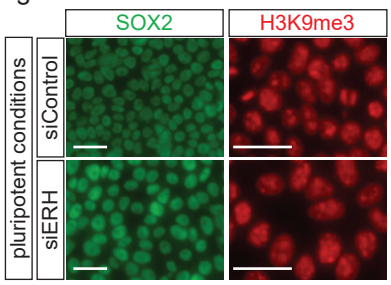
c Browser track showing limitations of peak calling to capture domain architecture of H3K9me3



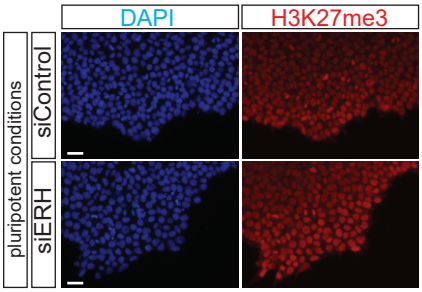
d Pluripotency maintenance in hESCs after repeated siRNA treatments



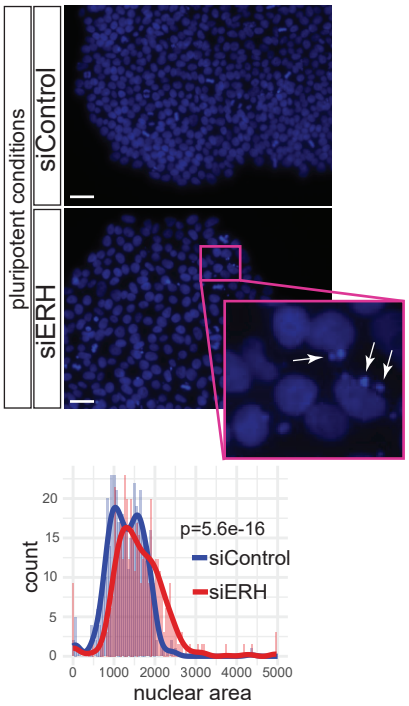
e Higher magnification showing nuclear organization of SOX2 and H3K9me3



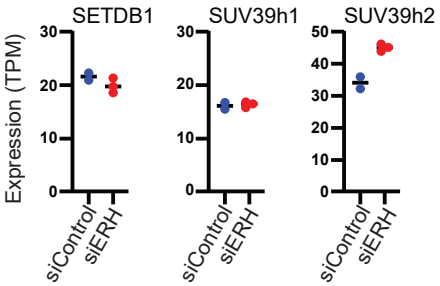
f H3K27me3 immunofluorescence in siControl and siERH H1 hESCs



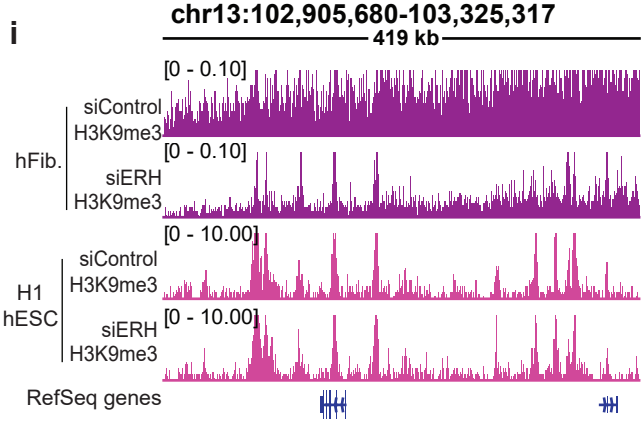
g Quantification of nuclear area in siControl and siERH H1 hESCs



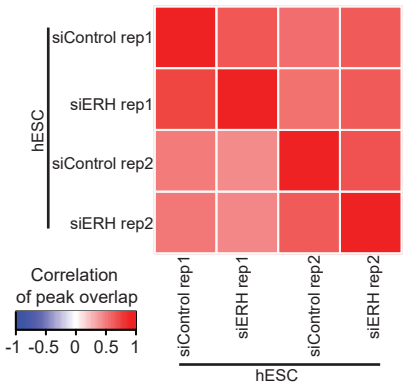
h H3K9me3 HMT expression in H1 hESCs



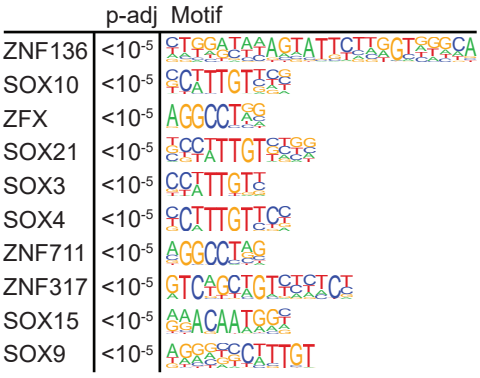
i



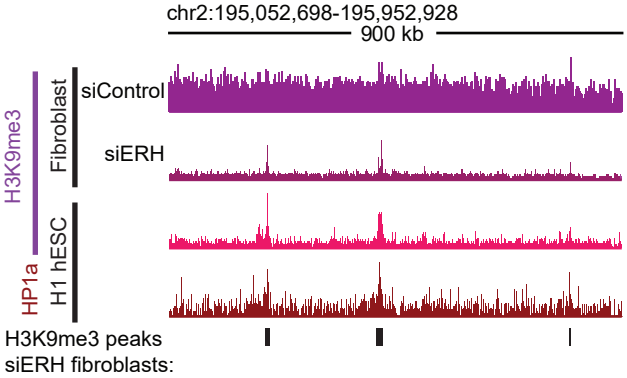
j



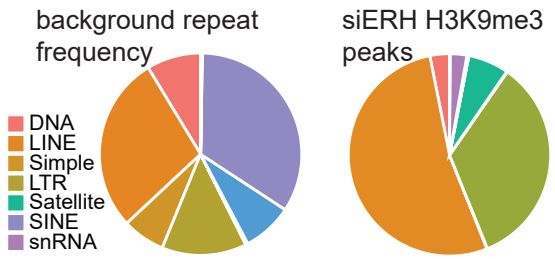
k Top 10 statistically enriched transcription factor motifs at siERH fibroblast H3K9me3 peaks



l Browser track of H3K9me3 in siControl and siERH compared to H3K9me3 and Hp1a in H1 hESC



m Proportion of repeat class at siERH fibroblast H3K9me3 peaks vs background abundance



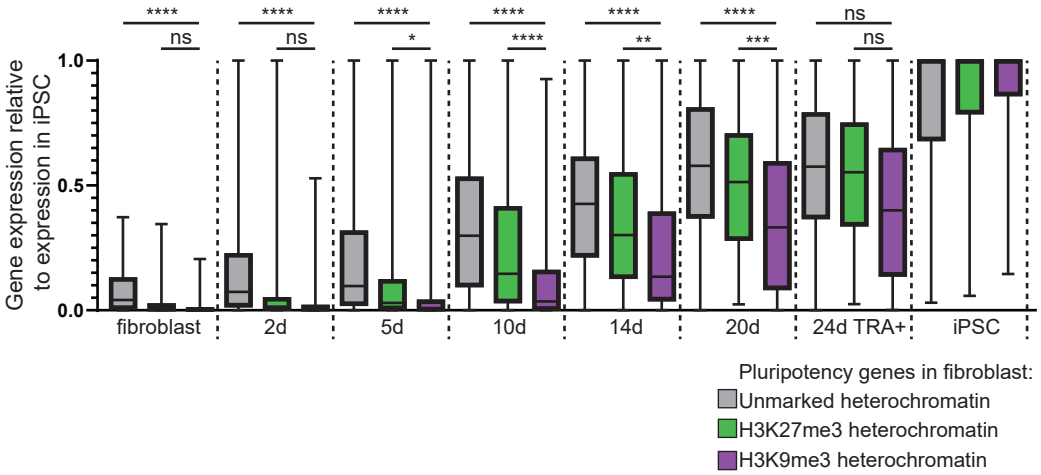
1103 Extended Data Fig. 2. Kinetics of iPSC gene activation is effected by heterochromatin marks.

1104 **a.** Expression of genes activated during fibroblasts to iPSC reprogramming timepoints[26] that
1105 are located in unmarked, H3K27me3 or H3K9me3 marked heterochromatin as annotated
1106 by Becker et al., 2017[5].

1107

Katznelson et al., Extended Data Figure 2

a delayed activation iPSC genes in fibroblast H3K9me3 domains during reprogramming



1108

1109 Extended Data Fig. 3. ERH depletion enhances SOX2 binding in H3K9me3 domains.

1110 **a.** Venn diagram showing overlap between SOX2 peak calls in siControl and siERH
1111 fibroblasts+OSKM.

1112 **b.** Browser track of example hESC SOX2 and NANOG peak coincident with fibroblast siERH
1113 H3K9me3 loss and SOX2 peak gain.

1114 **c.** Characterization by fibroblast H3K9me3 status of SOX2 binding sites gained in siControl
1115 and siERH fibroblasts+OSKM.

1116 **d.** SOX2 CUT&RUN average profile in siControl and siERH fibroblasts+OSKM for genome
1117 regions identified as being differentially bound by reprogramming transcription factors [20].

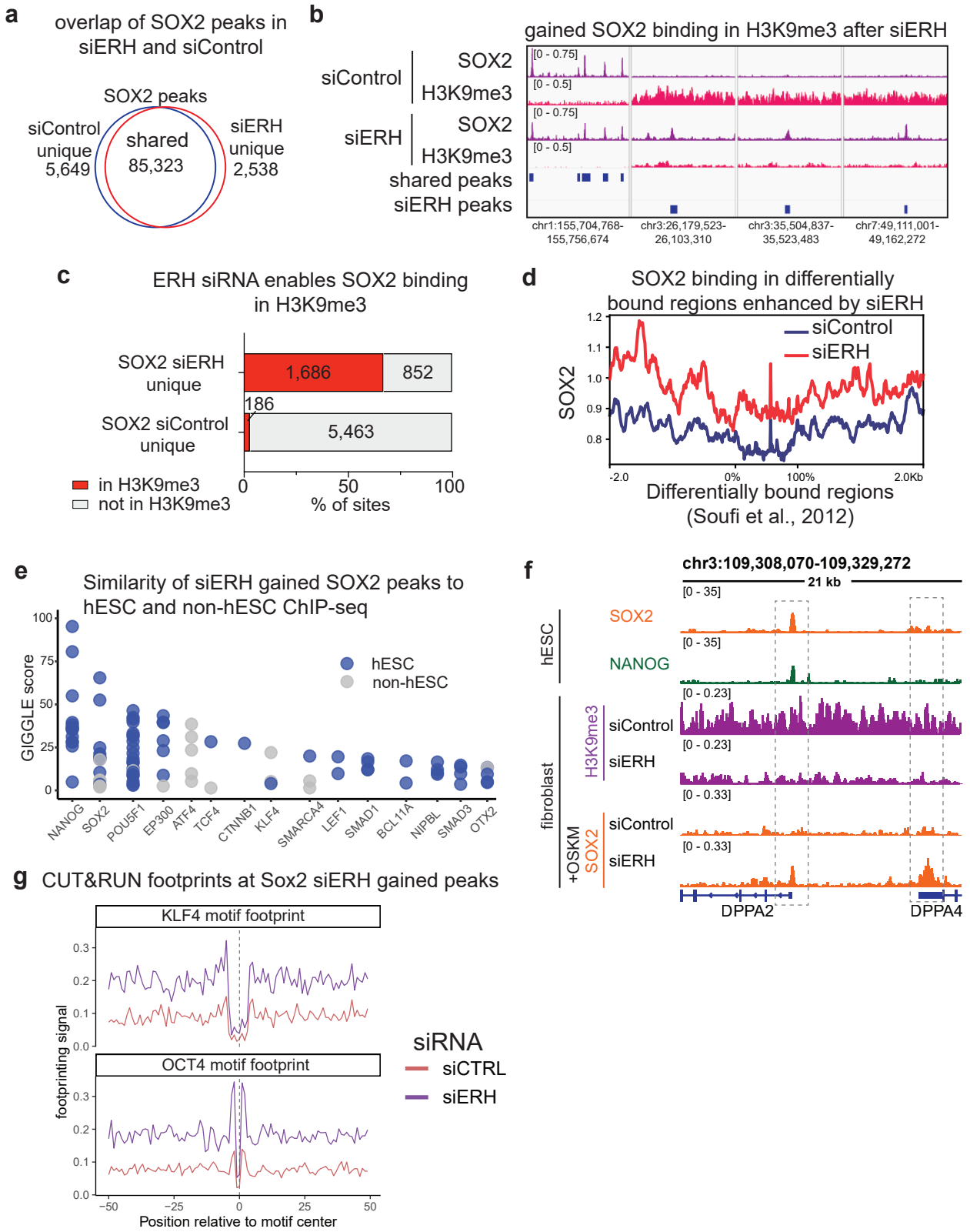
1118 **e.** Rank ordered GIGGLE enrichment scores. Each circle indicates similarity between siERH
1119 SOX2-gained sites and a ChIP-seq peak set from cistromeDB. Blue circles are peak sets
1120 profiled in hESCs, white circles are peaks from a non-embryonic origin.

1121 **f.** Browser track of H3K9me3 in siControl and siERH fibroblasts and RNA in siControl and
1122 siERH fibroblasts+OSKM at DPPA4.

1123 **g.** KLF4 and OCT4 footprint analysis at SOX2 siERH gained peaks.

1124

Katznelson et al., Extended Data Figure 3



1125 Extended Data Fig. 4. ERH depletion enhances gene activation of pluripotency

1126 **a.** Browser track of H3K9me3 loss and gene activation of DPPA4 in siERH treated human
1127 fibroblasts+OSKM.

1128 **b.** Example 8 cell genes activated in siERH+OSKM (n=2 per condition) but not activated
1129 during eight time points of iPSC reprogramming time course[26].

1130 **c.** Comparison of expression (FPKM) of naïve genes KLF5 and DNMT3L of siControl and
1131 siERH fibroblasts+OSKM from this study (n=2 per condition) with pluripotent and naïve
1132 human ESC data from Bi et al., 2022.

1133 **d.** Browser tracks of H3K9me3 in siControl and siERH fibroblasts and RNA in siControl and
1134 siERH fibroblasts+OSKM, at a locus of an LTR7Y-HERHV transposable element and
1135 nearby gene.

1136 **e.** Volcano plot showing differential repeat expression in siERH vs siControl
1137 fibroblast+OSKM for repeats annotated in[6] as repressed by H3K9me3 in human
1138 blastocysts. (TETranscripts, DESeq2, padj<=0.05, log2fc>0.5).

1139 **f.** Representative phase image of siControl and siERH treated fibroblasts+OSKM 2 days
1140 after infection with OSKM sendai virus.

1141 **g.** Representative phase image showing cell confluency of siControl and siERH
1142 fibroblasts+OSKM cells 2 weeks after initial OSKM infection with titrated siRNA
1143 concentration. All treatments were plated at identical cell numbers 1 day prior to infection.

1144 **h.** Plot of iPSC colony count from siControl and siERH fibroblasts+OSKM cells 3 weeks after
1145 initial OSKM infection with titrated siRNA concentration (n=2 per condition). Each value
1146 represents the number of colonies generated in 1 well of a 24 well plate.

1147 **i.** Alkaline phosphatase staining of siControl and siERH +/- sip53 treated fibroblasts+OSKM
1148 21 days after infection.

1149 **j.** Expression (TPM) of RIPK2 in siControl and siERH fibroblasts+OSKM 48 hours after
1150 infection (n=2 per condition).

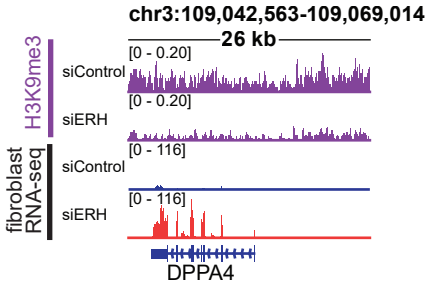
1151 **k.** Plot of iPSC colony count 3 weeks after initial OSKM infection for siControl and siERH
1152 cells reprogrammed in media containing the inhibitors GSK2983559, Ac-DEVD-CHO and
1153 GSK872 (n=4 per condition).

1154 **l.** Plot of iPSC colony count (n=2 per condition) and imaging of siERH fibroblasts+OSKM 3
1155 weeks after infection and treated with indicated combinations of inhibitors or DMSO
1156 control.

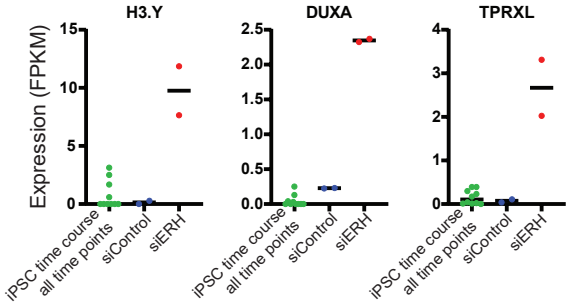
1157

Katznelson et al., Extended Data Figure 4

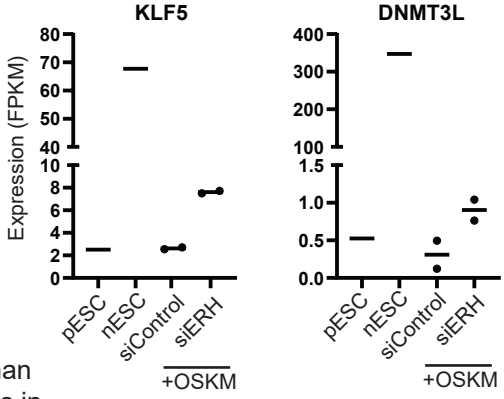
a Browser track of H3K9me3 and expression change at example naïve pluripotency gene



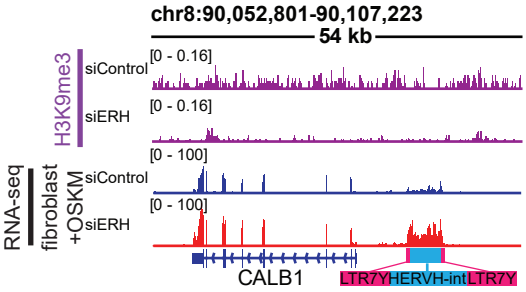
b Expression of naïve pluripotency genes during iPSC reprogramming and in siControl and siERH fibroblasts+OSKM



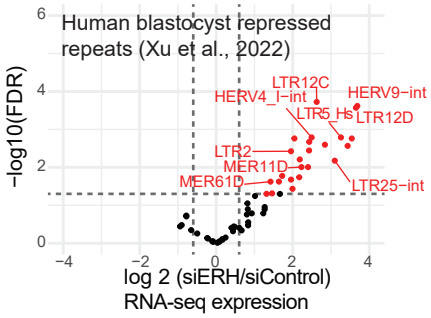
c Average expression in plutipotent (pESC) and naive (nESC) cells from Bi et al., 2022 compared to siCTRL and siERH +OSKM fibroblasts



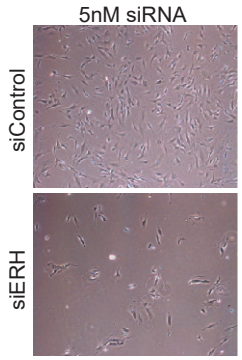
d H3K9me3 and expression change at example gene and proximal repeat



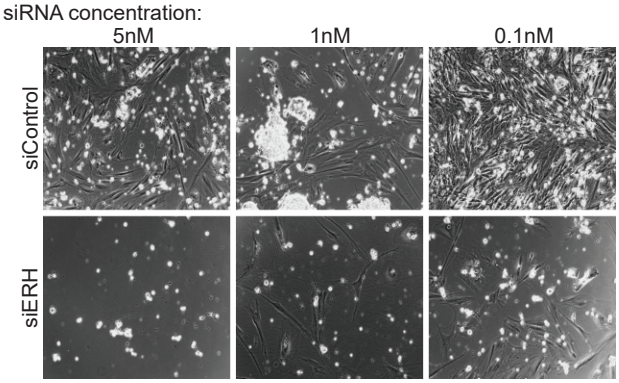
e Expression changes of human blastocyst repressed repeats in siERH vs siControl fibroblasts+OSKM



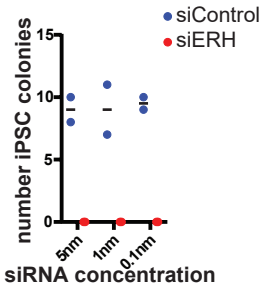
f 2 days post infection



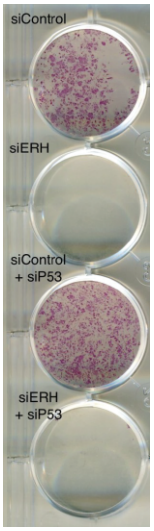
g 7 days post infection



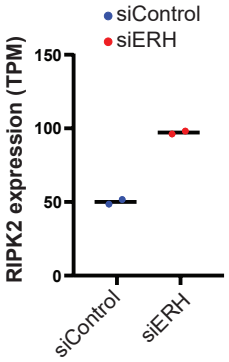
h iPSC colony number 14 days post infection



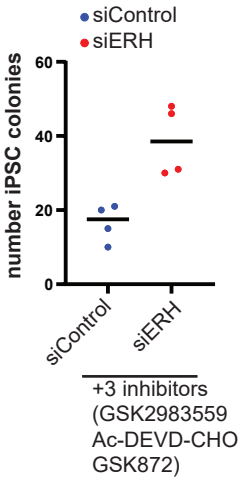
i



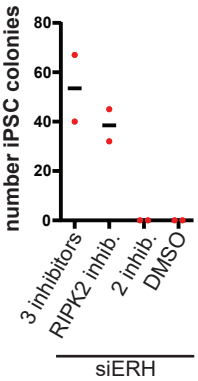
j

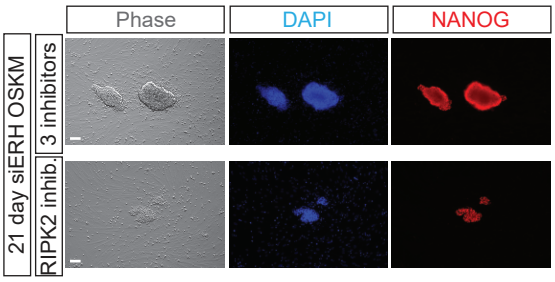


k



l





1158 Extended Data Fig. 5. Characterization of ERH and the effects of ERH depletion in mouse
1159 blastocysts

1160 **a.** ERH RNA expression in reads per kilobase million (RPKM) during stages of mouse pre-
1161 implantation development, RNA-seq data from[33]. unpaired two tail t-test; n=6.

1162 **b.** Fluorescent imaging of ERH (magenta) and quantification in ICM and trophectoderm cells
1163 (n=12 for ICM and n=13 for Trophectoderm, unpaired t-test, ns=not significant).

1164 **c.** ERH (magenta) and DAPI (teal) staining in mouse blastocysts injected at the 1 cell stage
1165 with siControl and siERH siRNAs showing protein depletion. Scale bars, 10μm.

1166 **d.** Representative DIC images of mouse blastocysts injected at the 1 cell stage with siControl
1167 and siERH siRNAs, two examples for each shown.

1168 **e.** Fluorescent imaging of trophectoderm marker CDX2 (green), pluripotency marker OCT4
1169 (magenta), with Phalloidin (orange) to define cell boundary and quantification of OCT4
1170 levels in siControl siERH blastocysts with DAPI (blue). (n=8 per condition, t-test, ns=not
1171 significant).

1172 **f.** Representative DAPI staining of siControl and siERH blastocysts. White arrows indicated
1173 DNA shedding.

1174 **g.** Expression of Erh in siControl and siERH treated mouse blastocysts by RNA-seq. (n=3
1175 per condition, t-test, $p>0.0001$).

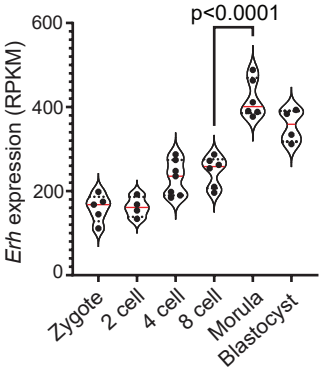
1176 **h.** TPM expression values from RNA-seq for primitive endoderm genes. (n=3 per condition,
1177 t-test)

1178 **i.** TPM expression values from RNA-seq for epiblast genes. (n=3 per condition, t-test,
1179 ns=not significant)

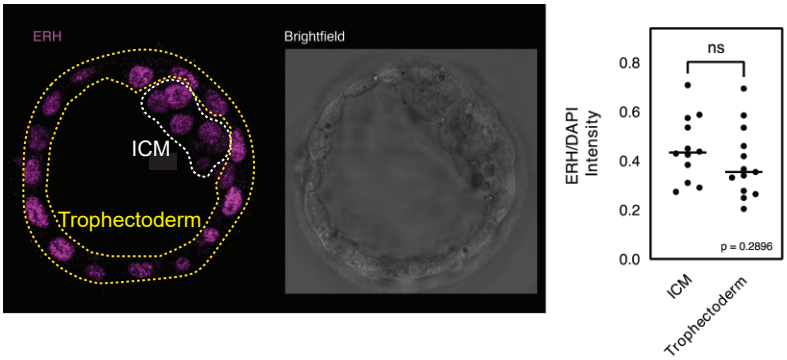
1180

Katznelson et al., Extended Data Figure 5

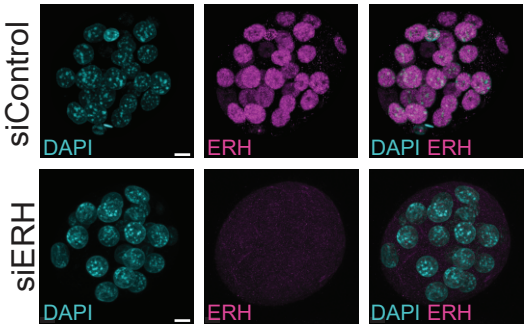
a *Erh* expression during mouse preimplantation



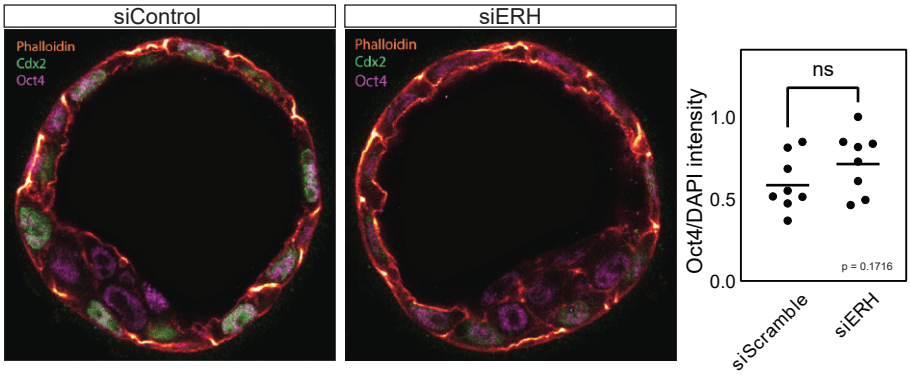
b ERH immunofluorescence in mouse blastocyst comparing protein levels in trophectoderm and ICM cells



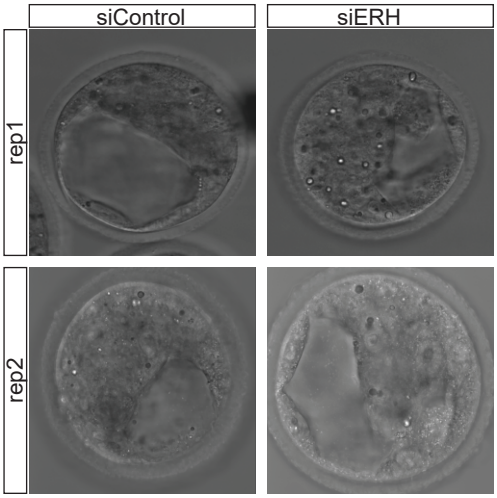
c ERH protein staining in mouse blastocysts injected with siRNA at the 1-cell stage



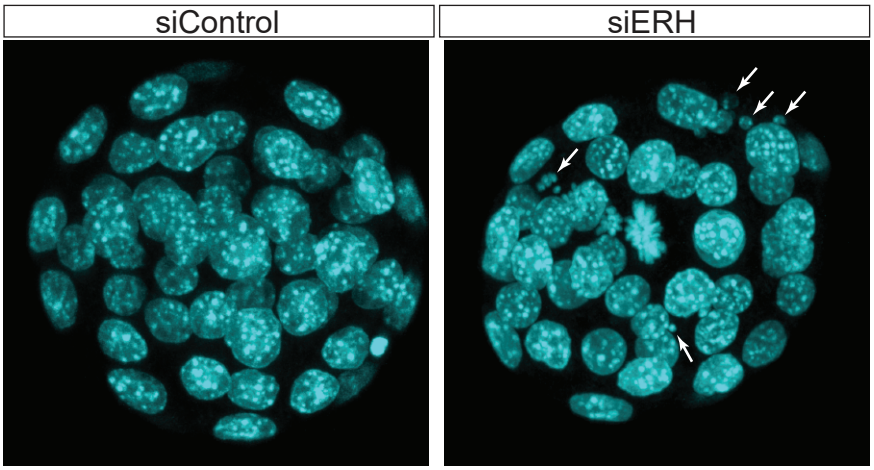
e Comparison of OCT4 protein levels between siControl and siERH mouse blastocysts



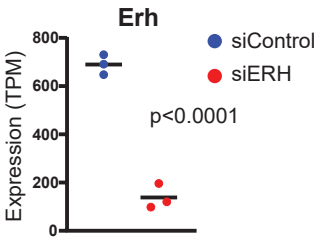
d Example brightfield images of mouse blastocysts injected at the 1 cell stage with siControl or siERH siRNAs



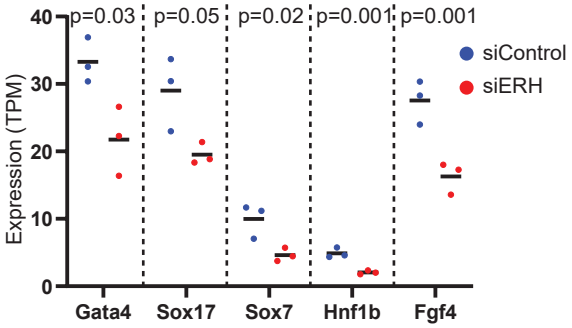
f DAPI staining in siControl and siERH blastocysts showing increased DNA shedding



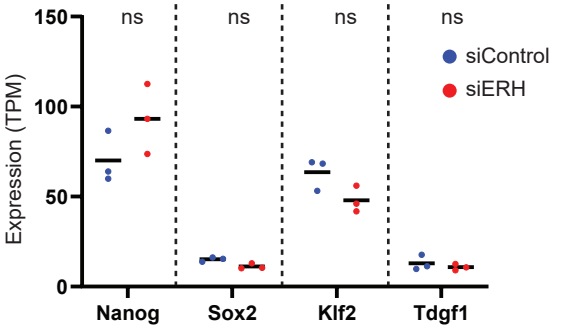
g *Erh* expression in siControl and siERH mouse blastocysts



h Expression of primitive endoderm genes in siControl and siERH mouse blastocysts



i Expression of epiblast genes in siControl and siERH mouse blastocysts



1181 Extended Data Fig. 6. ERH is involved in silencing of pluripotency genes during differentiation

1182 **a.** Imaging of H1 hESCs in pluripotent and 1 day endoderm differentiated showing changes
1183 in colony morphology (phase), DAPI (blue), silencing of SOX2 (green) and induction of
1184 endoderm marker SOX17 (red). Scale bars, 100µm.

1185 **b.** Quantification of nuclear NANOG immunofluorescence intensity in siControl vs siERH 1
1186 day endoderm differentiated H1 hESCs. (unpaired t-test).

1187 **c.** ERH transcripts per million (TPM) values from RNA-seq in siControl and siERH H1 hESCs
1188 in untreated, endoderm (DE) and ectoderm (EC) differentiation conditions.

1189 **d.** Volcano plot showing differential gene expression in siERH vs siControl H1 in pluripotency
1190 conditions. Inset pie chart indicates the number of genes upregulated and down regulated
1191 in siERH vs siControl H1 (DESeq2, padj<=0.05, log2fc>0.5).

1192 **e.** Volcano plot showing differential expression of genes upregulated in naïve vs primed
1193 hESCs (List from Messmer et al., 2019[43]) in siERH vs siControl H1 in pluripotency
1194 conditions. Inset pie chart indicates the number of genes upregulated and down regulated
1195 in siERH vs siControl H1 (DESeq2, padj<=0.05, log2fc>0.5).

1196 **f.** Volcano plot showing differential gene expression in siERH vs siControl 1 day endoderm
1197 for genes down regulated in siControl DE vs siControl H1. Inset pie chart indicates the
1198 number of genes upregulated and down regulated in siERH vs siControl DE (DESeq2,
1199 padj<=0.05, log2fc>0.5).

1200 **g.** Volcano plot showing differential gene expression in siERH vs siControl 2 day ectoderm
1201 for genes down regulated in siControl EC vs siControl H1. Inset pie chart indicates the
1202 number of genes upregulated and down regulated in siERH vs siControl EC (DESeq2,
1203 padj<=0.05, log2fc>0.5).

1204 **h.** Volcano plot showing differential expression of genes upregulated in naïve vs primed
1205 hESCs (List from Messmer et al., 2019[43]) in siERH vs siControl H1 hESCs in 1 day
1206 endoderm conditions. Inset pie chart indicates the number of genes upregulated and down
1207 regulated in siERH vs siControl H1 (DESeq2, padj<=0.05, log2fc>0.5).

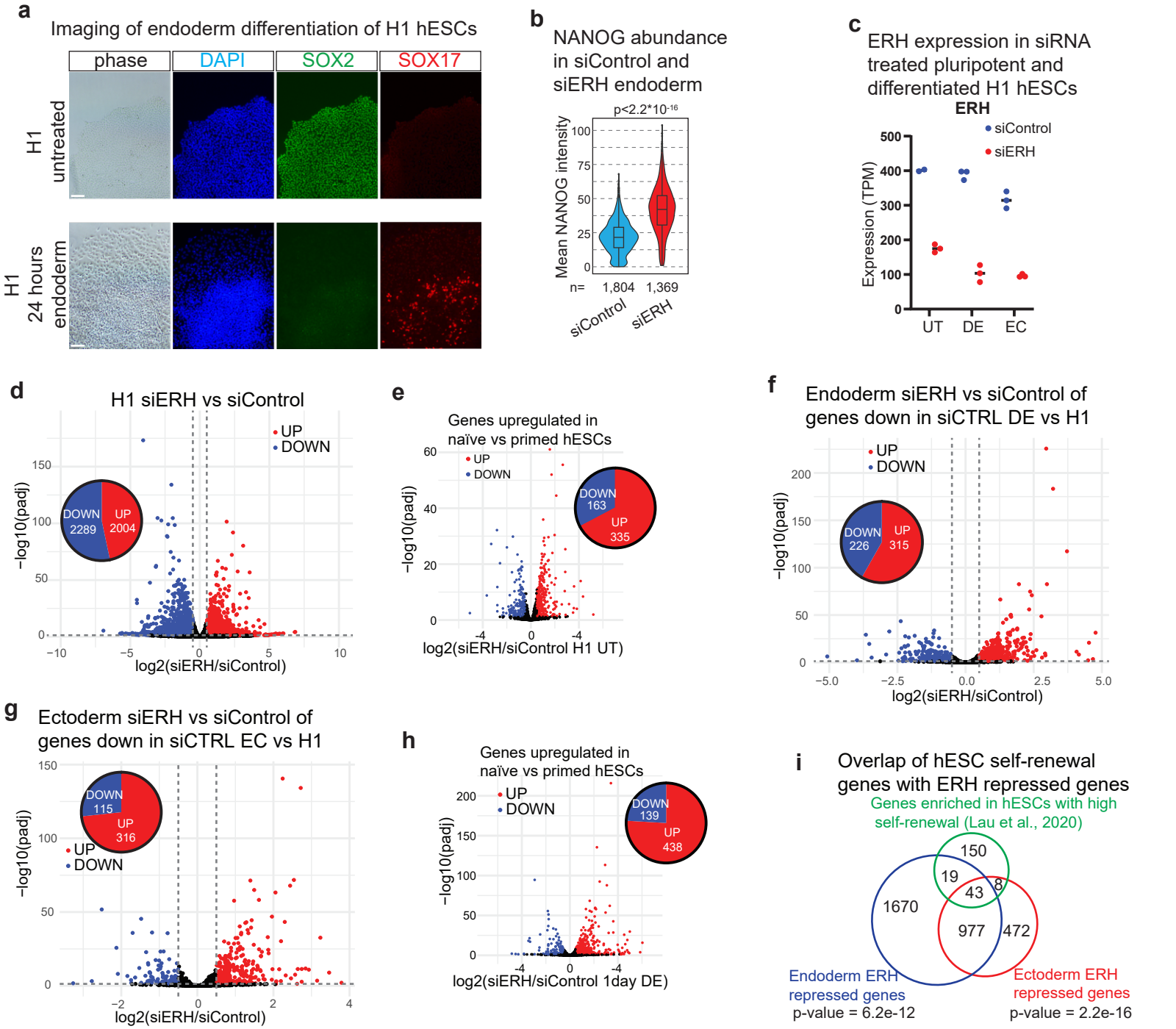
1208 **i.** Intersection of genes enriched in hESCs with high self-renewal[93] with genes repressed by
1209 ERH during endoderm and ectoderm differentiation. (hypergeometric test)

1210 **j.** Graphs of naïve gene RNA-seq, transcripts per million (TPM) values from RNA-seq
1211 replicates shown (DESeq2 adjusted p-value).

1212 **k.** Graphs of key naïve gene expression in siControl and siERH H1 hESCs compared to
1213 FPKM values for naïve and primed ESCs from Bi et al., 2022.

1214 For all RNA-seq n=3, except for siCTRL UT where n=2.

Katznelson et al., Extended Data Figure 6



1215 Extended Data Fig. 7. ERH represses alternative lineage gene activation during hESC
1216 differentiation

1217 **a.** GO analysis (PantherDB, statistical overrepresentation) of genes upregulated
1218 (padj<=0.05, log2 fold change >= 0.5) in siERH vs siControl H1 hESCs 2 day ectoderm
1219 differentiation.

1220 **b.** Example meiotic and gametogenic genes activated during endoderm and ectoderm
1221 differentiation by ERH depletion, transcripts per million (TPM) values from RNA-seq
1222 replicates shown (DESeq2 adjusted p-value).

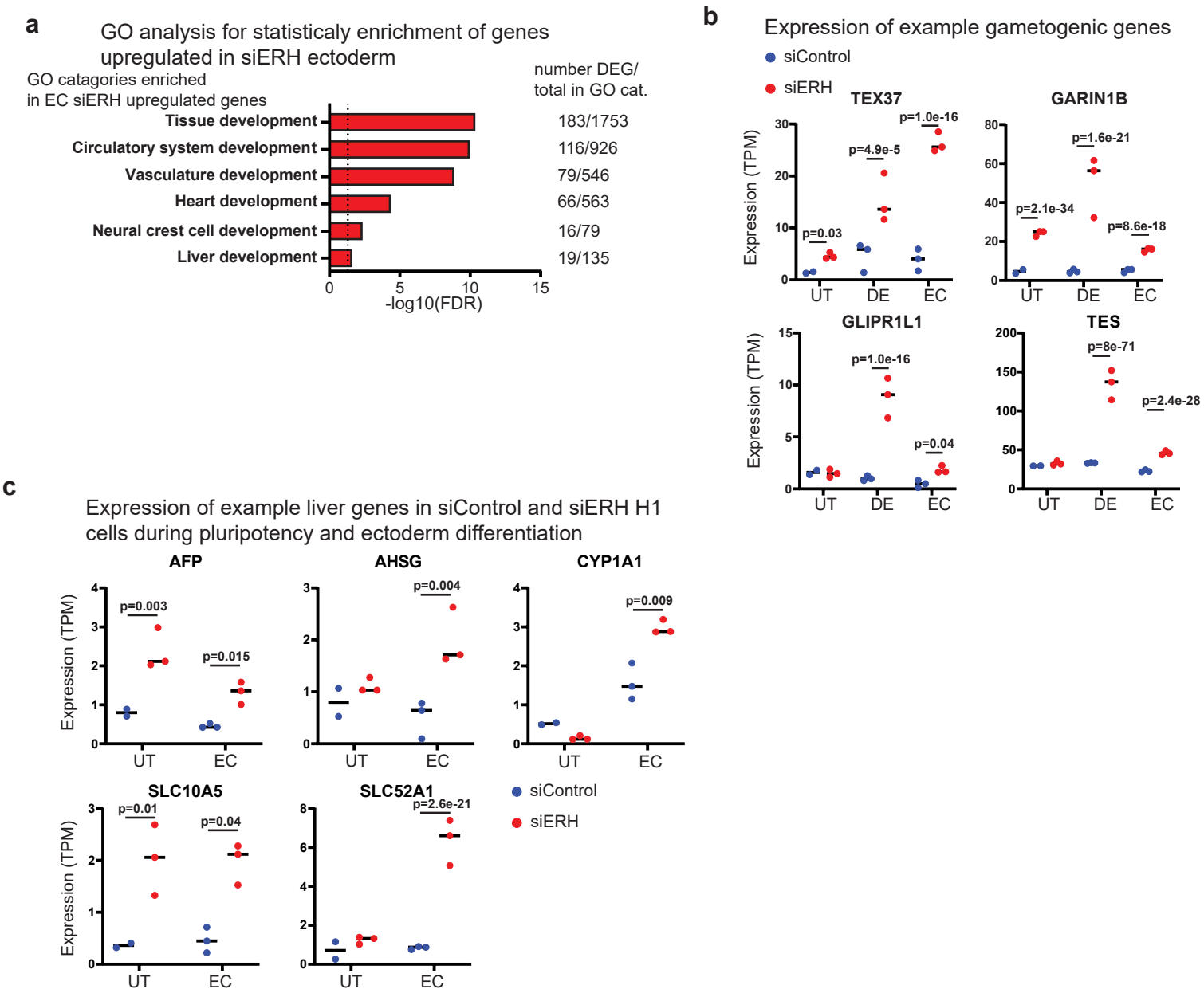
1223 **c.** Example alternative lineage genes activated during ectoderm differentiation by ERH
1224 depletion, transcripts per million (TPM) values from RNA-seq replicates shown (DESeq2
1225 adjusted p-value).

1226 For all RNA-seq n=3, except for siCTRL UT where n=2.

1227

1228

Katznelson et al., Extended Data Figure 7



1229 Extended Data Fig. 8. ERH represses evolutionarily young transposable elements with regulatory
1230 function

1231 **a.** H3K9me3 data from H1 hESCs (GSM1003585) and hESC derived endoderm
1232 (GSM916057) at gene bodies +/- 1kb for upregulated (adjusted pvalue<=0.05;
1233 log2foldchange>=0.5) by ERH depletion during 1 day endoderm differentiation.

1234 **b.** SVA class element expression plots in reads per million (RPM) from RNA-seq data
1235 analyzed by TETranscripts. (DESeq2 adjusted p-value).

1236 **c.** LTR7Y expression plot from RNA-seq data analyzed by TETranscripts in untreated (UT)
1237 and endoderm (DE) differentiated H1 hESCs treated with siControl or siERH. (DESeq2
1238 adjusted p-value).

1239 **d.** LTR5_Hs and HERVK expression plots in reads per million (RPM) from RNA-seq data
1240 analyzed by TETranscripts. (DESeq2 adjusted p-value).

1241 **e.** Browser track showing RNA-seq of example repeat element and nearby gene that is
1242 expressed in pluripotency, silenced in siControl DE and EC but expressed in siERH DE
1243 and EC conditions.

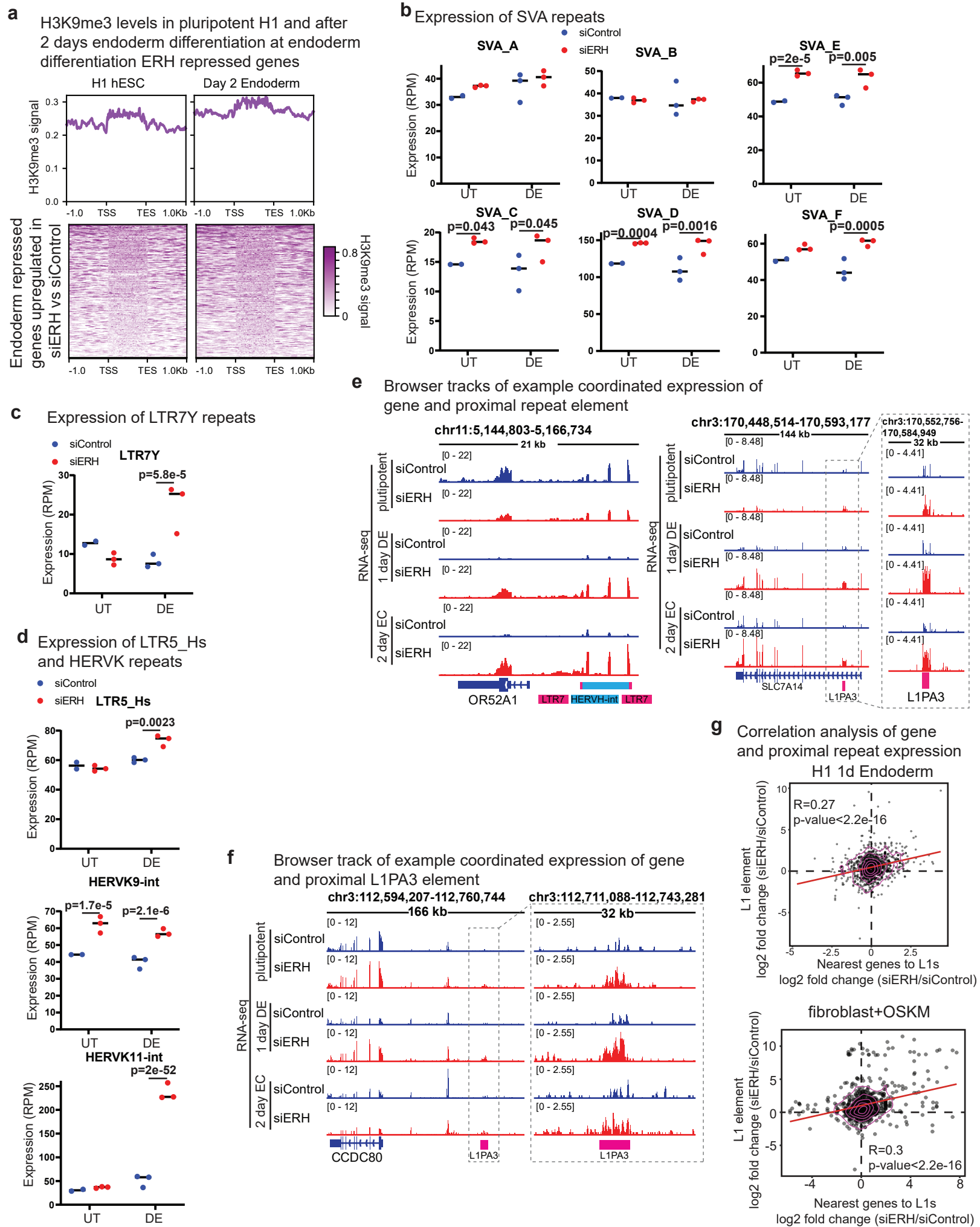
1244 **f.** Browser track showing RNA-seq of example repeat element and nearby gene that is
1245 expressed in pluripotency, silenced in siControl DE and EC but expressed in siERH DE
1246 condition.

1247 **g.** Plot of expression change of specific L1 element loci vs expression change of
1248 corresponding nearest gene within 25kb in 1 day endoderm. R, Pearson correlation.

1249 For all RNA-seq n=3, except for siCTRL UT where n=2.

1250

Katznelson et al., Extended Data Figure 8



1251 Extended Data Fig. 9. ERH dependent repression of H3K9me3 marked repeat elements

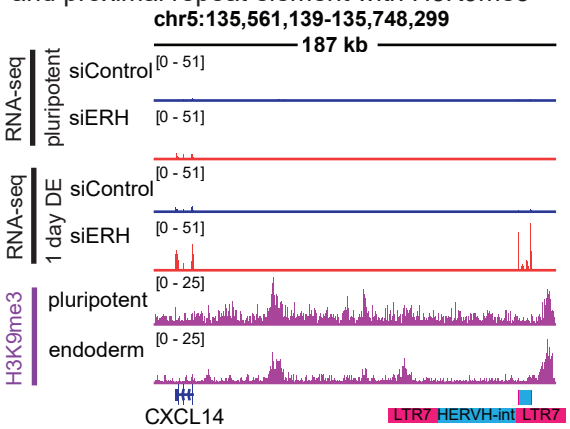
1252 **a.** Browser track showing RNA-seq of example repeat element and nearby gene in siControl
1253 and siERH treated H1 hESCs after 1 day of endoderm differentiation and H3K9me3 tracks
1254 from H1 hESCs (GSM1003585) and hESC derived endoderm (GSM916057).

1255 **b.** Plots of H3K9me3 CUT&RUN signal at L1 elements in endoderm differentiated H1 hESCs
1256 treated with siControl or siERH. Data from loci at all L1 element loci with at least 2-fold
1257 increased RNA expression upon ERH depletion.

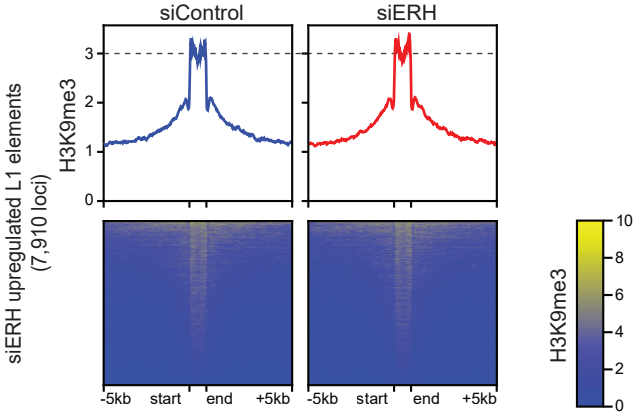
1258 **c.** Browser track showing example repeat element gaining H3K9me3 during endoderm
1259 differentiation in siControl and gaining less H3K9me3 in siERH. RNA-seq track showing
1260 proximal gene to repeat element being expressed in siERH endoderm.

Katznelson et al., Extended Data Figure 9

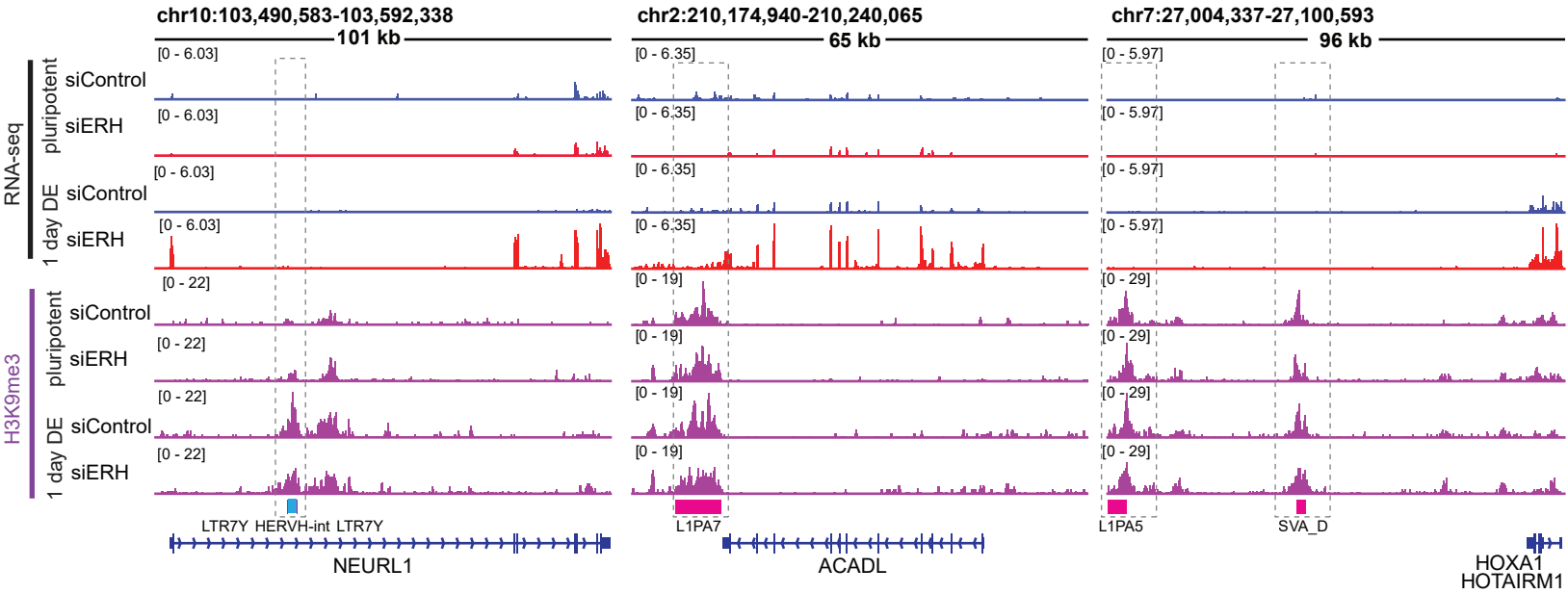
a Browser track of example coordinated expression of gene and proximal repeat element with H3K9me3



b H3K9me3 CUT&RUN signal in siControl and siERH endoderm differentiated H1 hESCs at siERH upregulated L1 element loci at all or endoerm H3K9me3 gained loci



c Browser tracks of example repeats and proximal genes upregulated by ERH knockdown during endoderm differentiation with and without decreased H3K9me3



1261 Extended Data Fig. 10. ERH protein levels are higher in differentiated cells and ERH is recruited
1262 to the chromatin by RNA and RNA binding proteins.

1263 **a.** ERH expression (TPM) from RNA-seq in siControl treated human fibroblasts, pluripotent
1264 hESCs, 1 day endoderm differentiated hESCs (DE), and 2 day ectoderm differentiated
1265 hESCs (EC). (n=2 for siControl fibroblasts and UT H1 hESCs, n=3 for siControl DE and
1266 EC H1 hESCs).

1267 **b.** ERH staining in human fibroblasts and pluripotent H1 hESCs, DAPI (blue), ERH (red).
1268 Staining and imaging performed in parallel using same conditions and microscope settings
1269 for both samples.

1270 **c.** ERH staining in pluripotent H1 hESCs and 1 day endoderm differentiated H1 hESCs, DAPI
1271 (blue), ERH (red). Scale bars, 50μm. Staining and imaging performed in parallel using
1272 same conditions and microscope settings for both samples.

1273 **d.** Western blot after RNase A treatment during soluble nuclear protein isolation. Western for
1274 ERH and YTHDC1 in nuclear and chromatin fractions with and without RNase. RBMX and
1275 H3 included as known RNA-sensitive and RNA-insensitive, chromatin-associated
1276 proteins.

1277 **e.** Quantification of immunofluorescence (IF) of nuclear ERH signal after siYTHDC1 followed
1278 by washout, in human fibroblasts.

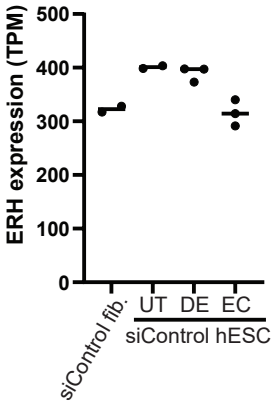
1279

1280

Katznelson et al., Extended Data Figure 10

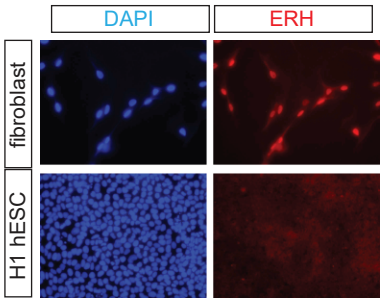
a

ERH expression by RNA-seq in human fibroblasts and human hESCs



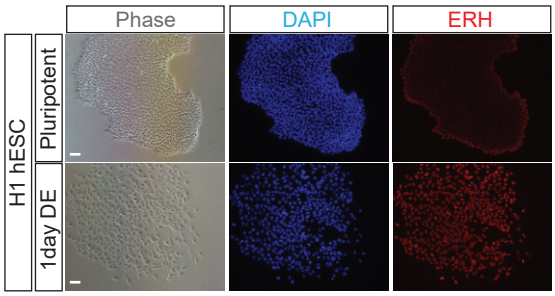
b

ERH in human fibroblasts and H1 hESCs stained in parallel, imaged using same settings



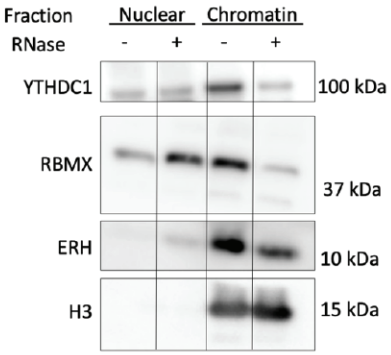
c

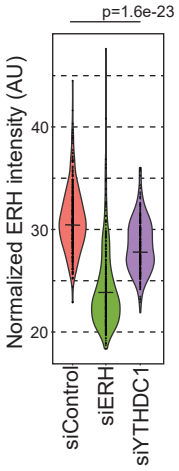
ERH in H1 hESCs in pluripotent and 1 day endoderm differentiaiton conditions, stained in parallel, imaged using same settings



d

RNase A treatment of soluble nuclear and chromatin fractions



e

Quantification of ERH intensity by immunofluorescence after permeabilization and washout for indicated siRNA