**Title**

Heterochromatin protein ERH represses alternative cell fates during early mammalian differentiation

**Author list and affiliations**

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

*These authors contributed equally

[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

K. Zaret, zaret@pennmedicine.upenn.edu

R. McCarthy, Ryan.McCarthy@seattlechildrens.org

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

**Summary**

During development, H3K9me3 heterochromatin is dynamically rearranged, silencing repeat elements and protein coding genes to restrict cell identity. Enhancer of Rudimentary Homolog (ERH) is an evolutionarily conserved protein originally characterized in fission yeast and recently shown to be required for H3K9me3 maintenance in human fibroblasts, but its function during development remains unknown. Here, we show that ERH is required for proper segregation of the inner cell mass and trophectoderm cell lineages during mouse development by repressing totipotent and alternative lineage programs. During human embryonic stem cell (hESC) differentiation into germ layer lineages, ERH is crucial for silencing naïve and pluripotency genes, transposable elements, and alternative lineage genes. Strikingly, ERH depletion in somatic cells reverts the H3K9me3 landscape to an hESC state and enables naïve and pluripotency gene and transposable element activation during iPSC reprogramming. Our findings reveal a role for ERH in initiation and maintenance of developmentally established gene repression.

## Introduction

Heterochromatin marked by H3K9me3 represses the expression of many protein-coding genes and DNA repeat elements[1–3]. Yet despite its stability, H3K9me3-heterochromatin is dynamically established and rearranged during embryogenesis[1,4]. Following fertilization, the human zygote divides three times to 8 cells while the cells remain totipotent - possessing the ability to contribute to all embryonic and extraembryonic cell lineages[5]. Zygotic genome activation (ZGA), marked by initial transcription of key genes and transposable elements[6], occurs at the 1-2 cell stage in mice and the 4-8 cell stage in humans[6], and requires demethylation of H3K9me3 marks in oocytes[7,8]. Protein-coding genes and a subset of repeats activated in ZGA are subsequently silenced through H3K9me3 establishment as the embryo loses totipotency and differentiates into the first two developmentally-specified lineages, the inner cell mass (ICM) and the trophectoderm, forming the blastocyst[4,9,10]. The trophectoderm will form the placenta, while the pluripotent cells of the ICM will further segregate to specify the primitive endoderm, which will form the extraembryonic endoderm, and the epiblast, which will give rise to the germ layers ectoderm, mesoderm, and endoderm. As development proceeds, H3K9me3 is established on lineage-inappropriate genes and repeat elements [4,9,11], with the highest number of genes marked by H3K9me3 at the germ-layer stage[1]. Throughout germ layer specification to mature somatic lineages, H3K9me3 is removed from genes that become active and deposited onto early developmental genes that become repressed[1]. Genes marked by H3K9me3 are refractory to transcription factor binding, preventing activation of alternative lineage genes, thereby reinforcing cell identity[3,12,13]. The mechanisms operating upstream in the initial targeting of histone methyltransferases to deposit H3K9me3 and the maintenance of these developmentally established H3K9me3 domains remain poorly understood. In this study, we investigate how *enhancer of rudimentary homolog* (ERH), a primary protein factor other than histone methyltransferases, contributes to H3K9me3 dynamics and gene repression during early stages of mammalian development.

ERH is an evolutionarily conserved protein that was originally discovered in fission yeast to regulate sporulation, the primary form of cell differentiation, via H3K9 methylation [14,15]. We subsequently found that ERH is a dominant regulator of H3K9me3 maintenance at genes and repeat elements in human fibroblasts [16]. In *S. pombe,* Erh1 forms a complex with the RNA binding protein Mmi1, which is then recruited by binding of nascent meiotic transcripts[14]. Through RNAi-independent mechanisms, Erh1-Mmi1 causes the degradation of its bound RNA transcripts, the establishment of H3K9 methylation, and consequent repression of meiotic and developmental genes, transposons, and rDNA[14,15]. In *S. pombe,* RNAi-dependent H3K9me3 requires an RNA-dependent RNA polymerase[17], however RNA-dependent RNA polymerases are not found in vertebrates[18]. In a remarkable example of evolutionary conservation, we recently found that ERH in terminally-differentiated human cells represses meiotic and germ cell genes, and further expanded its function in mammals to include the majority of regions that exhibit H3K9me3 heterochromatin, including genes and repeat elements[16]. However, the role of ERH in a developmental context in mammalian cells is unknown.

In this study, we investigate the role of ERH in gene repression to regulate cell fate and lineage commitment during stages of early cell lineage segregation. Specifically we focus on developmental stages where repression by H3K9me3 has been implicated - during trophectoderm and inner cell mass segregation in the blastocyst [4,9,10], pluripotency to germ layer lineage differentiation[1], and maintenance of repression in terminally-differentiated somatic cells[19]. By genetic perturbation, imaging, and gene expression profiling we identify a conserved role for ERH across mouse and human systems in the repression of early naïve and alternative lineage genes. From these findings we reasoned that ERH may also be required to maintain repression in terminally-differentiated cells of developmentally repressed genes. To investigate this hypothesis, we depleted ERH during somatic reprogramming to iPSC and identified regions of ERH-dependent H3K9me3. We find that ERH is necessary for the maintenance of the somatic H3K9me3 state in fibroblasts and upon ERH loss, there is reversion to a pluripotency-

like H3K9me3 state which is permissive to binding of pluripotency transcription factors, which in turn activate naive and alternative lineage genes and repeat elements. Together, our findings demonstrate a broad and fundamental role for ERH during developmentally acquired H3K9me3-mediated gene repression and maintenance of this repression in somatic cells.

## Results

### ERH is necessary for trophectoderm development, the first embryonic differentiation

To explore ERH functions during pre-implantation development, we first analyzed publicly available RNA-seq data of mouse embryogenesis timepoints from zygote to blastocyst[20]. ERH mRNA is present at all stages of pre-implantation development, but significantly increased from the 8-cell stage and remains highly expressed in the blastocyst (Figure 1A). We then immunostained mouse embryos for ERH protein at the zygote, 2-cell, 8-cell and blastocyst stages of pre-implantation development, which revealed that ERH is expressed and nuclear-localized at all stages analyzed, with increasing abundance in the early and late blastocyst stages (Figure 1B, Fig. S1A), consistent with its mRNA levels (Figure 1A). At the blastocyst stage, ERH levels are similar between cells in the ICM and trophectoderm (Figure S1A).

To investigate if ERH is involved in lineage specification during pre-implantation development, we depleted ERH levels by siRNA micro-injection at the 1-cell stage of mouse embryos, compared to a nontargeting siRNA control, and observed specific ERH depletion at the protein level by the blastocyst stage (Figure 1C). ERH depletion did not cause major morphological changes in blastocyst structure, as both siControl and siERH blastocysts form a cavity and segregate ICM and trophectoderm cells (Figure S1B-C). Staining blastocysts for key transcription factors involved in lineage specification revealed similar levels of the pluripotency marker OCT4 in ICM and TE cells between siControl and siERH blastocysts (Figure S1C). However, the levels of the trophectoderm-determining transcription factor CDX2[21] were significantly reduced in the trophectoderm (Figure 1D,E). We further validated the

trophectoderm-specific gene expression defect in ERH-depleted blastocysts by staining for the trophectoderm marker Keratin 8 (K8)[22], which was also significantly decreased (Figure 1F,G Thus, ERH is required for the establishment of proper trophectoderm differentiation.

**ERH maintains H3K9me3 levels and gene repression in 2-cell embryos and blastocysts**

The role of ERH in H3K9me3 heterochromatin regulation in *S. pombe*[14,15] and human fibroblasts[16] suggests that alterations in gene silencing mechanisms may underlie the defect in trophectoderm differentiation found in ERH knockdown embryos. Indeed, we observed a 32% decrease (p=0.0025) in H3K9me3 in blastocysts of ERH-depleted embryos (Figure 2A,B, Table S1).

To determine which genes may be impacted by ERH depletion, we performed RNA-seq on mouse blastocysts injected at the 1-cell stage with siControl or siERH (Table S 2); which confirmed an ~80% knockdown of ERH (Figure S2A). Gene Ontology (GO) analysis of upregulated genes indicated an activation of meiotic genes (Figure 2C; Table S 3), consistent with the role of ERH as a conserved repressor of meiotic genes via H3K9me3 in both *S. pombe*[14] and humans[16]. Notably, upon ERH depletion we observed preferential activation of DNA methylation-sensitive germline genes recently classified as being repressed during preimplantation development by *SETDB1, PRC1.6/RING1*[23]*,* and CpG methylation during post-implantation embryos[24] (Figure S2B).

We classified genes upregulated in ERH-depleted blastocysts into 4 groups, based upon their expression during normal preimplantation development (Figure 2D). Group 1 consists of transcripts present at the zygote stage, including maternally inherited transcripts, which become silenced or degraded as development progresses. Group 2 consists of transcripts highly expressed in the 2 to 8 cell stages and low or absent in the zygote and blastocyst stage. Group 3 consists of transcripts absent or lowly expressed at all stages measured. Group 4 consists of transcripts with highest expression in the blastocyst stage. Of the ERH-repressed genes (i.e.,

genes whose expression goes up upon ERH knockdown), 25 normally have peak expression during the 2 to 8 cell stage (group 2) and 54 are normally not expressed during preimplantation development (group 3). Among the group 2 genes was *Pou6f1*, whose degradation by the RNA exosome was shown to be necessary for *Cdx2* transcription during mouse blastocyst formation[25]. We observed a failure to repress, at the blastocyst stage, many 2 cell-specific genes activated during zygotic genome activation[10] (Figure S2C), including upregulation of all nine *Zscan4* paralogs in two of three siERH replicates (Figure S2D), which are typically expressed exclusively at the 2 cell stage[26]. Genes repressed by ERH in group 3 are typically expressed in diverse lineages (Figure 2E). The results indicate that ERH is critical for repression of a subset of 2 cell specific genes as well as genes normally repressed during blastocyst formation and segregation of ICM and trophectoderm cell lineages.

We hypothesized that if the ERH-dependent gene repression is mediated by H3K9me3, we would expect that ERH-repressed genes would gain H3K9me3 during preimplantation development. By analyzing published H3K9me3 ChIP-seq datasets[9] and comparing the H3K9me3 status in 2-cell vs trophectoderm and ICM cells, we found that a subset (34 of 149) of the genes repressed by ERH in mouse blastocysts (DESeq2, adjusted p>=0.05, log2 fold change >=0.5) normally gain H3K9me3 in the trophectoderm or ICM (Figure 2F). Examples of H3K9me3 and gene repression changes induced by ERH depletion status in group 1 genes include *Slc25a31* a metabolic gene expressed in oocytes and sperm[27]; in group 2 genes include *D1Pas1* previously identified as a 4 cell specific gene[28] and in group 3 genes include *Cyp46a1* a brain specific neuroprotective enzyme for cholesterol degradation[29](Figure S2E). Nearly half of the ERH repressed genes have been previously annotated as gaining H3K9me3[1] during mouse ICM to germ layer development (Figure S2F), more than observed for the genes downregulated in siERH blastocysts (Figure S2G). ERH- repressed genes were enriched for H3K9me3 peaks in ENCODE histone modification data from multiple mouse samples[30] (Figure S2H, top 7 p-

value mouse datasets shown). We conclude that aberrant gene activation observed in ERH-depleted mouse blastocysts is due to a failure to properly establish H3K9me3 silencing.

We further examined how the depletion of ERH may impact the specification of the ICM to epiblast and primitive endoderm. GO analysis of downregulated genes, during ERH depletion, indicated that early differentiation genes, including key markers of primitive endoderm such as Gata4 and Sox17 (Fig S2I), were enriched among the downregulated genes (Figure 2G; Table S 3), indication a failure to activate the primitive endoderm transcriptional network. Conversely, key epiblast genes such as Nanog and Tdgf1 were not significantly changed by ERH depletion (Figure S2J). We conclude that ERH is necessary to silence a subset of 2-8 cell stage genes as well as preventing spontaneous activation of later developmental genes during the blastocyst stage.

**ERH is necessary to silence naïve and pluripotency genes during human ESC differentiation**

ERH depletion in H1 hESCs cultured in pluripotency-maintaining conditions caused no observable changes in morphology or abundance of pluripotency factors (Figure 3A, Figure S3A). H3K9me3 (Figure 3B) levels were not changed in siERH treated hESCs; reminiscent of the partial effect of siERH on blastocysts (Figure 2B). We also observed no change in H3K27me3 in human ES cells treated with siERH (Figure S3A). Despite no apparent disruption of the pluripotent state or global H3K9me3 levels, we observed an increase in the area of the nucleus and increased nuclear blebs in ERH depleted hESCs (Figure S3B), which are changes associated with chromatin decompaction[31]. ERH depletion in self-renewing hESCs resulted in substantial differential expression of genes (Figure S3C; Table S 2), despite not disrupting pluripotency per se (Figure 3A). Global analysis of genes identified as upregulated in naïve vs primed hESCs[32] showed that ERH depletion in hESCs caused upregulation of genes associated with naïve pluripotency (Figure 3C).

Next, we wished to determine if ERH depletion impacted the ability of hESCs to differentiate. Activin A and the GSK-3 Inhibitor CHIR99021 together are capable of differentiating hESCs to an endoderm identity[33]. As expected, within 24 hours of treatment of undifferentiated H1 hESCs, we observed repression of pluripotency factors and initial expression of endodermal markers in a subset of cells (Figure S3D). Yet depletion of ERH in H1 hESCs followed by 24 hour endoderm differentiation showed elevated NANOG protein levels (Figure 3D; Figure S3E) and reduced emergence of SOX17+ cells (Figure 3C). To assess if this observation was lineage specific, we also performed differentiation to ectoderm, with control and ERH depleted cells, for 2 days with retinoic acid treatment. We observed sustained expression of the pluripotency factor OCT4 and decreased abundance of the ectodermal transcription factor PAX6 (Figure 3E).

To investigate genome-wide expression changes caused by ERH depletion, we performed RNA-seq of siRNA control and ERH depleted H1 hESCs after 1 day of endoderm or 2 days of ectoderm differentiation (Figure S3F; Table S 4). Notably, genes normally repressed in the siControl endoderm and ectoderm differentiations were preferentially upregulated by ERH depletion during endoderm (Figure S3G) and ectoderm (Figure S3H) differentiation. During endoderm differentiation of hESCs we observed that ERH depletion preferentially upregulated genes identified as higher in naïve vs primed hESCs[32] (Figure 3F), beyond what was observed in the pluripotent state (Figure 3C). For example, the pluripotency factor NANOG, previously demonstrated to be regulated by H3K9me3[34], was significantly upregulated by ERH depletion in all conditions (Figure 3G), commensurate with our immunofluorescence observations (Figure 3D, Figure S3D). Additional genes recently published as expressed in human pluripotent stem cells with a high capacity for self-renewal[35] were enriched in genes upregulated during endoderm and ectoderm differentiation after ERH depletion, including *DUSP5/6, NODAL, LEFTY1/2*, and *DACT1* (Figure S3I; Table S 2). Genes associated with 2 cell stage and naïve pluripotency including *ZSCAN4*[26], *KLF5*[36], *ESRRB*[37], *NR5A2*[38] and *NANOGP1*[39] were

upregulated during siERH treated hESC differentiation (Figures 3H and S3J). We conclude that ERH is required to repress pluripotency and maintain repression of naïve genes during hESC pluripotency exit.

**ERH regulates transposable elements with known cis-regulatory function in human ESCs**

Regulation of developmental gene networks are driven in part by transposable elements that act as regulatory regions for transcription factor binding[36,40–43]. To assess if transposable elements are repressed by ERH in hESC during pluripotency and differentiation, we analyzed transposable elements in RNA-seq data from siControl and siERH treated H1 hESCs during self-renewal and differentiation (Table S 5). Transposable elements upregulated by ERH knockdown in H1, definitive endoderm, and ectoderm states included several classes of L1 elements and the hominid-specific[44] SINE VNTR and Alu (SVA) elements (Figure 4A). Satellite repeats were also derepressed (Figure 4A). Differentiation-specific ERH-repressed transposable element classes, shared between definitive endoderm and ectoderm conditions, included human morula-blastocyst specific *LTR7Ys*[36] and human 8 cell stage expressed *HERVKs*[45] (Figure 4A). After stratifying L1 elements by their evolutionary age[46], we found that ERH preferentially represses evolutionarily young L1 repeats in pluripotent and differentiating hESCs [46] (Figure 4B). Among the most upregulated young repeats upon ERH depletion during differentiation are the *LTR7Ys* (Figure S4A), which are normally bound by NANOG and KLF5, both of which are also upregulated by ERH knockdown (Figure 3E,F) to increase naïve gene expression[36]. Also upregulated by siERH were 8 cell stage *HERVK9-int* and *HERVK11* transposable elements and their LTR regulatory element *LTR5_Hs* (Figure S4B) as well as *SVA* transposable elements (Figure S4C) which are specific to the naïve state[40,47].

Next, we wished to identify instances of apparent co-expression of genes and nearby repeat elements. By filtering our data for genes expressed during pluripotency that exhibit ERH-dependent repression during differentiation, we identified instances with proximal evolutionarily

young transposable elements that also exhibited induction in siERH conditions (Figure 4C and S4D-E). Globally, the locus specific expression changes in siERH endoderm of L1 repeats correlated with the gene expression of the closest gene within 25 kb (Figure S4F), demonstrating coregulation of genes and proximal L1 repeats. The ERH repressed L1 elements gain H3K9me3 during pluripotency to endoderm differentiation (Figure 4D). This contrasts with H3K9me3 levels at gene bodies of genes upregulated in siERH DE vs siControl DE, which only gain a modest increase in H3K9me3 levels during pluripotency to endoderm differentiation (Figure S4G); indicating the role of H3K9me3 and ERH at this cell fate change occurs at transposable element regions and not gene bodies. We conclude that ERH silences repeat elements that apparently function as regulatory elements for promoting pluripotency and naïve states.

**ERH restricts lineage commitment during human ESC differentiation**

During differentiation, a key function of H3K9me3 heterochromatin is to repress lineage-incompatible genes[1,9]. Since we observed an upregulation of lineage-incompatible genes in ERH knockdown blastocysts, we next investigated if, during hESC differentiation, ERH depletion causes defects in restricting alternative lineage genes. GO analysis of genes upregulated by ERH depletion during endoderm and ectoderm differentiation confirmed activation of transcriptional networks typically not induced by the specific differentiation treatment (Figure 5A and S5A; Table S6). ERH depletion upregulated meiotic and gametogenic genes in all tested developmental stages (Figure S5B), consistent with previous observations in human fibroblasts[16]. To further assess the cell lineages in which the upregulated genes function, we generated composite expression profiles of the genes upregulated by ERH depletion during endoderm (Figure 5B) and ectoderm differentiation (Figure 5C) from expression data downloaded from the BBI Descartes human single cell atlas[48], and found that heart, muscle and liver genes are preferentially activated. Directed differentiation of hESCs with Activin A and

WNT activation, which typically induces definitive endoderm (DE) at high efficiency[33], instead induced both early and late mesodermal genes[49] when ERH was depleted (Figure 5D-E). Notable precocious gene activation in siERH definitive endoderm included *NXK2-5, NPPB* and *FLNC*, typically expressed in hearth muscle, as well as *TAGLN,* a canonical smooth muscle marker (Figure 5E). Increased levels of BRACHYURY protein, a key determinant of early mesoderm[50], was confirmed by immunofluorescence (Figure 5F). Off target gene activation during ectoderm differentiation included genes typically activated in liver development (Figure S5C). Thus, while ERH appears not to be crucial to maintain hESC pluripotency, it is essential to restrict cell identity for proper germ layer differentiation.

To assess the role of ERH in the repression of lineage specific transposable elements during differentiation, we utilized recent annotations of these elements for the early cell lineages[43] (Figure 5G). Notably, *MER11A* and *MER11C* elements, which are silenced by H3K9me3 during in vitro endoderm differentiation[43], were significantly upregulated in siERH endoderm differentiation (Figure 5G and S5D). Transposable elements activated by specific transcription factors[41,43], such as *MER11C* by Brachyury and *LTR6B* by GATA6, demonstrated increased activation in the siERH endoderm where the transcription factors are expressed, but were not activated under pluripotent conditions or during ectoderm differentiation where the factors are silent (Figure 5G). We were able to map reads to specific repeat element sites and observed instances where the repeat element and adjacent genes were upregulated following ERH depletion in regions marked by H3K9me3 in hESC derived endoderm (Figure 5H and S5E). ERH therefore represses evolutionarily young and lineage-specific repeats, which in turn during differentiation repress pluripotency and alternative lineages.

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

As ERH is essential for proper cell identity after exit from pluripotency, yet is dispensable in the maintenance of hESCs, we assessed how it controls heterochromatin in a terminally-differentiated, somatic cell background. We previously found that megabase-sized H3K9me3 heterochromatin domains present in human primary fibroblasts are absent in hESCs[12]. We also found that in human fibroblasts, depleting ERH during reprogramming to hepatocyte-like cells (hiHeps) improved the activation of hepatic genes located in fibroblast H3K9me3 heterochromatin domains [16]. Given the role of ERH in maintaining the somatic regulatory program by repression of alternative lineages, we assessed how H3K9me3 is disrupted after ERH loss by CUT&RUN for H3K9me3 after siRNA knockdown in human primary fibroblasts (Figure 6A).

Consistent with previous studies, we observed that H3K9me3 is globally lost in human fibroblasts after ERH knockdown (Figure S6A). Assessment of H3K9me3 on a genome browser showed that while large megabase-scale domains were diminished, residual peaks were retained in the ERH knockdown (Fig 6B). Residual peaks were found enriched at LINEs, LTRs, and Satellite repeats (Figure S6B) and were also observed at 3' UTRs of ZNF genes (Figure S6C). Interestingly, the residual H3K9me3 peaks highly overlap with H3K9me3 sites in H1 ESCs (Figure 6B, C). We validated that the residual sites are bona-fide marked by H3K9me3, rather than another chromatin protein with antibody cross-reactivity, by showing overlap with ESC peaks of HP1a (Figure 6C, D, S6D). A Pearson correlation of H3K9me3 peaks from BJ fibroblasts after siControl or siERH, as well as in H1-hESCs, supports a shift in the H3K9me3 heterochromatin landscape to an embryonic state (Figure 6E). Residual H3K9me3 peaks were highly enriched for SOX and ZNF binding motifs (Figure S6E). We conclude that ERH is required to maintain the broad H3K9me3 domains established in somatic cells but does not impact H3K9me3 peaks that persist from the pluripotent state. We conclude that distinct mechanisms regulate the maintenance of the observed, ERH-dependent H3K9me3 domains and the ERH-independent residual H3K9me3 peaks.

**ERH-dependent H3K9me3 heterochromatin suppresses SOX2 binding and activation of naïve specific genes and repeats during OSKM reprogramming**

Since H3K9me3 represses iPSC reprogramming by blocking transcription factor binding[12], and pluripotency genes in H3K9me3 heterochromatin show delayed activation compared to genes in heterochromatin marked with H3K27me3 or neither mark (Figure S7A), we assessed the impact of ERH depletion in the context of initiating transcription factor binding and gene activation during induced pluripotent stem cell (iPS) reprogramming (Figure 6A). We chose to focus on a 48 hour time point after initiation of reprogramming by infection with sendai virus to express Oct4, Sox2, Klf4, and c-Myc (Figure 6A) as this time point typically proceeds major transcriptional changes[51] and the reprogramming transcription factors are unable to bind H3K9me3-marked hESC targets[12].

While the majority of SOX2 binding sites were shared between siControl and siERH fibroblasts+OSKM, 2,538 binding sites were gained (Figure S7B). We observed siERH unique SOX2 peaks at loci losing H3K9me3 after ERH depletion (Figure 7A), and that siERH induced SOX2 binding was elevated at sites previously annotated as resistant to reprogramming factor binding in fibroblasts[12] (Figure S7C). We asked if the enhanced SOX2 binding after ERH knockdown occurred at endogenous targets by comparing the genomic locations of siERH-gained SOX2 sites to published ChIP-seq peaks in cistromeDB[52,53]. By analyzing the SOX2 siERH-gained peaks in the cistromeDB toolkit[53], we found that these sites had the most similarity to ChIP-seq peak sets of SOX2 and its cofactors such as NANOG and OCT4 in embryonic stem cells (Figure S7D). Inspection of genome browser tracks identified specific loci where siERH-gained SOX2 binding corresponded to hESC SOX2 and NANOG binding, and occurred within domains of reduced H3K9me3 (Figure 7B; Figure S7E). iPSC genes embedded in H3K9me3 domains and normally refractory to activation in early reprogramming were generally and broadly activated after siERH (Figure S7F). Further, after only 48 hours of

reprogramming during ERH knockdown, genes associated with the human 8-cell like cell (8CLC) state[54] were activated, including *ZSCAN4, DUXA, DPPA4, ARGFX, TPRXL*[55]*, KLF5, TRIM49*, and the histone variant *H3.Y*[56] (Figure 7C; Figure S7G). We compared the activation levels of these genes in siERH+OSKM with published RNA-seq datasets[57] and found that *H3.Y, DUXA* and *TPRXL* do not normally get activated at similar levels to siERH+OSKM at any stage during iPSC reprogramming (Figure S7H).

Analysis of evolutionarily young L1 repeats, which we observed to be repressed by ERH during hESC pluripotency and differentiation (Figure 3B and 3E), showed that ERH was also critical for maintaining repression of the repeats during iPSC reprogramming (Figure 7D). Furthermore, significant correlation was observed between the activation of L1 repeats and L1 proximal genes (Figure S7I). We observed that *LTR7* and *HERVH* repeats, which are expressed in and required to maintain pluripotency[58,59], demonstrated decreased H3K9me3 after ERH depletion and increased expression in siERH+OSKM (Figure 7E), some of which we also observed to be repressed by ERH during hESC differentiation (Figure S7J). *LTR7Y* repeats, which promote naïve pluripotency[36], exhibited similar ERH-dependent H3K9me3 and repression (Figure S7K). We observed instances where genes and adjacent *LTR7Y* repeats upregulated by ERH depletion during hESC differentiation were also upregulated by siERH+OSKM (Figure 7F; Figure S7L). Notably, residual H3K9me3 peaks retained at *LTR7Y* repeats were insufficient to maintain repression (Figure 7F). Furthermore, we observed activation upon ERH+OSKM of repeat elements recently demonstrated to be active specifically during the human 8 cell-like cell stage and later H3K9me3 marked and silenced in human blastocysts[4] (Figure S7M). These results demonstrate that ERH maintains H3K9me3 repression of a subset of 8-cell stage and pluripotency genes and repeats in terminally differentiated human cells.

Despite increased activation of pluripotency and 8-cell genes at 48 hours after initiation of OSKM-based reprogramming, during siERH, by 7 days we observed substantial cell death (Figure S8A). This is in contrast to our previous published observations where ERH depletion

without the addition of reprogramming factors and with the addition of hepatic, rather than iPS, reprogramming factors did not lead to marked cell death or block successful reprogramming[16]. The extent of cell death was lessened by decreasing levels of ERH siRNA but did not change the number of emerging iPSC colonies (Figure S8B), indicating that ERH is necessary for later stages of iPSC reprogramming, despite promoting a barrier to initial gene activation. It has recently been shown that activation of 8 cell-like state in pluripotent cells also induces cell death[54], suggesting that the repression of the 8 cell program by ERH may be necessary for survival during iPSC reprogramming.

In addition to activation of pluripotency and 8 cell-like cell genes with siERH+OSKM, we also observed activation of alternative lineage genes, especially those associated with neuronal lineages (Figure 7G). The neuronal gene signature was reprogramming factor-dependent, as it was not observed with ERH depletion in fibroblasts alone, in the absence of the OSKM reprogramming factors (Figure 7G; data for noTF condition downloaded from[16]). Analysis of lineage specific transposable element[43] expression, after ERH knockdown during OSKM reprogramming, demonstrated widespread activation (Figure S8C) potentially acting in the precocious activation of nearby alternative lineage genes such as the expression of oligodendrocyte *OLMALINC*[60] (Figure S8D). We conclude that ERH maintains repression of early 8-cell, pluripotency, and alternative lineage genes and repeats in human fibroblasts by maintaining H3K9me3 heterochromatin domains.

## Discussion

We investigated the role of ERH in gene repression during the earliest stages of mammalian lineage segregation and in the maintenance of repression of pluripotency and alternative lineage genes in terminally differentiated cells. H3K9me3 is dynamically rearranged during preimplantation development[4,9] and later stages of cell lineage commitment[1] to repress expression of genes belonging to alternative cell fates, however the mechanisms that modulate

H3K9me3 profiles, beyond direct studies of the relevant methyltransferases, have been unclear. We that ERH controls H3K9me3 for repression of naïve and alternative lineage genes during the cell fate commitment of totipotent cells to ICM and trophectoderm in mouse blastocysts as well as of hESCs to the endoderm and ectoderm cell lineages, to enforce cell fate and prevent anomalous gene activation. The 33% loss of H3K9me3 signals in early embryos with an ERH knockdown is less than the 50% loss of H3K9me3 we observed during comparable ERH siRNA knockdowns in human fibroblasts[16], suggesting that H3K9me3 may be less dependent on ERH in early development than terminally differentiated somatic cells. Regardless, between the three systems we employed: mouse preimplantation development, hESC differentiation, and fibroblast iPSC reprogramming, we observe that ERH commonly repressed a shared set of naïve genes such as *ZSCAN4* across all systems, indicating that ERH not only is involved in the initial repression of this early transcriptional program but also in maintaining repression in terminally differentiated cells. We find it remarkable that ERH, which is the sole apparent mechanism for repressing meiotic genes in S. pombe, is conserved in repressing germ cell genes in mammals[16] and, based on our present study, has expanded to control diverse other types of developmental genes and stages.

In a developmental context, the failure to repress 2-cell, 8-cell and pluripotency genes after ERH knockdown results in delays of specific-lineage gene expression, including for primitive endoderm genes in mouse blastocysts and various endoderm and ectoderm genes during hESC differentiation. Furthermore, we demonstrated that ERH is necessary to maintain repression of alternative lineage genes, as ERH depletion elicited precocious expression of genes typical of later cell differentiation in mouse blastocysts and human hESCs as well as activation of neuronal genes during initial iPSC reprogramming. ERH-dependent developmental gene regulation likely occurs not only through changes in direct repression of genes but also through the regulation of transposable elements, which are canonically repressed by H3K9me3 and serve as key developmental regulatory elements[40,41,43]. We observed that evolutionarily

young transposable elements, including the human specific *LTR7Ys*[36], are especially upregulated by ERH depletion during hESC differentiation and iPSC reprogramming. Much of the dysregulation of lineage specific genes upon ERH depletion during hESC differentiation may be attributable to increased accessibility of transcription factor binding sites on de-repressed repeats[36,40,43].

A key finding of our analysis is that ERH depletion in human fibroblasts reverted the H3K9me3 chromatin landscape to an hESC-like state. This hESC-like H3K9me3 landscape functionally permits binding of the SOX2 factor in fibroblasts to endogenous hESC SOX2 targets, and activation of 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 these residual H3K9me3 peaks, which are enriched for zinc finger motifs, may be H3K9me3 nucleation sites and that ERH is essential for the spread of H3K9me3 from these sites to establish the typical broad domains observed in human somatic cells. That many residual H3K9me3 peaks in ERH depleted fibroblasts are seen in the pluripotent state suggests that the somatic H3K9me3 domains may be nucleated with a non-repressive H3K9me3 peak very early in development, only to be expanded to form a repressive domain at a developmentally appropriate time. Importantly, we observe increased activation of many genes and transposable elements that harbor a residual H3K9me3 peak, yet lose broad H3K9me3, indicating that the 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*[61–63], leaving further investigation needed into how H3K9me3 nucleation and spread occurs in multicellular organisms.

Comparing the role ERH plays in H3K9me3 and cell fate regulation in *S. pombe* and mammalian systems provides insight into how H3K9me3 heterochromatin has been adapted through evolution to the purpose of cell fate regulation in complex multicellular mammals. We

observe in mouse development and hESC systems that ERH is functionally required in multiple developmental contexts to repress meiotic and gametogenic genes, as it is in *S. pombe*[14], although the gene members targeted by ERH in this pathway has evolved and expanded in mammals. Our findings on meiotic gene repression are supported by the recent identification of ERH in a screen for proteins silencing germline genes in mESCs[64]. Other aspects of *Erh1* function in *S. pombe,* such as the targeting of RNAs for degradation by the RNA exosome[14], appear to be conserved as we observed upregulation of developmental genes typically downregulated by RNA exosome degradation[25]. Evolutionarily divergent functions of ERH may be driven by evolutionarily young transposable elements that retain RNA expression similar to how other classes of transposable elements have influenced the evolution of Krüppel associated box Zinc finger proteins DNA dependent targeting of H3K9me3[65,66].

We found that ERH participates in gene repression at the earliest stages of development and that it is essential for maintaining this repression even in terminally differentiated cells. Our findings advance the understanding of how heterochromatic repression is established and maintained during development and identify genes repressed by this mechanism. We propose that understanding mechanisms by which ERH is employed to direct H3K9me3 to specific loci in a developmentally coordinated fashion will be crucial to modulating developmental H3K9me3 rearrangements to achieve greater control of cell identity.

**Limitations of the study**

In this study, we identified a role for ERH in initiating and maintaining H3K9me3 heterochromatin in preimplantation embryonic development, primary germ layer specification, and terminally differentiated somatic cells. Due to challenges in genetically manipulating the 1 cell embryo and primary human fibroblasts, we employed siRNAs to deplete ERH protein, which limits our ability to study the immediate impact of ERH loss. As we observed H3K9me3 disruption and broad gene activation after siRNAs targeting ERH, we cannot exclude that some

transcriptional changes are due to secondary effects. While an siRNA approach allowed us to assess ERH in diverse developmental and cellular systems, we look forward to future studies applying acute degradation systems, which will further mechanistic insights of ERH function on heterochromatin.

**Acknowledgements**

R.L.M. was supported by an NIH NIDDK K01 postdoctoral fellowship DK117970-01. N.P. was supported by NIGMS (GM139970-01) and NICHD (HD102013-01A1). NIHGM36477 to K.S.Z.

**Author contributions**

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

**Declaration of Interests**

The authors have no conflicts of interest to declare.

**Figure titles and legends**

Figure 1. ERH is expressed during pre-implantation mouse development and critical for normal blastocyst development

(A) ERH RNA expression in reads per kilobase million (RPKM) during stages of mouse pre-implantation development, RNA-seq data from[20]. unpaired two tail t-test; n=6.

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

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

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

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

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

(G) Quantification of CDX2 intensity in siControl vs siERH blastocysts (unpaired t-test, n=9 blastocysts, n=8 blastocysts for each siRNA).

See also Figure S1 and Table S1.

Figure 2. ERH represses 2 cell and alternative lineage genes in mouse blastocysts

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

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

(C) GO analysis (PantherDB) reporting significantly enriched GO biological processes (FDR<=0.05) for genes significantly upregulated (DESeq2 adjusted p-value<=0.05, log2fc>=0.5) in siERH vs siControl mouse blastocysts.

(D) ERH repressed blastocyst genes grouped based upon normal expression patterns exhibited during mouse pre-implantation development re-analyzed from [20].

(E) Analysis showing diverse lineages where genes in group 3 are typically expressed, lineage annotations from BBI Descarts mouse development database[67].

(F) H3K9me3 profiles and heatmaps from 2-cell, trophectoderm and ICM[9] for genes significantly upregulated (adjusted p-value<=0.05, log2fc>=0.5) in siERH vs siControl mouse blastocysts.

(G) GO analysis (PantherDB[68]) reporting significantly enriched GO biological processes (FDR<=0.05) for genes significantly downregulated (adjusted p-value<=0.05, log2fc<=-0.5) in siERH vs siControl mouse blastocysts.

See also Figure S2 and Table S2,S3.

Figure 3. ERH is dispensable for pluripotency maintenance but is involved in silencing of pluripotency and naïve genes during differentiation

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

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

(C) Volcano plot showing differential gene expression in siERH vs siControl H1 untreated (UT) for genes annotated in[32] as upregulated in naïve vs primed hESCs. Inset pie chart indicates the number of genes upregulated and down regulated in siERH vs siControl (DESeq2, padj<=0.05, log2fc>0.5).

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

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

(F) Volcano plot showing differential gene expression in siERH vs siControl 1 day endoderm for genes annotated in[32] as upregulated in naïve vs primed hESCs. Inset pie chart indicates the number of genes upregulated and down regulated in siERH vs siControl (DESeq2, padj<=0.05, log2fc>0.5).

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

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

See also Figure S3 and Table S4.

Figure 4. ERH represses transposable elements that act as gene regulatory regions

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

(B) Expression change and significance, TEtranscripts[69] analysis, of L1 LTR transposable elements in by ERH depletion in H1 hESCs untreated (UT), 1 day endoderm (DE) or 2 day ectoderm (EC) listed in order of evolutionary age as determined in[46].

(C) Browser track showing RNA-seq of example evolutionarily young L1PA3 repeat and nearby gene in siControl and siERH treated pluripotency hESCs, 1 day DE and 2 day EC.

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

See also Figure S4 and Table S5.

Figure 5. ERH represses alternative lineage gene activation during hESC differentiation

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

(B) Radar plot showing cumulative expression profile across indicated human tissues, data downloaded from BBI Descartes database[70], for genes upregulated (padj<=0.05, log2 fold change >= 0.5) in siERH vs siControl H1 hESCs 1 day endoderm differentiation.

(C) Radar plot showing cumulative expression profile across indicated human tissues, data downloaded from BBI Descartes database[70], for genes upregulated (padj<=0.05, log2 fold change >= 0.5) in siERH vs siControl H1 hESCs 2 day ectoderm differentiation.

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

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

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

(G) Expression change and significance, TEtranscripts[69] analysis, of transposable elements by ERH depletion in H1 hESCs untreated (UT), 1 day endoderm (DE) or 2 day ectoderm (EC) grouped by lineage/stage at which they are typically expressed as determined in[43]. Negative log2 fold changes are represented as unfilled circles.

(H) Browser track showing RNA-seq of example repeat element and nearby genes in siControl and siERH treated H1 hESCs after 1 day of endoderm differentiation and

H3K9me3 tracks from H1 hESCs (GSM1003585) and hESC derived endoderm (GSM916057).

See also Figure S5 and Table S6.

Figure 6. ERH depletion reverts fibroblast H3K9me3 patterns to an hESC like state

(A) Schematic of siRNA treatments and timing for H3K9me3 CUT&RUN experiments shown in Figure 6 and SOX2 CUT&RUN and RNA-seq experiments shown in Figure 7.

(B) Browser track of H3K9me3 levels in siControl and siERH treated human fibroblasts and H1 hESCs (GSM1003585) at the multi-domain level, domain level and gene level. Example H3K9me3 peaks common between siERH fibroblasts and H1 hESCs are indicated by red triangles.

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

(D) Browser track comparing siControl and siERH fibroblast H3K9me3 with H1 hESC H3K9me3 patterns. Peak locations called by MACS2 analysis on siERH H3K9me3 data are shown.

(E) Correlation of H3K9me3 from BJ fibroblasts after siControl or siERH with H1 hESC H3K9me3.

See also Figure S6 and Table S7.

Figure 7. ERH depletion enhances gene activation of pluripotency and alternative lineage genes during iPSC reprogramming

(A) Heatmap of H3K9me3 CUT&RUN from siControl and siERH human fibroblasts and SOX2 CUT&RUN from siControl and siERH fibroblasts+OSKM at siERH unique SOX2 peak sites.

(B) Browser track of example hESC SOX2 and NANOG peak coincident with fibroblast siERH H3K9me3 loss and SOX2 peak gain.

(C) Expression of ERH and example naïve pluripotency genes activated in 48 hours OSKM by ERH depletion, transcripts per million (TPM) values from RNA-seq replicates shown (DESeq2 adjusted p-value).

(D) Expression change and significance, TEtranscripts[69] analysis, of L1 LTR transposable elements in by ERH depletion in fibroblasts+OSKM listed in order of evolutionary age as determined in[46].

(E) Profile plots of H3K9me3 and RNA levels from LTR7 and HERVH repeats in siERH vs siControl fibroblasts+OSKM.

(F) Browser track showing H3K9me3 in siControl and siERH fibroblasts and RNA levels in siControl+OSKM and siERH+OSKM for gene and adjacent LTR7Y naïve pluripotency repeat.

(G) Radar plot showing cumulative expression profile across indicated human tissues, data downloaded from BBI Descartes database, for genes upregulated (padj<=0.05, log2 fold change >= 0.5) in siERH vs siControl fibroblasts+OSKM and fibroblasts+NoTF[16].

See also Figures S7,S8 and Tables S8,S9.

## STAR Methods

### Key resource table

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| Anti-ERH antibody | Abcam | ab166620 |
| Anti-H3 (trimethyl K9) antibody | Abcam | ab8898 |
| Anti-H3 (trimethyl K27) antibody | Active Motif | 39155 |
| Anti-Keratin-8 antibody | DSHB | TROMA-I |
| Anti-CDX2 antibody | Invitrogen | ZC007 |
| Anti-Phalloidin-Alexa Fluor 555 | Invitrogen | A34055 |
| Anti-SOX2 antibody | Cell Signaling | 23064 |
| Anti-OCT4 antibody | Abcam | ab19857 |

| | | |
|---|---|---|
| Anti-PAX6 antibody | BD Biosciences | 561462 |
| Anti-Brachyury antibody | Cell Signaling | 81694 |
| Phalloidin-Alexa Fluor 555 | Invitrogen | A34055 |
| Goat anti-Rabbit Alexa Fluor 488 | Invitrogen | A32731 |
| Goat anti-Rat Alexa Fluor 488 | Invitrogen | A-11006 |
| Goat anti-Rabbit Alexa Fluor 647 | Invitrogen | A32733 |
| Goat anti-Mouse Alexa Fluor 647 | Invitrogen | A32728 |
| Donkey anti-Rabbit Alexa Fluor 594 | Invitrogen | A32754 |
| Donkey anti-Mouse Alexa Fluor 488 | Invitrogen | A32766 |
| Chemicals, peptides, and recombinant proteins | | |
| Activin A | Peprotech | 120-14P |
| CHIR99021 | Cayman Chemical | 13122 |
| Retinoic Acid | Cayman Chemical | 11017 |
| Matrigel | Millipore Sigma | CLS354277 |
| mTeSR1 | Stem Cell Tech | 85850 |
| DPBS | ThermoFisher | 14190144 |
| Eagle's Minimal Essential Medium | Millipore Sigma | M2279 |
| Fetal Bovine Serum | Hyclone | SH30071 |
| GlutaMAX | ThermoFisher | 35050061 |
| OptiMEM | ThermoFisher | 31985070 |
| Accutase | Stem Cell Tech | 07920 |
| Y-27632 ROCK inhibitor | Cayman Chemical | 10005583 |
| KSOM+AA | Merck | MR-121-D |
| M2 medium | Merck | MR-015-D |
| Lipofectamine RNAiMAX Transfection Reagent | ThermoFisher | 13778150 |
| TRIzol | ThermoFisher | 15596026 |
| Formaldehyde | Electron Microscopy Sciences | 15710 |
| Triton X-100 | Millipore Sigma | T8787 |
| DAPI | Millipore Sigma | 10236276001 |
| Mineral oil | Millipore Sigma | M-5310 |
| Bovine serum albumin | Gemini Bio | 700-100P |
| Donkey serum | Millipore Sigma | D9663 |
| Magnetic concanavalin A beads | Antibodies-online | ABIN6952467 |
| cOmplete, Mini, EDTA-free protease inhibitor cocktail | Millipore Sigma | 11836170001 |
| Digitonin | Millipore Sigma | D141-100MG |
| Spermidine | Millipore Sigma | S2626 |
| pA/G-MNase | EpiCypher | 15-1016 |
| RNase A | Millipore Sigma | 10109142001 |
| CytoTune-iPS 2.0 Sendai reprogramming kit | ThermoFisher | A16517 |
| Deposited data | | |
| Raw and analyzed data | This paper | GSE268901(RNAseq) GSE268961(CUT&RUN) |
| H3K9me3 ChIP-seq H1 hESC (referenced as #1) | ENCODE Project | GSM1003585 |
| H3K9me3 ChIP-seq H1 hESC (referenced as #2) | ENCODE Project | GSM1003585 |

| | | |
|---|---|---|
| H3K9me3 ChIP-seq hESC derived endoderm | Roadmap Epigenomics | GSM916057 |
| HP1alpha ChIP-seq H1 hESC | ENCODE Project | GSE123229 |
| SOX2 ChIP-seq in H9 hESC | Zhou et al.,[1] | GSE69479 |
| NANOG ChIP-seq H1 hESC | ENCODE Project | GSM803437 |
| Experimental models: Cell lines | | |
| Human: H1 hESC line | WiCell | WA01 |
| Human: BJ foreskin fibroblasts | ATCC | CRL-2522 |
| Experimental models: Organisms/strains | | |
| C57BL/6 mice | A*STAR | N/A |
| Recombinant DNA and RNA | | |
| Silencer Select Negative Control No. 1 siRNA | Thermo Fisher | 4390843 |
| Human ERH siRNA (Silencer Select) | Thermo Fisher | S4815 |
| AllStars negative control siRNA | Qiagen | 1027281 |
| Mouse ERH siRNA #1 | Qiagen | S104403294 |
| Mouse ERH siRNA #2 | Qiagen | S104403287 |
| Software and algorithms | | |
| R (v4.2.1) | R Core Team, 2022 | https://www.r-project.org/ |
| Bowtie2 (v2.3.4.1) | Langmead et al.,[2] | https://github.com/BenLangmead/bowtie2 |
| STAR (2.7.10b) | Dubin et al.,[3] | https://github.com/alexdobin/STAR |
| featureCounts | Liao et al.,[4] | https://rnnh.github.io/bioinfo-notebook/docs/featureCounts.html |
| Samtools (v1.16.1) | Danecek et al.,[5] | https://www.htslib.org/ |
| Bedtools (v2.27.1) | Quinlan and Hall[6] | https://bedtools.readthedocs.io/en/latest/ |
| DESeq2 (v1.38.3) | Love et al.,[7] | https://bioconductor.org/packages/release/bioc/html/DESeq2.html |
| ChIPSeqSpikeInFree (R v3.6.3) | Jin et al.,[8] | https://github.com/stjude/ChIPseqSpikeInFree |
| SEACR (v1.3) | Meers et al.,[9] | https://github.com/FredHutch/SEACR |
| CistromeDB toolkit | Zheng et al.,[10] | http://dbtoolkit.cistrome.org/ |
| Intervene | Khan et al.,[11] | https://github.com/asntech/intervene |
| ImageJ | Schneider et al.,[12] | https://imagej.net/ij/ |

| | | |
|---|---|---|
| Deeptools (v3.5.0) | Ramirez et al.,[13] | https://github.com/deeptools/deepTools |
| HOMER | Heinz et al.,[14] | http://homer.ucsd.edu/homer/motif/ |
| TETranscripts (v2.2.3) | Jin et al.,[15] | https://github.com/mhammell-laboratory/TEtranscripts |
| Igv | Robinson et al.,[16] | https://igv.org/ |
| Prism | GraphPad Software | https://www.graphpad.com/features |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact Ryan McCarthy (ryan.mccarthy@seatttlechildrens.org).

### Data and code availability

- RNA-seq and CUT&RUN datasets collected in this study have been deposited in the GEO of NCBI and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. 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).
- 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.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Cell lines culture and condition

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

**Mouse embryo work**

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

**METHOD DETAILS**

**Mouse embryo immunofluorescence and microscopy**

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

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

**Mouse embryo siRNA injection experiments**

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

**Differentiation of hESCs**

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

endoderm differentiation of hESCs, mTeSR1 medium was removed and cells were cultured in DMEM/F12 (Thermo Fisher, 11320033) supplemented with 10% (v/v) Fetal Bovine Serum (Hyclone, SH30071) with 100ng/ml Activin A (Peprotech, 120-14P) and 3uM CHIR99021 (Cayman Chemical, 13122). For ectoderm differentiation of hESCs, mTeSR medium was removed and cells were cultured in DMEM/F12 with 5uM Retinoic Acid (Cayman Chemical, 11017).

**Cell culture siRNA transfection experiments**

All ERH knockdown experiments were performed with two cycles of siRNA transfection (Silencer Select siRNA, Thermo Fisher, s4815), three days apart. Transfections were performed in fibroblasts using 5nM final concentration of siRNA and in hESCs using a 10nM final concentration of siRNA and 2.5ul/ml final concentration of Lipofectamine RNAiMAX (Thermo Fisher #31985070).

**Reprogramming factor expression experiments**

We utilized the CytoTune-iPS 2.0 Sendai reprogramming kit (ThermoFisher, A16517) to express Oct4, SOX2, KLF4 and cMyc in BJ fibroblasts. Cells were treated according to the manufacturer's instructions three days after the second of two cycles of siRNA transfection.

**Cell culture immunofluorescence and microscopy**

For immunofluorescence of human fibroblasts and hESCs cells were washed briefly with PBS, and fixed in 4% paraformaldehyde in PBS for 10 minutes. Fixed cells were washed 3 times with PBS, permeabilized with ice-cold 0.1% Triton-X in PBS for 10 minutes, and washed twice with PBS-T (PBS, 0.05% Tween-20). Samples were blocked with 3% donkey serum (Millipore Sigma, D9663) in PBS for 2 hours at room temperature and incubated overnight at 4$^{o}$C using the following concentrations: anti-H3K9me3 (1:300, Abcam ab8898), anti-H3K27me3 (1:300,

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

**Total RNA isolation and sequencing**

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

**CUT&RUN experiments**

CUT&RUN was performed as described in[71], with digitonin and NaCl in wash buffers optimized for permeabilization of BJ fibroblasts and retainment of SOX2 on chromatin as performed in[72]. Next, BJ fibroblasts were detached from cell culture plates with Accutase (STEMCELL

Technologies 07920), washed, bound to magnetic concanavalin A beads, and permeabilized with a dig-wash buffer (0.1% digitonin, 20 mM HEPES pH 7.5, 75 mM NaCl, 0.5 mM Spermidine, EDTA-free protease inhibitor). Bead-bound cells were incubated with a SOX2 (CST, clone D9B8N, 23064, 1:100) or H3K9me3 antibody (abcam ab8898, 1:100) at 4C overnight. The following morning, the cells were washed twice with dig-wash buffer, incubated with pA/G-MNase for an hour at 4C, and washed twice more. After chilling cells on an ice block for 5 minutes, MNase was activated for 30 minutes by addition of 2mM $CaCl_2$. The reaction was stopped by adding an equal volume of 2X STOP buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 0.1% Digitonin, 0.05 mg/mL glycogen, 5 mg/mL RNase A). Digested chromatin was extracted from permeabilized cells at 37C for 30 minutes, and DNA was purified by phenol-chloroform. Libraries were prepared as previously described[73]. Two to three biological replicates were prepared for each condition.

**CUT&RUN data processing**

Paired-end CUT&RUN reads were mapped to the human genome build hg38 using bowtie2[74] (v2.3.4.1) with run parameters --local --very-sensitive --no-unal --no-mixed --no-discordant -I 10 -X 1000. A quality-filtered BAM file was generated with the command samtools (v1.16.1) view -q 5 -f 2 -bS. Global scaling factors (SF) for siERH and siControl H3K9me3 data sets were calculated using the default parameters of ChIPSeqSpikeInFree[75] (R v3.6.3). Bigwig files for visualization were generated from indexed BAM files with command bamCoverage --smoothLength 10 --normalizeUsing CPM --scaleFactor 1/SF --ignoreForNormalization chrM (deeptools 3.5.0). BED files were generated from BAMs using the bedtools command bamToBed (bedtools v2.27.1). BED files were converted to bedgraph format using the bedtools command genomecov. Peaks were called on the bedgraph files with SEACR[76] (v1.3), using the stringent setting and selecting the top 0.01% of regions by AUC. Peak lists for downstream analyses were derived from individual replicates or the union of overlapping sites from biological

triplicates, as indicated in the text and figures. Enrichment analysis of known DNA binding motifs was performed using HOMER[77].

**Heatmaps**

Heatmaps were generated with deeptools[78] version 3.5.0. Counts matrices were generated with command computeMatrix reference-point --missingDataAsZero --referencePoint center. Figures were produced with the plotHeatmap command.

**Correlation plots**

To assess genome wide similarity between hBJ siERH H3K9me3, hBJ siControl H3K9me3, and H1-hESC H3K9me3 data sets, the Pearson correlation of peak overlap was calculated using the Intervene tool[79]. In brief, peak files for individual H3K9me3 replicates were processed with the command intervene pairwise --I *H3K9me3.bed --compute fract --htype color. Visualization of Pearson correlation of fractional overlap between peak sets were generated on https://intervene.shinyapps.io/intervene.

**RNA-seq data analysis**

RNA-seq reads (paired-end) were mapped to the mm10 or hg38 genome for mouse and human experiments respectively using STAR[80] (2.7.10b). The mapped reads to genes were counted using featureCounts[81]. For counting repeat mapping reads the TEcount function from TEtranscripts[69] was run on STAR aligned bam files using the hg38 repeatmasker gtf file annotation. The output count files were analyzed by DESeq2[82] to detect differentially expressed genes (Benjamini adjusted p-value < 0.05, log2 fold change > 1 or log2 fold change > 1.5). Genes and repeats with fewer than 10 total mapped reads were excluded from DESeq2 and TETranscripts analysis. TPM, FPKM and RPM were calculated using an in-house R script.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### 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[83]. 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[68] (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 p-values corrected for multiple testing using the Banjamini and Hochberg method. Reproducibility was confirmed by independent experiments.

**Supplemental information titles and legends**

Figure S1. Characterization of ERH and the effects of ERH depletion in mouse blastocysts

(A) Fluorescent imaging of ERH (magenta) and quantification in ICM and trophectoderm cells. (unpaired t-test, ns=not significant)

(B) Representative DIC images of mouse blastocysts injected at the 1 cell stage with siControl and siERH siRNAs, two examples for each shown.

(C) Fluorescent imaging of trophectoderm marker CDX2 (green), pluripotency marker OCT4 (magenta), with Phalloidin (orange) to define cell boundary and quantification of OCT4 levels in siControl siERH blastocysts with DAPI (blue). (t-test, ns=not significant).

Figure S2. ERH repression of 2-cell and developmental genes in mouse blastocysts.

(A) Expression of Erh in siControl and siERH treated mouse blastocysts by RNA-seq. (t-test, p>0.0001).

(B) Venn diagram showing the overlap in gene lists of mouse blastocyst siERH upregulated and downregulated genes with 137 DNA methylation sensitive germline genes identified by[24]. (hypergeometric test, p=0.64e-12).

(C) Venn diagram showing the overlap in gene lists of mouse blastocyst siERH upregulated genes classified into group2 (expressed in the 2- and 8-cell stages) with mouse ZGA genes annotated by[54]. (hypergeometric test, p=1e-10).

(D) TPM expression values from RNA-seq of 9 Zscan4 paralogs in siControl and siERH treated mouse blastocysts.

(E) Example gene tracks of genes significantly upregulated in siERH vs siControl blastocysts (adjusted p-value<=0.05, log2fc>=0.05) with H3K9me3 tracks for 2 cell, trophectoderm and ICM cells from[9].

(F) Pie chart showing the fraction of siERH upregulated genes (adjusted p-value<=0.05, log2fc>=0.05) which have been annotated as gaining H3K9me3[1] in 0-3 lineages of early endoderm, mesoderm or ectoderm.

(G) Same as (F) but for siERH downregulated genes (adjusted p-value<=0.05, log2fc<=-0.05).

(H) Top 7 results from Enrichr analysis based upon p-value of ENCODE histone modifications enriched at siERH upregulated genes (adjusted p-value<=0.05, log2fc>=0.05).

(I) TPM expression values from RNA-seq for primitive endoderm genes. (t-test, *p<=0.05)

(J) TPM expression values from RNA-seq for epiblast genes. (t-test, ns=not significant)

Figure S3. ERH is dispensable for pluripotency maintenance but is involved in silencing of pluripotency and naïve genes during differentiation

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

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

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

(D) Imaging of endoderm differentiation showing changes in colony morphology (phase), DAPI (blue), silencing of SOX2 (green) and induction of endoderm marker SOX17 (red). Scale bars, 100µm.

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

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

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

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

(I) Intersection of genes enriched in hESCs with high self-renewal[35] with genes repressed by ERH during endoderm and ectoderm differentiation. (hypergeometric test)

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

Figure S4. ERH represses transposable elements that act as gene regulatory regions

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

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

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

(D) Browser track showing RNA-seq of example repeat element and nearby gene that is expressed in pluripotency, silenced in siControl DE and EC but expressed in siERH DE and EC conditions.

(E) Browser track showing RNA-seq of example repeat element and nearby gene that is expressed in pluripotency, silenced in siControl DE and EC but expressed in siERH DE condition.

(F) Plot of expression change of specific L1 element loci vs expression change of corresponding nearest gene within 25kb in 1 day endoderm. R, Pearson correlation.

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

Figure S5. ERH represses alternative lineage gene activation during hESC differentiation

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

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

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

(D) MER11A and MER11C expression plots from RNA-seq data analyzed by TETranscripts. (DESeq2 adjusted p-value).

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

Figure S6. ERH depletion reverts fibroblast H3K9me3 patterns to an hESC like state

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

(B) Occurrence of repeats from RepEnrich database, and repeat occurrence at sites of siERH fibroblasts H3K9me3 peaks.

(C) Example browser track of siERH H3K9me3 peak at 3' region of a ZNF gene.

(D) Browser tracks showing H3K9me3 (GSM1003585) and HP1a (GSM1701825) in H1 hESCs and H3K9me3 in siControl and siERH human fibroblasts.

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

Figure S7. ERH depletion enhances gene activation of pluripotency

(A) Expression of genes activated during fibroblasts to iPSC reprogramming timepoints[57] that are located in unmarked, H3K27me3 or H3K9me3 marked heterochromatin as annotated by[3].

(B) Venn diagram showing overlap between SOX2 peak calls in siControl and siERH fibroblasts+OSKM.

(C) SOX2 CUT&RUN average profile in siControl and siERH fibroblasts+OSKM for genome regions identified as being differentially bound by reprogramming transcription factors [12].

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

(E) Browser track of example hESC SOX2 and NANOG peak coincident with fibroblast siERH H3K9me3 loss and SOX2 peak gain.

(F) Volcano plot of expression changes (DESeq2) in siERH vs siControl fibroblasts+OSKM of iPSC activated genes located in fibroblast H3K9me3 domains.

(G) Browser track of H3K9me3 in siControl and siERH fibroblasts and RNA in siControl and siERH fibroblasts+OSKM at DPPA4.

(H) Example 8 cell genes activated in siERH+OSKM but not activated during eight time points of iPSC reprogramming time course[57].

(I) Plot of expression change of specific L1 element loci vs expression change of corresponding nearest gene within 25kb in fibroblasts+OSKM. R, Pearson correlation.

(J) Browser tracks of H3K9me3 in siControl and siERH fibroblasts and RNA in siControl and siERH fibroblasts+OSKM, untreated H1, 1 day DE and 2 day EC at a locus of an LTR7 repeat.

(K) Profile plots of H3K9me3 and RNA levels from LTR7 and HERVH repeats in siERH vs siControl fibroblasts+OSKM.

(L) Browser track of H3K9me3 in siControl and siERH fibroblasts and RNA in siControl and siERH fibroblasts+OSKM, untreated H1 and 1 day DE at an LTR7Y-HERVH transposable element and nearby gene.

(M) Volcano plot showing differential repeat expression in siERH vs siControl fibroblast+OSKM for repeats annotated in[4] as repressed by H3K9me3 in human blastocysts. (TETranscripts, DESeq2, padj<=0.05, log2fc>0.5).

Figure S8. ERH depletion activates alternative lineage genes and repeats during iPSC reprogramming

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

(B) Plot of iPSC colony count from siControl and siERH fibroblasts+OSKM cells 3 weeks after initial OSKM infection with titrated siRNA concentration. Each value represents the number of colonies generated in 1 well of a 24 well plate.

(C) Expression change and significance, TEtranscripts[69] analysis, of transposable elements by ERH depletion in fibroblasts+OSKM grouped by lineage/stage at which they are typically expressed as determined in[43]. Negative log2 fold changes are represented as unfilled circles.

(D) Browser track of H3K9me3 in siControl and siERH fibroblasts and RNA in siControl and siERH fibroblasts+OSKM at an L1PA4 repeat element and nearby neuronal gene.

**Supplemental Tables**

Table S1. Blastocyst immunofluorescence quantification values.

Table S2. Blastocyst RNA-seq DESeq2 siERH/siControl differential gene expression analysis results.

Table S3. Blastocyst siControl vs siERH GO analysis results.

Table S4. H1 hESC untreated, endoderm and ectoderm RNA-seq DESeq2 siERH/siControl differential gene expression analysis results.

Table S5. H1 hESC untreated, endoderm, and ectoderm TETranscripts siERH/siControl differential repeat expression analysis

Table S6. H1 hESC untreated, endoderm and ectoderm GO analysis results.

Table S7. H3K9me3 peak coordinates in siERH human fibroblasts.

Table S8. Fibroblast +OSKM 48 hours, siERH/siControl RNA-seq DESeq2 differential gene expression analysis results.

Table S9. Fibroblast +OSKM 48 hours, siERH/siControl TETranscripts differential repeat expression analysis results.

## References


1. Nicetto, D., Donahue, G., Jain, T., Peng, T., Sidoli, S., Sheng, L., Montavon, T., Becker, J.S., Grindheim, J.M., Blahnik, K., et al. (2019). H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification. Science *363*, 294–297. https://doi.org/10.1126/science.aau0583.

2. Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R.A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes & development *23*, 2484–2489. https://doi.org/10.1101/gad.1837309.

3. Becker, J.S., McCarthy, R.L., Sidoli, S., Donahue, G., Kaeding, K.E., He, Z., Lin, S., Garcia, B.A., and Zaret, K.S. (2017). Genomic and Proteomic Resolution of Heterochromatin and Its Restriction of Alternate Fate Genes. Mol Cell *68*, 1023-1037.e15. https://doi.org/10.1016/j.molcel.2017.11.030.

4. Xu, R., Li, S., Wu, Q., Li, C., Jiang, M., Guo, L., Chen, M., Yang, L., Dong, X., Wang, H., et al. (2022). Stage-specific H3K9me3 occupancy ensures retrotransposon silencing in human pre-implantation embryos. Cell Stem Cell *29*, 1051-1066.e8. https://doi.org/10.1016/j.stem.2022.06.001.

5. Wu, G., and Schöler, H.R. (2016). Lineage Segregation in the Totipotent Embryo. Curr Top Dev Biol *117*, 301–317. https://doi.org/10.1016/bs.ctdb.2015.10.014.

6. Schulz, K.N., and Harrison, M.M. (2019). Mechanisms regulating zygotic genome activation. Nat Rev Genet *20*, 221–234. https://doi.org/10.1038/s41576-018-0087-x.

7. Xu, R., Zhu, Q., Zhao, Y., Chen, M., Yang, L., Shen, S., Yang, G., Shi, Z., Zhang, X., Shi, Q., et al. (2023). Unreprogrammed H3K9me3 prevents minor zygotic genome activation and lineage commitment in SCNT embryos. Nat Commun *14*, 4807. https://doi.org/10.1038/s41467-023-40496-3.

8. Sankar, A., Lerdrup, M., Manaf, A., Johansen, J.V., Gonzalez, J.M., Borup, R., Blanshard, R., Klungland, A., Hansen, K., Andersen, C.Y., et al. (2020). KDM4A regulates the maternal-to-zygotic transition by protecting broad H3K4me3 domains from H3K9me3 invasion in oocytes. Nat Cell Biol *22*, 380–388. https://doi.org/10.1038/s41556-020-0494-z.

9. Wang, C., Liu, X., Gao, Y., Yang, L., Li, C., Liu, W., Chen, C., Kou, X., Zhao, Y., Chen, J., et al. (2018). Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. Nature cell biology *20*, 620–631. https://doi.org/10.1038/s41556-018-0093-4.

10. Yu, H., Chen, M., Hu, Y., Ou, S., Yu, X., Liang, S., Li, N., Yang, M., Kong, X., Sun, C., et al. (2022). Dynamic reprogramming of H3K9me3 at hominoid-specific retrotransposons during human preimplantation development. Cell Stem Cell *29*, 1031-1050.e12. https://doi.org/10.1016/j.stem.2022.06.006.

11. Burton, A., Brochard, V., Galan, C., Ruiz-Morales, E.R., Rovira, Q., Rodriguez-Terrones, D., Kruse, K., Le Gras, S., Udayakumar, V.S., Chin, H.G., et al. (2020). Heterochromatin

establishment during early mammalian development is regulated by pericentromeric RNA and characterized by non-repressive H3K9me3. Nat Cell Biol *22*, 767–778. https://doi.org/10.1038/s41556-020-0536-6.

12. Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell *151*, 994–1004. https://doi.org/10.1016/j.cell.2012.09.045.

13. Methot, S.P., Padeken, J., Brancati, G., Zeller, P., Delaney, C.E., Gaidatzis, D., Kohler, H., van Oudenaarden, A., Großhans, H., and Gasser, S.M. (2021). H3K9me selectively blocks transcription factor activity and ensures differentiated tissue integrity. Nat Cell Biol *23*, 1163–1175. https://doi.org/10.1038/s41556-021-00776-w.

14. Sugiyama, T., Thillainadesan, G., Chalamcharla, V.R., Meng, Z., Balachandran, V., Dhakshnamoorthy, J., Zhou, M., and Grewal, S.I.S. (2016). Enhancer of Rudimentary Cooperates with Conserved RNA-Processing Factors to Promote Meiotic mRNA Decay and Facultative Heterochromatin Assembly. Mol Cell *61*, 747–759. https://doi.org/10.1016/j.molcel.2016.01.029.

15. Xie, G., Vo, T.V., Thillainadesan, G., Holla, S., Zhang, B., Jiang, Y., Lv, M., Xu, Z., Wang, C., Balachandran, V., et al. (2019). A conserved dimer interface connects ERH and YTH family proteins to promote gene silencing. Nat Commun *10*, 251. https://doi.org/10.1038/s41467-018-08273-9.

16. McCarthy, R.L., Kaeding, K.E., Keller, S.H., Zhong, Y., Xu, L., Hsieh, A., Hou, Y., Donahue, G., Becker, J.S., Alberto, O., et al. (2021). Diverse heterochromatin-associated proteins repress distinct classes of genes and repetitive elements. Nat Cell Biol *23*, 905–914. https://doi.org/10.1038/s41556-021-00725-7.

17. Sugiyama, T., Cam, H., Verdel, A., Moazed, D., and Grewal, S.I.S. (2005). RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. Proc Natl Acad Sci U S A *102*, 152–157. https://doi.org/10.1073/pnas.0407641102.

18. Zong, J., Yao, X., Yin, J., Zhang, D., and Ma, H. (2009). Evolution of the RNA-dependent RNA polymerase (RdRP) genes: duplications and possible losses before and after the divergence of major eukaryotic groups. Gene *447*, 29–39. https://doi.org/10.1016/j.gene.2009.07.004.

19. Matoba, S., Liu, Y., Lu, F., Iwabuchi, K.A., Shen, L., Inoue, A., and Zhang, Y. (2014). Embryonic development following somatic cell nuclear transfer impeded by persisting histone methylation. Cell *159*, 884–895. https://doi.org/10.1016/j.cell.2014.09.055.

20. Gao, Y., Liu, X., Tang, B., Li, C., Kou, Z., Li, L., Liu, W., Wu, Y., Kou, X., Li, J., et al. (2017). Protein Expression Landscape of Mouse Embryos during Pre-implantation Development. Cell reports *21*, 3957–3969. https://doi.org/10.1016/j.celrep.2017.11.111.

21. Ralston, A., and Rossant, J. (2008). Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. Dev Biol *313*, 614–629. https://doi.org/10.1016/j.ydbio.2007.10.054.

22. Lim, H.Y.G., Alvarez, Y.D., Gasnier, M., Wang, Y., Tetlak, P., Bissiere, S., Wang, H., Biro, M., and Plachta, N. (2020). Keratins are asymmetrically inherited fate determinants in the mammalian embryo. Nature *585*, 404–409. https://doi.org/10.1038/s41586-020-2647-4.

23. Mochizuki, K., Sharif, J., Shirane, K., Uranishi, K., Bogutz, A.B., Janssen, S.M., Suzuki, A., Okuda, A., Koseki, H., and Lorincz, M.C. (2021). Repression of germline genes by PRC1.6 and SETDB1 in the early embryo precedes DNA methylation-mediated silencing. Nat Commun *12*, 7020. https://doi.org/10.1038/s41467-021-27345-x.

24. Dahlet, T., Argüeso Lleida, A., Al Adhami, H., Dumas, M., Bender, A., Ngondo, R.P., Tanguy, M., Vallet, J., Auclair, G., Bardet, A.F., et al. (2020). Genome-wide analysis in the mouse embryo reveals the importance of DNA methylation for transcription integrity. Nat Commun *11*, 3153. https://doi.org/10.1038/s41467-020-16919-w.

25. Wu, D., and Dean, J. (2023). RNA exosome ribonuclease DIS3 degrades Pou6f1 to promote mouse pre-implantation cell differentiation. Cell Rep *42*, 112047. https://doi.org/10.1016/j.celrep.2023.112047.

26. Falco, G., Lee, S.-L., Stanghellini, I., Bassey, U.C., Hamatani, T., and Ko, M.S.H. (2007). Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. Dev Biol *307*, 539–550. https://doi.org/10.1016/j.ydbio.2007.05.003.

27. Nicholls, P.K., Schorle, H., Naqvi, S., Hu, Y.-C., Fan, Y., Carmell, M.A., Dobrinski, I., Watson, A.L., Carlson, D.F., Fahrenkrug, S.C., et al. (2019). Mammalian germ cells are determined after PGC colonization of the nascent gonad. Proc Natl Acad Sci U S A *116*, 25677–25687. https://doi.org/10.1073/pnas.1910733116.

28. Hamatani, T., Carter, M.G., Sharov, A.A., and Ko, M.S.H. (2004). Dynamics of global gene expression changes during mouse preimplantation development. Dev Cell *6*, 117–131. https://doi.org/10.1016/s1534-5807(03)00373-3.

29. Boussicault, L., Alves, S., Lamazière, A., Planques, A., Heck, N., Moumné, L., Despres, G., Bolte, S., Hu, A., Pagès, C., et al. (2016). CYP46A1, the rate-limiting enzyme for cholesterol degradation, is neuroprotective in Huntington's disease. Brain *139*, 953–970. https://doi.org/10.1093/brain/awv384.

30. Xie, Z., Bailey, A., Kuleshov, M.V., Clarke, D.J.B., Evangelista, J.E., Jenkins, S.L., Lachmann, A., Wojciechowicz, M.L., Kropiwnicki, E., Jagodnik, K.M., et al. (2021). Gene Set Knowledge Discovery with Enrichr. Curr Protoc *1*, e90. https://doi.org/10.1002/cpz1.90.

31. Stephens, A.D., Liu, P.Z., Banigan, E.J., Almassalha, L.M., Backman, V., Adam, S.A., Goldman, R.D., and Marko, J.F. (2018). Chromatin histone modifications and rigidity affect nuclear morphology independent of lamins. Mol Biol Cell *29*, 220–233. https://doi.org/10.1091/mbc.E17-06-0410.

32. Messmer, T., von Meyenn, F., Savino, A., Santos, F., Mohammed, H., Lun, A.T.L., Marioni, J.C., and Reik, W. (2019). Transcriptional Heterogeneity in Naive and Primed Human Pluripotent Stem Cells at Single-Cell Resolution. Cell Rep *26*, 815-824.e4. https://doi.org/10.1016/j.celrep.2018.12.099.

33. Teo, A.K.K., Valdez, I.A., Dirice, E., and Kulkarni, R.N. (2014). Comparable generation of activin-induced definitive endoderm via additive Wnt or BMP signaling in absence of serum. Stem Cell Reports *3*, 5–14. https://doi.org/10.1016/j.stemcr.2014.05.007.

34. Dubois, A., Vincenti, L., Chervova, A., Greenberg, M.V.C., Vandormael-Pournin, S., Bourc'his, D., Cohen-Tannoudji, M., and Navarro, P. (2022). H3K9 tri-methylation at Nanog times differentiation commitment and enables the acquisition of primitive endoderm fate. Development *149*, dev201074. https://doi.org/10.1242/dev.201074.

35. Lau, K.X., Mason, E.A., Kie, J., De Souza, D.P., Kloehn, J., Tull, D., McConville, M.J., Keniry, A., Beck, T., Blewitt, M.E., et al. (2020). Unique properties of a subset of human pluripotent stem cells with high capacity for self-renewal. Nat Commun *11*, 2420. https://doi.org/10.1038/s41467-020-16214-8.

36. Ai, Z., Xiang, X., Xiang, Y., Szczerbinska, I., Qian, Y., Xu, X., Ma, C., Su, Y., Gao, B., Shen, H., et al. (2022). Krüppel-like factor 5 rewires NANOG regulatory network to activate human naive pluripotency specific LTR7Ys and promote naive pluripotency. Cell Rep *40*, 111240. https://doi.org/10.1016/j.celrep.2022.111240.

37. Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Göttgens, B., Niwa, H., and Smith, A. (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell Stem Cell *11*, 491–504. https://doi.org/10.1016/j.stem.2012.06.008.

38. Gassler, J., Kobayashi, W., Gáspár, I., Ruangroengkulrith, S., Mohanan, A., Gómez Hernández, L., Kravchenko, P., Kümmecke, M., Lalic, A., Rifel, N., et al. (2022). Zygotic genome activation by the totipotency pioneer factor Nr5a2. Science *378*, 1305–1315. https://doi.org/10.1126/science.abn7478.

39. Maskalenka, K., Alagöz, G., Krueger, F., Wright, J., Rostovskaya, M., Nakhuda, A., Bendall, A., Krueger, C., Walker, S., Scally, A., et al. (2023). NANOGP1, a tandem duplicate of NANOG, exhibits partial functional conservation in human naïve pluripotent stem cells. Development *150*, dev201155. https://doi.org/10.1242/dev.201155.

40. Pontis, J., Planet, E., Offner, S., Turelli, P., Duc, J., Coudray, A., Theunissen, T.W., Jaenisch, R., and Trono, D. (2019). Hominoid-Specific Transposable Elements and KZFPs Facilitate Human Embryonic Genome Activation and Control Transcription in Naive Human ESCs. Cell Stem Cell *24*, 724-735.e5. https://doi.org/10.1016/j.stem.2019.03.012.

41. Wu, F., Liufu, Z., Liu, Y., Guo, L., Wu, J., Cao, S., Qin, Y., Guo, N., Fu, Y., Liu, H., et al. (2022). Species-specific rewiring of definitive endoderm developmental gene activation via endogenous retroviruses through TET1-mediated demethylation. Cell Rep *41*, 111791. https://doi.org/10.1016/j.celrep.2022.111791.

42. Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature *487*, 57–63. https://doi.org/10.1038/nature11244.

43. Pontis, J., Pulver, C., Playfoot, C.J., Planet, E., Grun, D., Offner, S., Duc, J., Manfrin, A., Lutolf, M.P., and Trono, D. (2022). Primate-specific transposable elements shape

transcriptional networks during human development. Nat Commun *13*, 7178. https://doi.org/10.1038/s41467-022-34800-w.

44. Wang, H., Xing, J., Grover, D., Hedges, D.J., Han, K., Walker, J.A., and Batzer, M.A. (2005). SVA elements: a hominid-specific retroposon family. J Mol Biol *354*, 994–1007. https://doi.org/10.1016/j.jmb.2005.09.085.

45. Grow, E.J., Flynn, R.A., Chavez, S.L., Bayless, N.L., Wossidlo, M., Wesche, D.J., Martin, L., Ware, C.B., Blish, C.A., Chang, H.Y., et al. (2015). Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. Nature *522*, 221–225. https://doi.org/10.1038/nature14308.

46. Khan, H., Smit, A., and Boissinot, S. (2006). Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. Genome Res *16*, 78–87. https://doi.org/10.1101/gr.4001406.

47. Theunissen, T.W., Friedli, M., He, Y., Planet, E., O'Neil, R.C., Markoulaki, S., Pontis, J., Wang, H., Iouranova, A., Imbeault, M., et al. (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. Cell Stem Cell *19*, 502–515. https://doi.org/10.1016/j.stem.2016.06.011.

48. Domcke, S., Hill, A.J., Daza, R.M., Cao, J., O'Day, D.R., Pliner, H.A., Aldinger, K.A., Pokholok, D., Zhang, F., Milbank, J.H., et al. (2020). A human cell atlas of fetal chromatin accessibility. Science *370*, eaba7612. https://doi.org/10.1126/science.aba7612.

49. Evseenko, D., Zhu, Y., Schenke-Layland, K., Kuo, J., Latour, B., Ge, S., Scholes, J., Dravid, G., Li, X., MacLellan, W.R., et al. (2010). Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. Proc Natl Acad Sci U S A *107*, 13742–13747. https://doi.org/10.1073/pnas.1002077107.

50. Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M., Trotter, M.W., Hemberger, M., Smith, J.C., et al. (2011). BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. Cell Stem Cell *9*, 144–155. https://doi.org/10.1016/j.stem.2011.06.015.

51. Koche, R.P., Smith, Z.D., Adli, M., Gu, H., Ku, M., Gnirke, A., Bernstein, B.E., and Meissner, A. (2011). Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell stem cell *8*, 96–105. https://doi.org/10.1016/j.stem.2010.12.001.

52. Taing, L., Dandawate, A., L'Yi, S., Gehlenborg, N., Brown, M., and Meyer, C.A. (2024). Cistrome Data Browser: integrated search, analysis and visualization of chromatin data. Nucleic Acids Res *52*, D61–D66. https://doi.org/10.1093/nar/gkad1069.

53. Zheng, R., Wan, C., Mei, S., Qin, Q., Wu, Q., Sun, H., Chen, C.-H., Brown, M., Zhang, X., Meyer, C.A., et al. (2019). Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. Nucleic Acids Res *47*, D729–D735. https://doi.org/10.1093/nar/gky1094.

54. Yu, X., Liang, S., Chen, M., Yu, H., Li, R., Qu, Y., Kong, X., Guo, R., Zheng, R., Izsvák, Z., et al. (2022). Recapitulating early human development with 8C-like cells. Cell Rep *39*, 110994. https://doi.org/10.1016/j.celrep.2022.110994.

55. Zou, Z., Zhang, C., Wang, Q., Hou, Z., Xiong, Z., Kong, F., Wang, Q., Song, J., Liu, B., Liu, B., et al. (2022). Translatome and transcriptome co-profiling reveals a role of TPRXs in human zygotic genome activation. Science *378*, abo7923. https://doi.org/10.1126/science.abo7923.

56. Taubenschmid-Stowers, J., Rostovskaya, M., Santos, F., Ljung, S., Argelaguet, R., Krueger, F., Nichols, J., and Reik, W. (2022). 8C-like cells capture the human zygotic genome activation program in vitro. Cell Stem Cell *29*, 449-459.e6. https://doi.org/10.1016/j.stem.2022.01.014.

57. Cacchiarelli, D., Trapnell, C., Ziller, M.J., Soumillon, M., Cesana, M., Karnik, R., Donaghey, J., Smith, Z.D., Ratanasirintrawoot, S., Zhang, X., et al. (2015). Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency. Cell *162*, 412–424. https://doi.org/10.1016/j.cell.2015.06.016.

58. Wang, J., Xie, G., Singh, M., Ghanbarian, A.T., Raskó, T., Szvetnik, A., Cai, H., Besser, D., Prigione, A., Fuchs, N.V., et al. (2014). Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. Nature *516*, 405–409. https://doi.org/10.1038/nature13804.

59. Lu, X., Sachs, F., Ramsay, L., Jacques, P.-É., Göke, J., Bourque, G., and Ng, H.-H. (2014). The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. Nat Struct Mol Biol *21*, 423–425. https://doi.org/10.1038/nsmb.2799.

60. Mills, J.D., Kavanagh, T., Kim, W.S., Chen, B.J., Waters, P.D., Halliday, G.M., and Janitz, M. (2015). High expression of long intervening non-coding RNA OLMALINC in the human cortical white matter is associated with regulation of oligodendrocyte maturation. Mol Brain *8*, 2. https://doi.org/10.1186/s13041-014-0091-9.

61. Obersriebnig, M.J., Pallesen, E.M.H., Sneppen, K., Trusina, A., and Thon, G. (2016). Nucleation and spreading of a heterochromatic domain in fission yeast. Nat Commun *7*, 11518. https://doi.org/10.1038/ncomms11518.

62. Murawska, M., Greenstein, R.A., Schauer, T., Olsen, K.C.F., Ng, H., Ladurner, A.G., Al-Sady, B., and Braun, S. (2021). The histone chaperone FACT facilitates heterochromatin spreading by regulating histone turnover and H3K9 methylation states. Cell Rep *37*, 109944. https://doi.org/10.1016/j.celrep.2021.109944.

63. Greenstein, R.A., Jones, S.K., Spivey, E.C., Rybarski, J.R., Finkelstein, I.J., and Al-Sady, B. (2018). Noncoding RNA-nucleated heterochromatin spreading is intrinsically labile and requires accessory elements for epigenetic stability. Elife *7*. https://doi.org/10.7554/eLife.32948.

64. Al Adhami, H., Vallet, J., Schaal, C., Schumacher, P., Bardet, A.F., Dumas, M., Chicher, J., Hammann, P., Daujat, S., and Weber, M. (2023). Systematic identification of factors involved in the silencing of germline genes in mouse embryonic stem cells. Nucleic Acids Res, gkad071. https://doi.org/10.1093/nar/gkad071.

65. Yang, P., Wang, Y., and Macfarlan, T.S. (2017). The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. Trends Genet *33*, 871–881. https://doi.org/10.1016/j.tig.2017.08.006.

66. Imbeault, M., Helleboid, P.-Y., and Trono, D. (2017). KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. Nature *543*, 550–554. https://doi.org/10.1038/nature21683.

67. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496–502. https://doi.org/10.1038/s41586-019-0969-x.

68. Thomas, P.D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L.-P., and Mi, H. (2022). PANTHER: Making genome-scale phylogenetics accessible to all. Protein Sci *31*, 8–22. https://doi.org/10.1002/pro.4218.

69. Jin, Y., Tam, O.H., Paniagua, E., and Hammell, M. (2015). TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. Bioinformatics *31*, 3593–3599. https://doi.org/10.1093/bioinformatics/btv422.

70. Cao, J., O'Day, D.R., Pliner, H.A., Kingsley, P.D., Deng, M., Daza, R.M., Zager, M.A., Aldinger, K.A., Blecher-Gonen, R., Zhang, F., et al. (2020). A human cell atlas of fetal gene expression. Science *370*, eaba7721. https://doi.org/10.1126/science.aba7721.

71. Skene, P.J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife *6*, e21856. https://doi.org/10.7554/eLife.21856.

72. Katznelson, A., and Zaret, K. (2024). A fluorescence-based protocol to quantitatively titrate CUT&RUN buffer components. STAR Protoc *5*, 102866. https://doi.org/10.1016/j.xpro.2024.102866.

73. Liu, N., Hargreaves, V.V., Zhu, Q., Kurland, J.V., Hong, J., Kim, W., Sher, F., Macias-Trevino, C., Rogers, J.M., Kurita, R., et al. (2018). Direct Promoter Repression by BCL11A Controls the Fetal to Adult Hemoglobin Switch. Cell *173*, 430-442.e17. https://doi.org/10.1016/j.cell.2018.03.016.

74. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

75. Jin, H., Kasper, L.H., Larson, J.D., Wu, G., Baker, S.J., Zhang, J., and Fan, Y. (2020). ChIPseqSpikeInFree: a ChIP-seq normalization approach to reveal global changes in histone modifications without spike-in. Bioinformatics *36*, 1270–1272. https://doi.org/10.1093/bioinformatics/btz720.

76. Meers, M.P., Tenenbaum, D., and Henikoff, S. (2019). Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling. Epigenetics Chromatin *12*, 42. https://doi.org/10.1186/s13072-019-0287-4.

77. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining

transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576–589. https://doi.org/10.1016/j.molcel.2010.05.004.

78. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160-165. https://doi.org/10.1093/nar/gkw257.

79. Khan, A., and Mathelier, A. (2017). Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. BMC Bioinformatics *18*, 287. https://doi.org/10.1186/s12859-017-1708-7.

80. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

81. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

82. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

83. Filippi-Chiela, E.C., Oliveira, M.M., Jurkovski, B., Callegari-Jacques, S.M., da Silva, V.D., and Lenz, G. (2012). Nuclear morphometric analysis (NMA): screening of senescence, apoptosis and nuclear irregularities. PLoS One *7*, e42522. https://doi.org/10.1371/journal.pone.0042522.

Figure 1

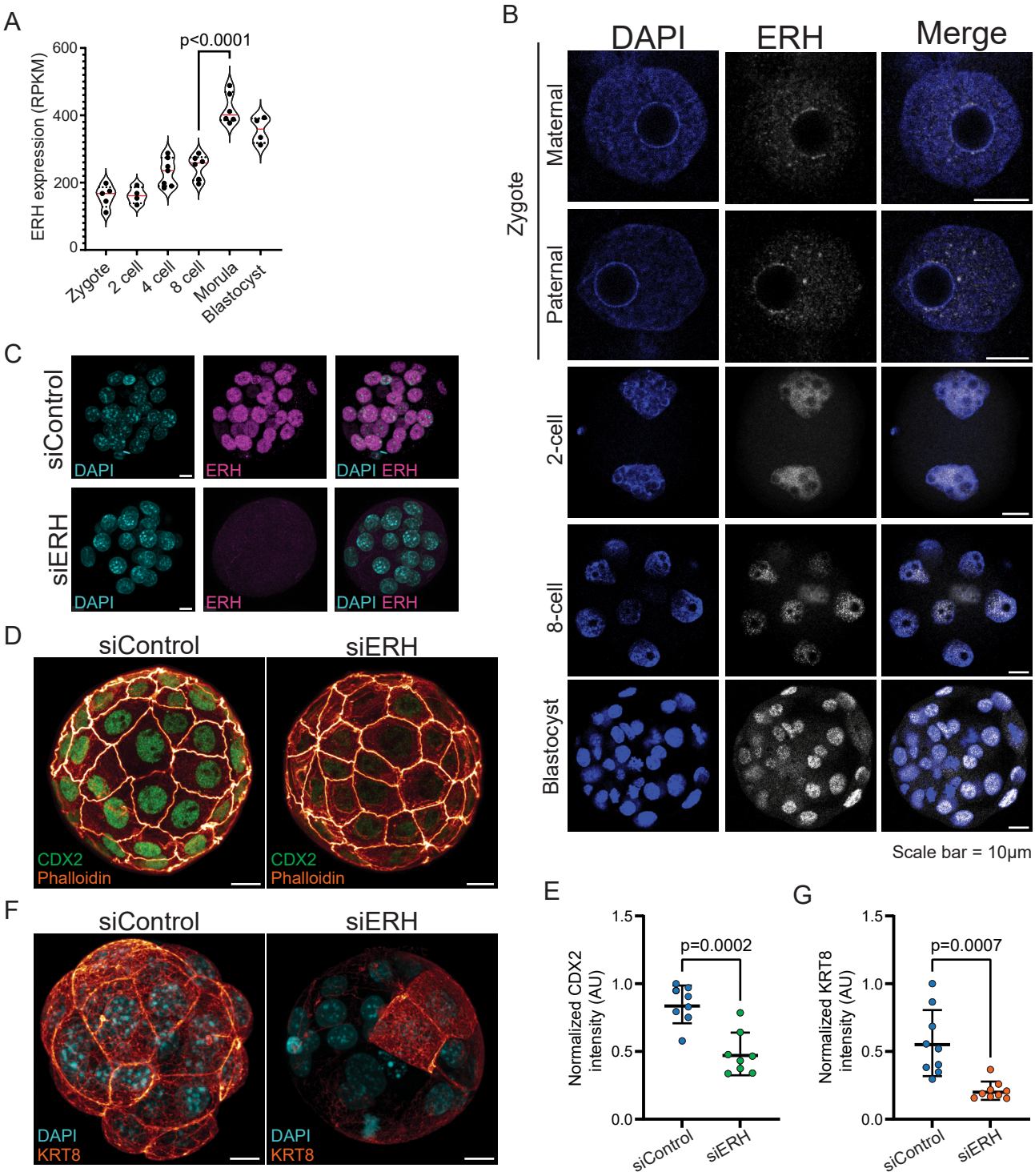

A
p<0.0001
ERH expression (RPKM)
600
400
200
0
Zygote
2 cell
4 cell
8 cell
Morula
Blastocyst
B
DAPI
ERH
Merge
Zygote
Maternal
Paternal
2-cell
8-cell
Blastocyst
Scale bar = 10μm
C
siControl
siERH
DAPI
ERH
DAPI ERH
D
siControl
siERH
CDX2
Phalloidin
F
siControl
siERH
DAPI
KRT8
E
Normalized CDX2 intensity (AU)
p=0.0002
1.5
1.0
0.5
0.0
siControl
siERH
G
Normalized KRT8 intensity (AU)
p=0.0007
1.5
1.0
0.5
0.0
siControl
siERH

Figure 2

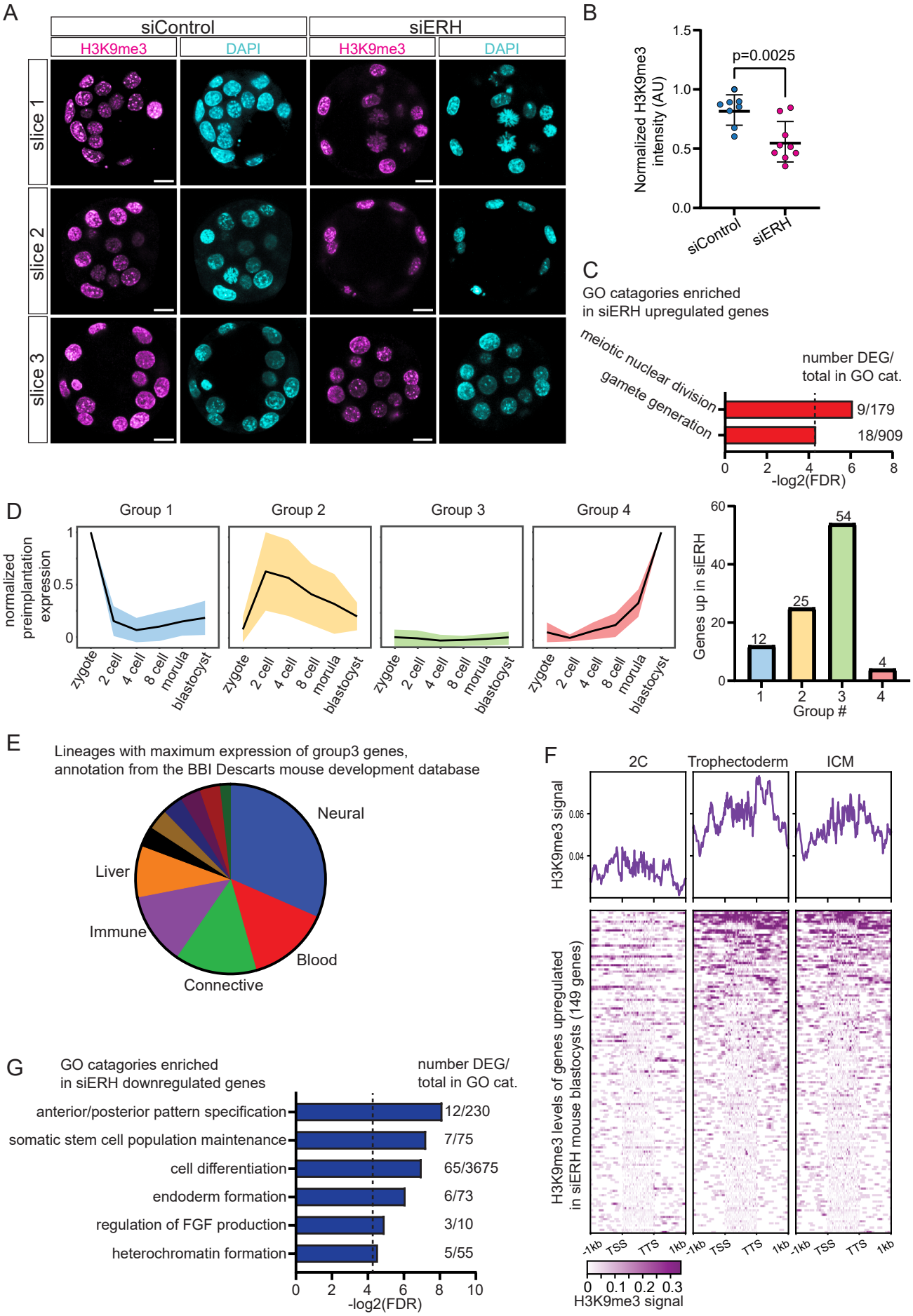

A
siControl
siERH
H3K9me3
DAPI
H3K9me3
DAPI
slice 1
slice 2
slice 3
B
Normalized H3K9me3 intensity (AU)
p=0.0025
1.5
1.0
0.5
0.0
siControl
siERH
C
GO catagories enriched in siERH upregulated genes
number DEG/ total in GO cat.
meiotic nuclear division
gamete generation
9/179
18/909
0
2
4
6
8
-log2(FDR)
D
Group 1
Group 2
Group 3
Group 4
normalized preimplantation expression
1
0.5
0
zygote
2 cell
4 cell
8 cell
morula
blastocyst
Genes up in siERH
60
40
20
0
12
25
54
4
1
2
3
4
Group #
E
Lineages with maximum expression of group3 genes, annotation from the BBI Descarts mouse development database
Neural
Liver
Immune
Blood
Connective
F
2C
Trophectoderm
ICM
H3K9me3 signal
0.06
0.04
H3K9me3 levels of genes upregulated in siERH mouse blastocysts (149 genes)
-1kb
TSS
TTS
1kb
0
0.1
0.2
0.3
H3K9me3 signal
G
GO catagories enriched in siERH downregulated genes
number DEG/ total in GO cat.
anterior/posterior pattern specification
somatic stem cell population maintenance
cell differentiation
endoderm formation
regulation of FGF production
heterochromatin formation
12/230
7/75
65/3675
6/73
3/10
5/55
0
2
4
6
8
10
-log2(FDR)

Figure 3

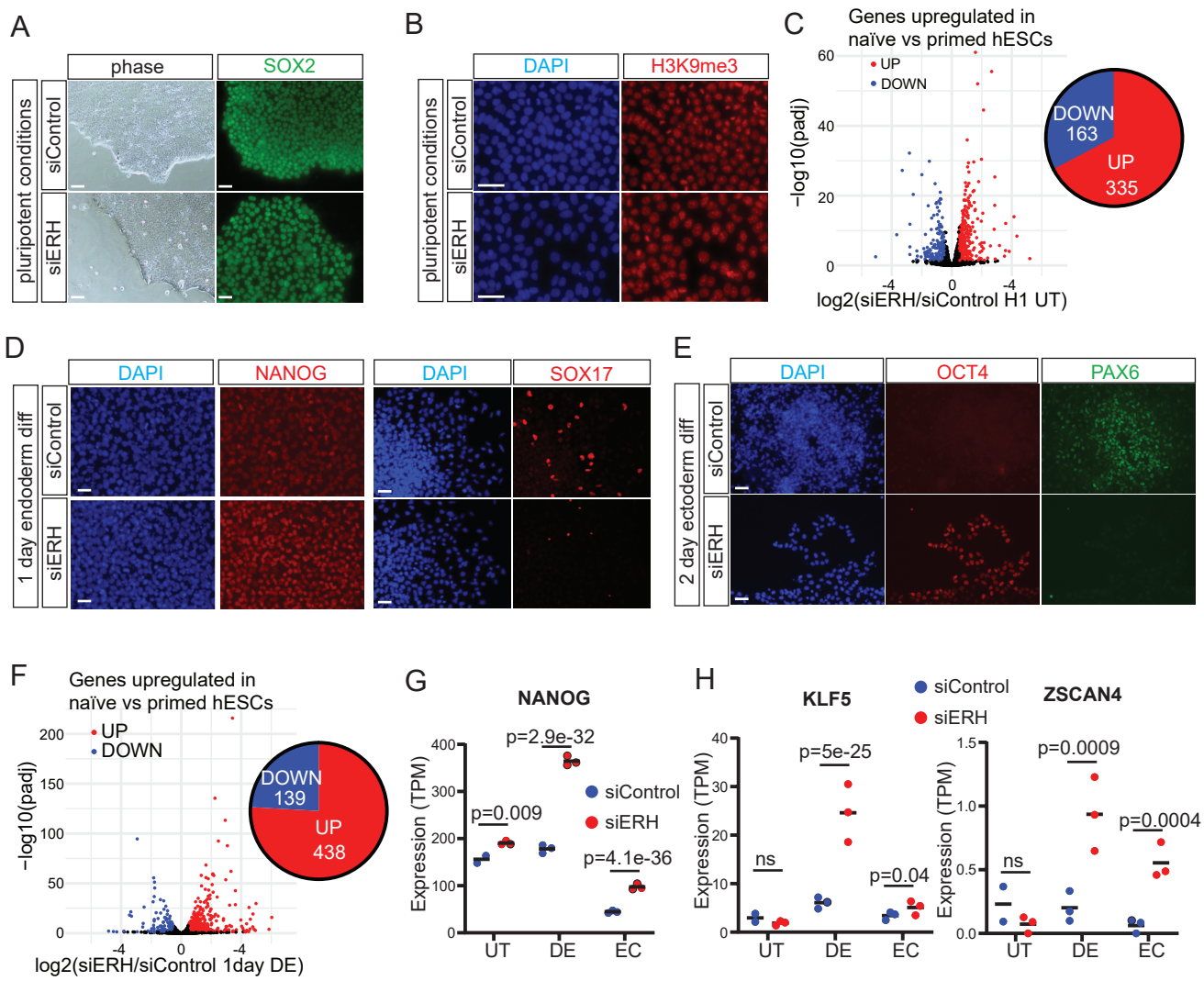

A
phase
SOX2
pluripotent conditions
siControl
siERH
B
DAPI
H3K9me3
pluripotent conditions
siControl
siERH
C
Genes upregulated in naïve vs primed hESCs
UP
DOWN
−log10(padj)
log2(siERH/siControl H1 UT)
DOWN 163
UP 335
D
DAPI
NANOG
DAPI
SOX17
1 day endoderm diff
siControl
siERH
E
DAPI
OCT4
PAX6
2 day ectoderm diff
siControl
siERH
F
Genes upregulated in naïve vs primed hESCs
UP
DOWN
−log10(padj)
log2(siERH/siControl 1day DE)
DOWN 139
UP 438
G
NANOG
Expression (TPM)
p=2.9e-32
p=0.009
p=4.1e-36
siControl
siERH
UT
DE
EC
H
KLF5
ZSCAN4
siControl
siERH
Expression (TPM)
p=5e-25
ns
p=0.04
p=0.0009
ns
p=0.0004
UT
DE
EC

Figure 4

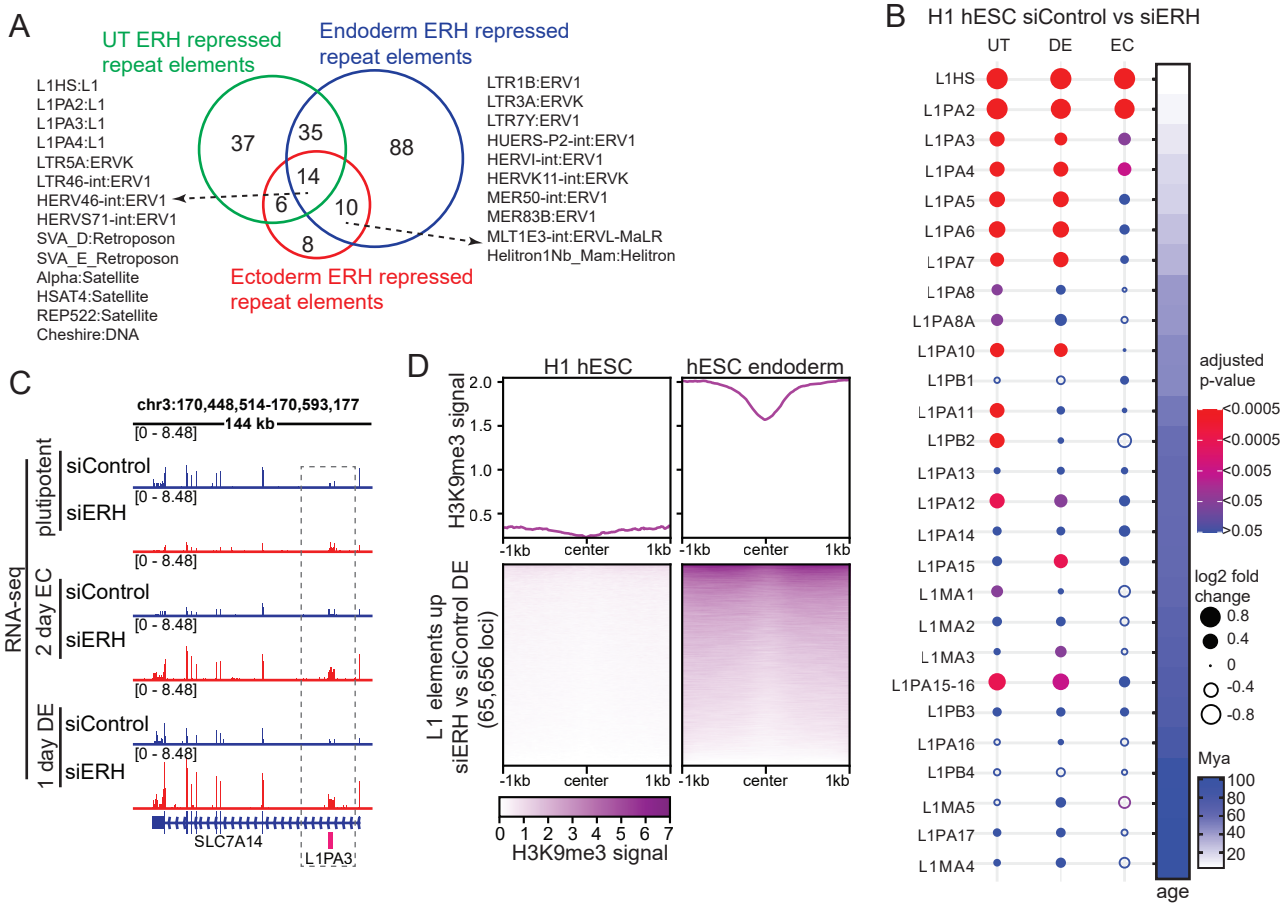

A
UT ERH repressed repeat elements
Endoderm ERH repressed repeat elements
Ectoderm ERH repressed repeat elements
37
35
88
14
6
10
8
L1HS:L1
L1PA2:L1
L1PA3:L1
L1PA4:L1
LTR5A:ERVK
LTR46-int:ERV1
HERV46-int:ERV1
HERVS71-int:ERV1
SVA_D:Retroposon
SVA_E_Retroposon
Alpha:Satellite
HSAT4:Satellite
REP522:Satellite
Cheshire:DNA
LTR1B:ERV1
LTR3A:ERVK
LTR7Y:ERV1
HUERS-P2-int:ERV1
HERVI-int:ERV1
HERVK11-int:ERVK
MER50-int:ERV1
MER83B:ERV1
MLT1E3-int:ERVL-MaLR
Helitron1Nb_Mam:Helitron
B
H1 hESC siControl vs siERH
UT
DE
EC
L1HS
L1PA2
L1PA3
L1PA4
L1PA5
L1PA6
L1PA7
L1PA8
L1PA8A
L1PA10
L1PB1
L1PA11
L1PB2
L1PA13
L1PA12
L1PA14
L1PA15
L1MA1
L1MA2
L1MA3
L1PA15-16
L1PB3
L1PA16
L1PB4
L1MA5
L1PA17
L1MA4
adjusted p-value
<0.0005
<0.0005
<0.005
<0.05
>0.05
log2 fold change
0.8
0.4
0
-0.4
-0.8
Mya
100
80
60
40
20
age
C
chr3:170,448,514-170,593,177
144 kb
[0 - 8.48]
RNA-seq
plutipotent
siControl
siERH
2 day EC
siControl
siERH
1 day DE
siControl
siERH
SLC7A14
L1PA3
D
H1 hESC
hESC endoderm
H3K9me3 signal
2.0
1.5
1.0
0.5
-1kb
center
1kb
L1 elements up siERH vs siControl DE (65,656 loci)
0 1 2 3 4 5 6 7
H3K9me3 signal

Figure 5

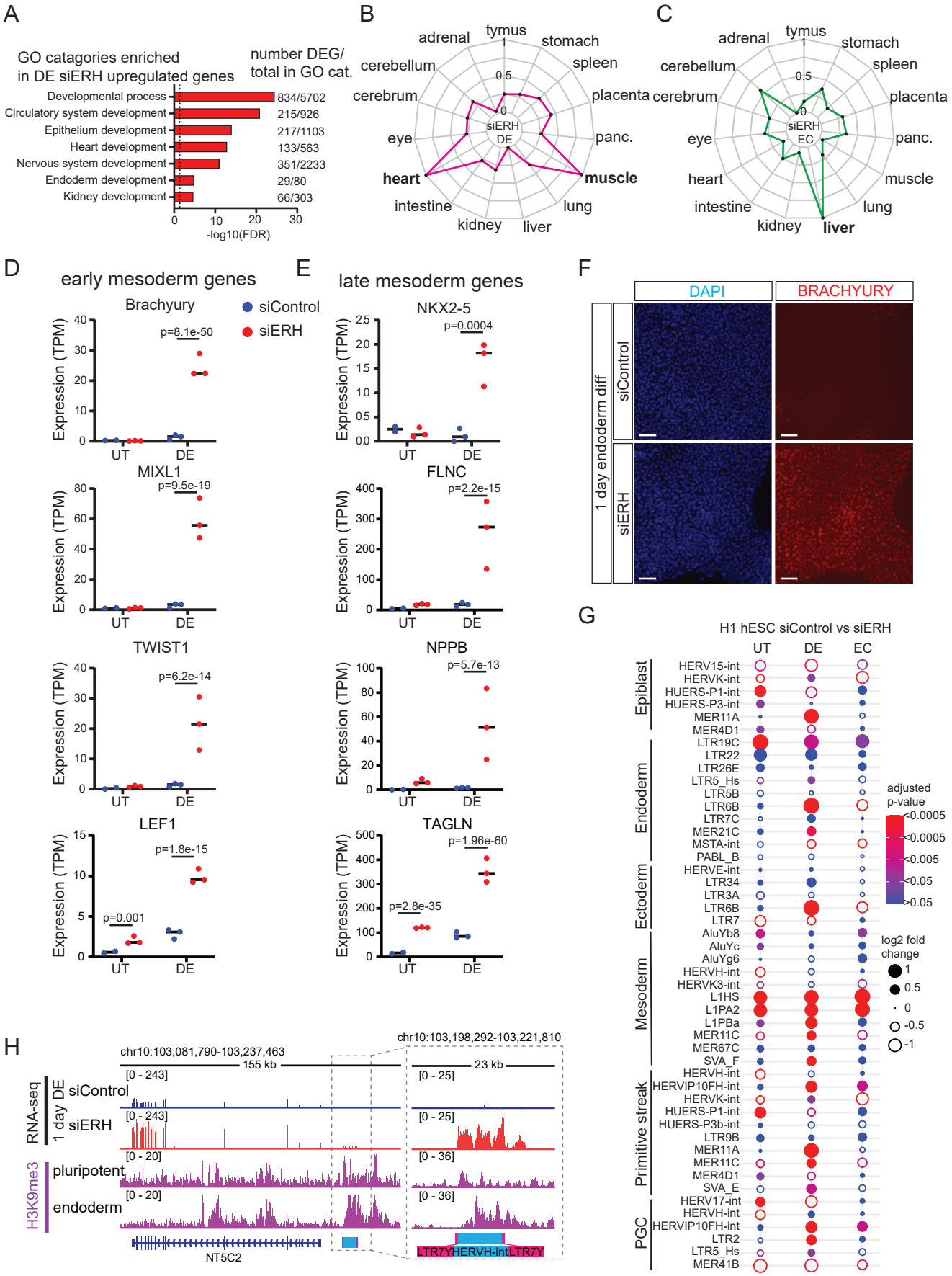

A
GO catagories enriched in DE siERH upregulated genes
number DEG/ total in GO cat.
Developmental process 834/5702
Circulatory system development 215/926
Epithelium development 217/1103
Heart development 133/563
Nervous system development 351/2233
Endoderm development 29/80
Kidney development 66/303
-log10(FDR)
B
siERH DE
C
siERH EC
tymus
stomach
spleen
placenta
panc.
muscle
lung
liver
kidney
intestine
heart
eye
cerebrum
cerebellum
adrenal
D
early mesoderm genes
Brachyury
MIXL1
TWIST1
LEF1
siControl
siERH
Expression (TPM)
UT
DE
E
late mesoderm genes
NKX2-5
FLNC
NPPB
TAGLN
F
DAPI
BRACHYURY
1 day endoderm diff
siControl
siERH
G
H1 hESC siControl vs siERH
UT
DE
EC
Epiblast
Endoderm
Ectoderm
Mesoderm
Primitive streak
PGC
adjusted p-value
log2 fold change
H
chr10:103,081,790-103,237,463
155 kb
chr10:103,198,292-103,221,810
23 kb
RNA-seq
1 day DE
siControl
siERH
H3K9me3
pluripotent
endoderm
NT5C2
LTR7YHERVH-intLTR7Y

Figure 6

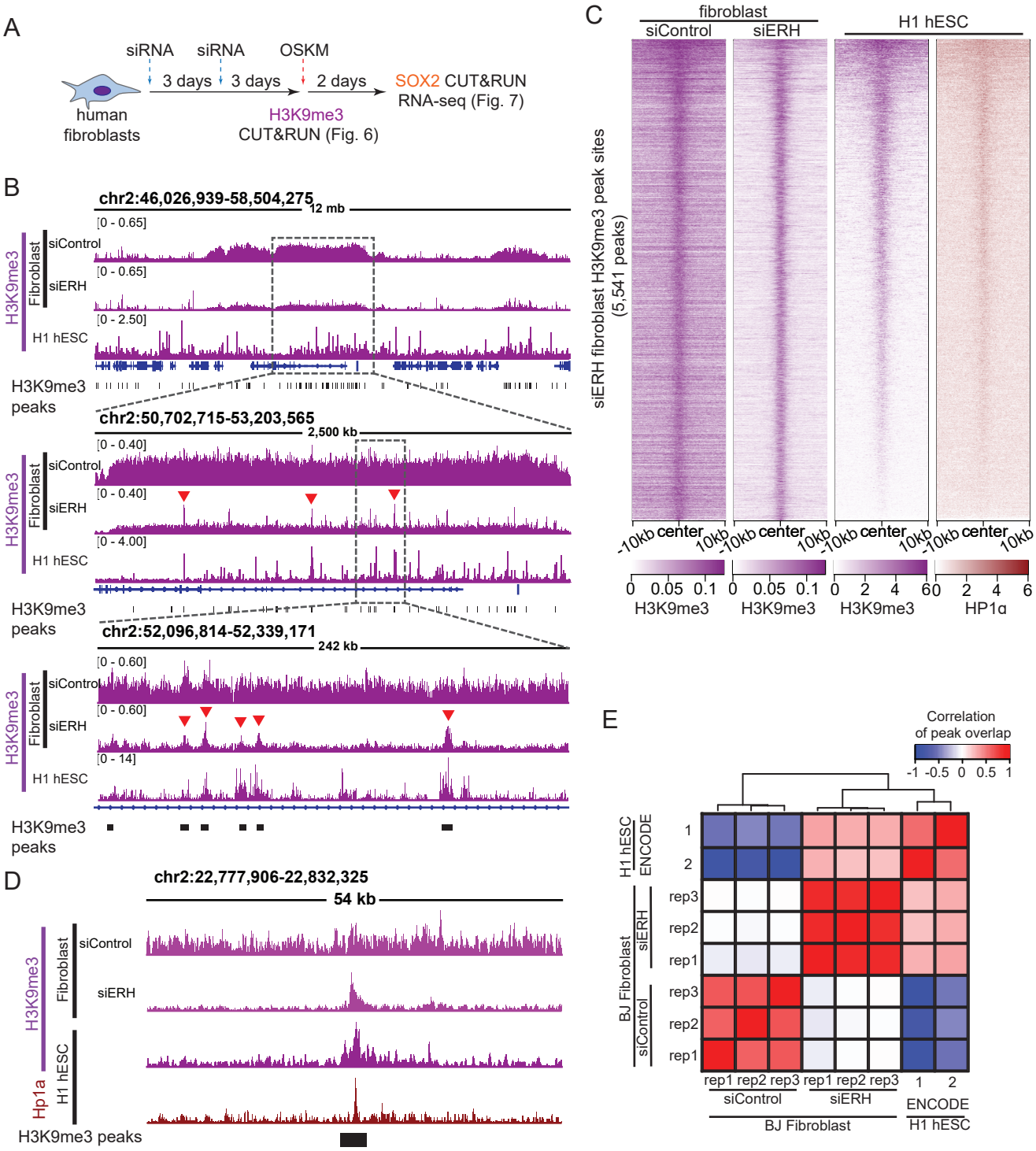

A
siRNA
siRNA
OSKM
3 days
3 days
2 days
human fibroblasts
H3K9me3 CUT&RUN (Fig. 6)
SOX2 CUT&RUN RNA-seq (Fig. 7)
B
chr2:46,026,939-58,504,275
12 mb
H3K9me3
Fibroblast
siControl
siERH
H1 hESC
H3K9me3 peaks
chr2:50,702,715-53,203,565
2,500 kb
chr2:52,096,814-52,339,171
242 kb
C
fibroblast
siControl
siERH
H1 hESC
siERH fibroblast H3K9me3 peak sites (5,541 peaks)
-10kb center 10kb
H3K9me3
HP1a
D
chr2:22,777,906-22,832,325
54 kb
Hp1a
E
Correlation of peak overlap
-1 -0.5 0 0.5 1
H1 hESC
ENCODE
BJ Fibroblast
siERH
siControl
rep1 rep2 rep3
BJ Fibroblast
ENCODE H1 hESC

Figure 7

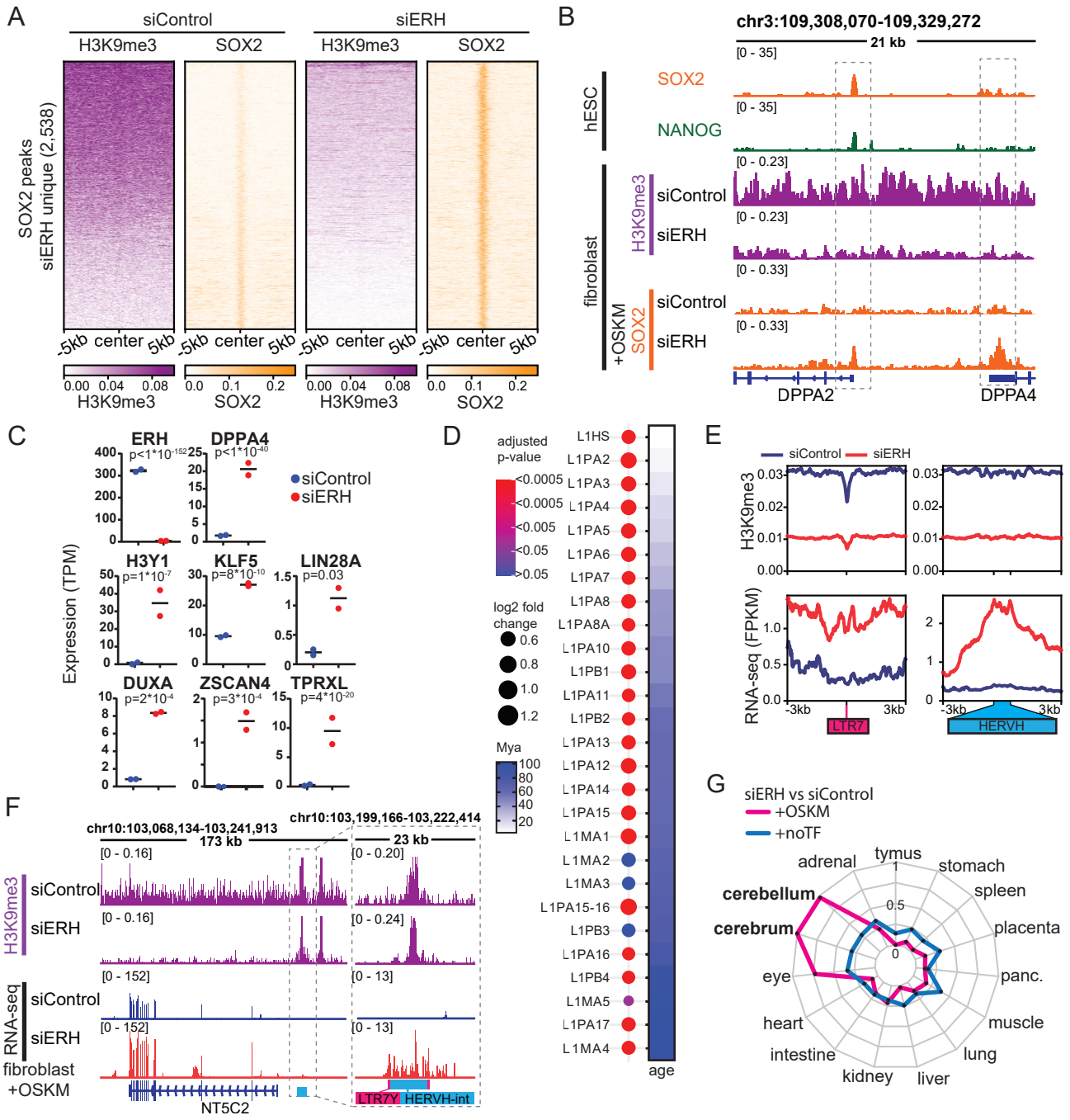
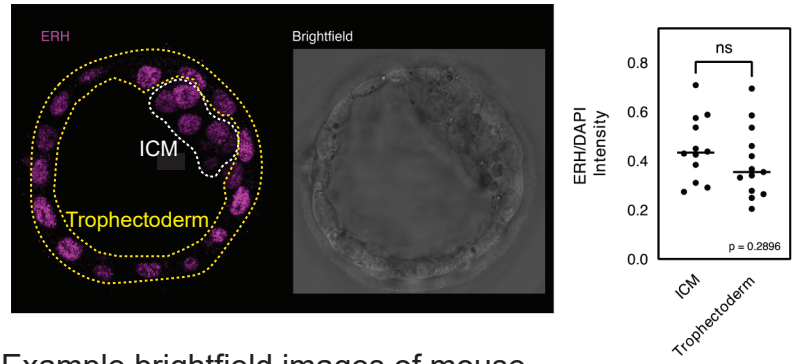
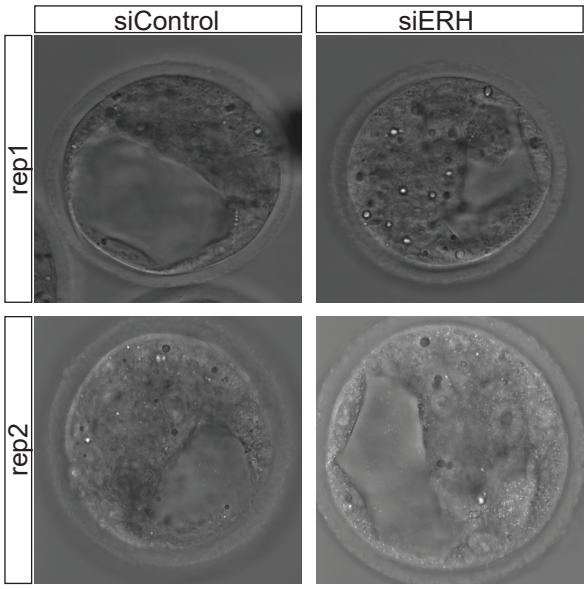

A
siControl
siERH
H3K9me3
SOX2
SOX2 peaks
siERH unique (2,538)
-5kb center 5kb
H3K9me3
SOX2
B
chr3:109,308,070-109,329,272
21 kb
hESC
SOX2
NANOG
fibroblast
H3K9me3
siControl
siERH
+OSKM
SOX2
DPPA2
DPPA4
C
ERH
DPPA4
H3Y1
KLF5
LIN28A
DUXA
ZSCAN4
TPRXL
siControl
siERH
Expression (TPM)
D
adjusted p-value
log2 fold change
Mya
age
E
siControl
siERH
H3K9me3
RNA-seq (FPKM)
LTR7
HERVH
F
chr10:103,068,134-103,241,913
173 kb
chr10:103,199,166-103,222,414
23 kb
H3K9me3
siControl
siERH
RNA-seq
fibroblast
+OSKM
NT5C2
LTR7Y
HERVH-int
G
siERH vs siControl
+OSKM
+noTF
tymus
stomach
spleen
placenta
panc.
muscle
lung
liver
kidney
intestine
heart
eye
cerebrum
cerebellum
adrenal

Figure S1

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




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




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

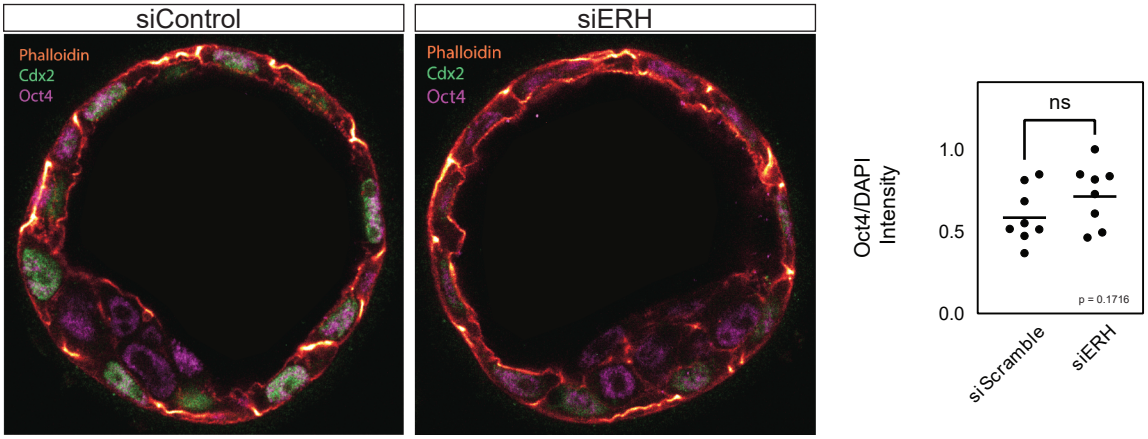


Figure S2

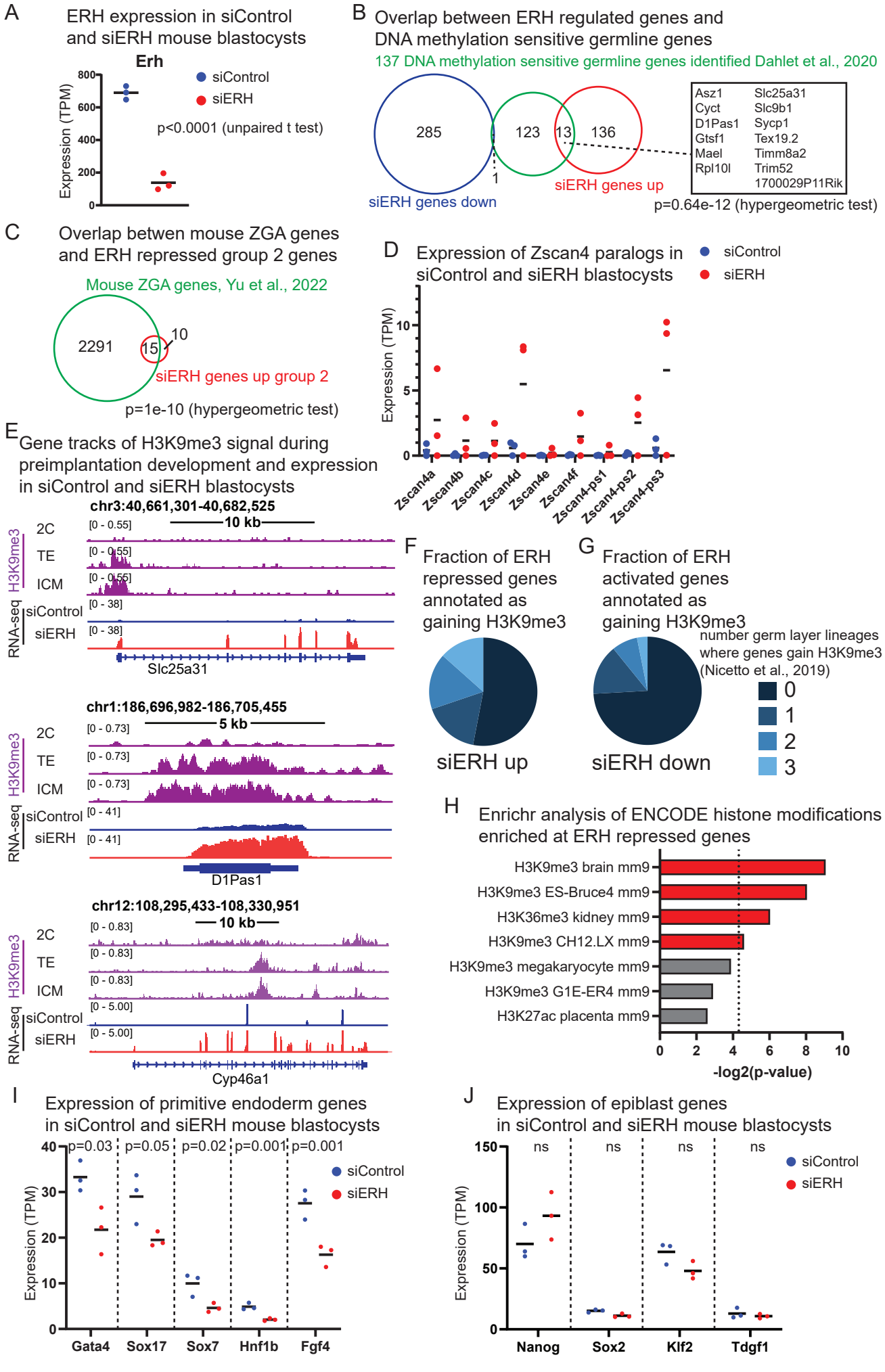

A
ERH expression in siControl and siERH mouse blastocysts
Erh
siControl
siERH
p<0.0001 (unpaired t test)
Expression (TPM)
B
Overlap between ERH regulated genes and DNA methylation sensitive germline genes
137 DNA methylation sensitive germline genes identified Dahlet et al., 2020
285
123
13
136
1
siERH genes down
siERH genes up
Asz1 Slc25a31
Cyct Slc9b1
D1Pas1 Sycp1
Gtsf1 Tex19.2
Mael Timm8a2
Rpl10l Trim52
1700029P11Rik
p=0.64e-12 (hypergeometric test)
C
Overlap betwen mouse ZGA genes and ERH repressed group 2 genes
Mouse ZGA genes, Yu et al., 2022
2291
15
10
siERH genes up group 2
p=1e-10 (hypergeometric test)
D
Expression of Zscan4 paralogs in siControl and siERH blastocysts
siControl
siERH
Expression (TPM)
Zscan4a
Zscan4b
Zscan4c
Zscan4d
Zscan4e
Zscan4f
Zscan4-ps1
Zscan4-ps2
Zscan4-ps3
E
Gene tracks of H3K9me3 signal during preimplantation development and expression in siControl and siERH blastocysts
chr3:40,661,301-40,682,525
10 kb
H3K9me3
2C
TE
ICM
RNA-seq
siControl
siERH
Slc25a31
chr1:186,696,982-186,705,455
5 kb
D1Pas1
chr12:108,295,433-108,330,951
10 kb
Cyp46a1
F
Fraction of ERH repressed genes annotated as gaining H3K9me3
siERH up
G
Fraction of ERH activated genes annotated as gaining H3K9me3
siERH down
number germ layer lineages where genes gain H3K9me3 (Nicetto et al., 2019)
0
1
2
3
H
Enrichr analysis of ENCODE histone modifications enriched at ERH repressed genes
H3K9me3 brain mm9
H3K9me3 ES-Bruce4 mm9
H3K36me3 kidney mm9
H3K9me3 CH12.LX mm9
H3K9me3 megakaryocyte mm9
H3K9me3 G1E-ER4 mm9
H3K27ac placenta mm9
-log2(p-value)
I
Expression of primitive endoderm genes in siControl and siERH mouse blastocysts
p=0.03
p=0.05
p=0.02
p=0.001
p=0.001
siControl
siERH
Expression (TPM)
Gata4
Sox17
Sox7
Hnf1b
Fgf4
J
Expression of epiblast genes in siControl and siERH mouse blastocysts
ns
ns
ns
ns
siControl
siERH
Expression (TPM)
Nanog
Sox2
Klf2
Tdgf1

Figure S3

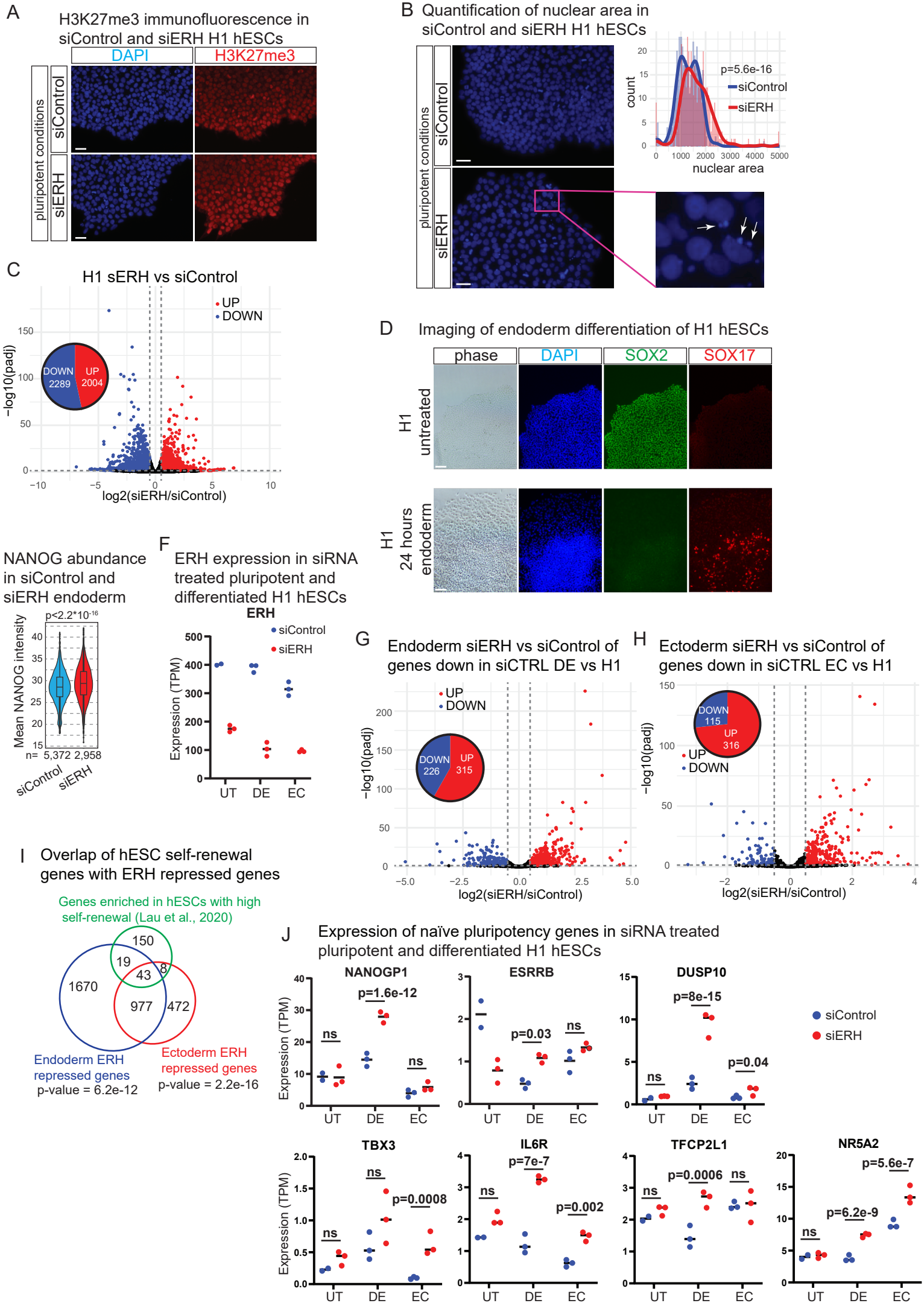

A
H3K27me3 immunofluorescence in siControl and siERH H1 hESCs
DAPI
H3K27me3
pluripotent conditions
siControl
siERH
B
Quantification of nuclear area in siControl and siERH H1 hESCs
p=5.6e-16
siControl
siERH
count
nuclear area
C
H1 sERH vs siControl
UP
DOWN
DOWN 2289
UP 2004
−log10(padj)
log2(siERH/siControl)
D
Imaging of endoderm differentiation of H1 hESCs
phase
DAPI
SOX2
SOX17
H1 untreated
H1 24 hours endoderm
E
NANOG abundance in siControl and siERH endoderm
p<2.2*10-16
Mean NANOG intensity
n= 5,372 2,958
siControl
siERH
F
ERH expression in siRNA treated pluripotent and differentiated H1 hESCs
ERH
siControl
siERH
Expression (TPM)
UT
DE
EC
G
Endoderm siERH vs siControl of genes down in siCTRL DE vs H1
UP
DOWN
DOWN 226
UP 315
−log10(padj)
log2(siERH/siControl)
H
Ectoderm siERH vs siControl of genes down in siCTRL EC vs H1
DOWN 115
UP 316
UP
DOWN
−log10(padj)
log2(siERH/siControl)
I
Overlap of hESC self-renewal genes with ERH repressed genes
Genes enriched in hESCs with high self-renewal (Lau et al., 2020)
150
19
43
8
1670
977
472
Endoderm ERH repressed genes
p-value = 6.2e-12
Ectoderm ERH repressed genes
p-value = 2.2e-16
J
Expression of naïve pluripotency genes in siRNA treated pluripotent and differentiated H1 hESCs
NANOGP1
p=1.6e-12
ns
ns
ESRRB
p=0.03
ns
DUSP10
p=8e-15
ns
p=0.04
siControl
siERH
Expression (TPM)
UT
DE
EC
TBX3
ns
p=0.0008
ns
IL6R
p=7e-7
ns
p=0.002
TFCP2L1
p=0.0006
ns
ns
NR5A2
p=5.6e-7
p=6.2e-9
ns

Figure S4

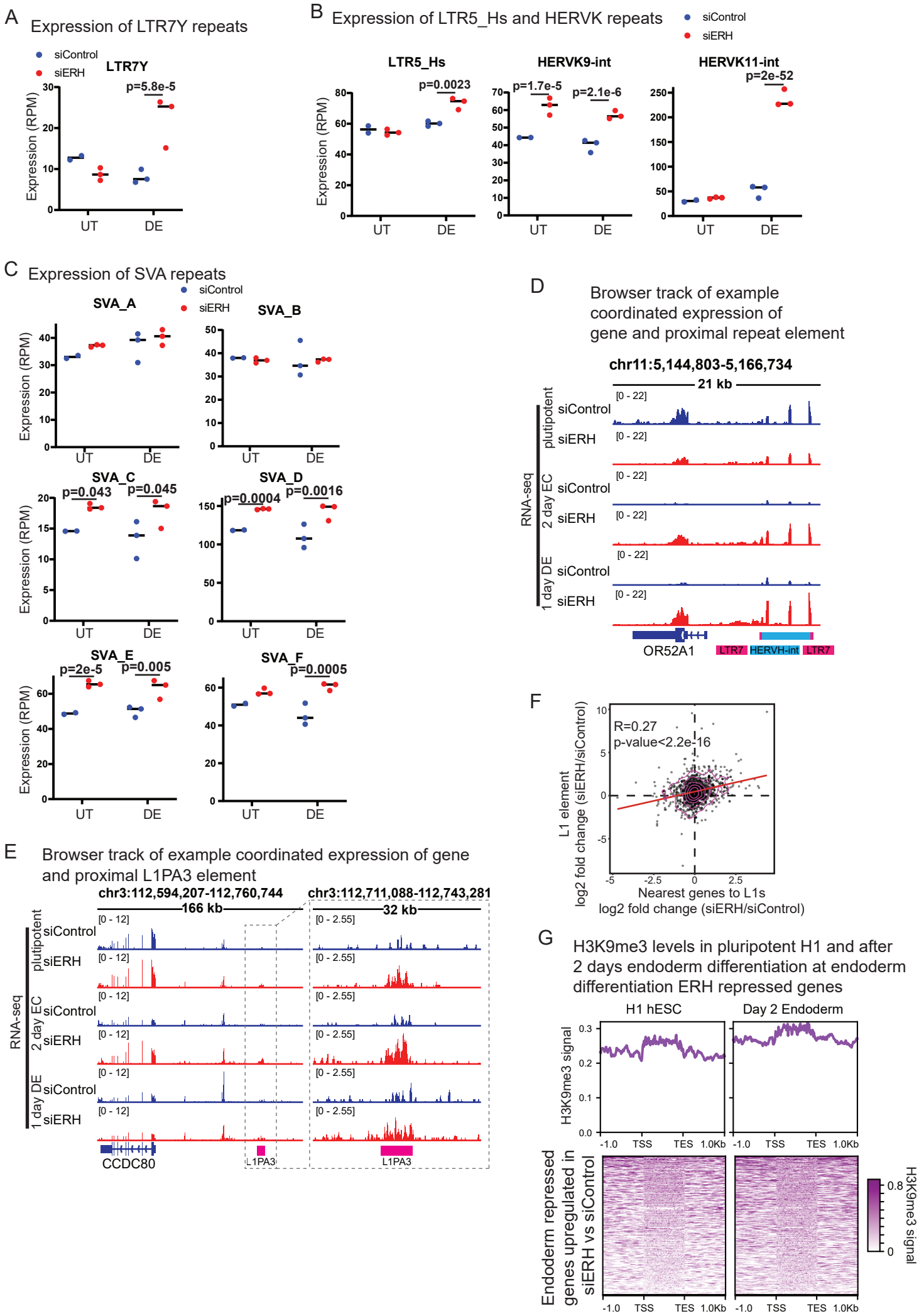
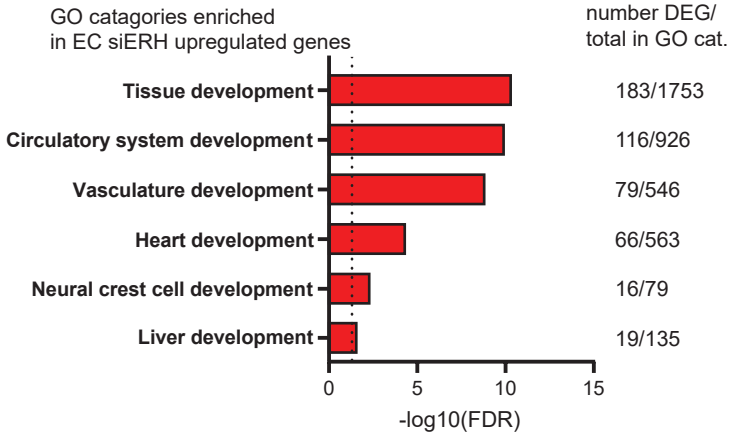

A
Expression of LTR7Y repeats
siControl
siERH
LTR7Y
p=5.8e-5
Expression (RPM)
UT
DE
B
Expression of LTR5_Hs and HERVK repeats
siControl
siERH
LTR5_Hs
p=0.0023
HERVK9-int
p=1.7e-5
p=2.1e-6
HERVK11-int
p=2e-52
Expression (RPM)
UT
DE
C
Expression of SVA repeats
siControl
siERH
SVA_A
SVA_B
SVA_C
p=0.043
p=0.045
SVA_D
p=0.0004
p=0.0016
SVA_E
p=2e-5
p=0.005
SVA_F
p=0.0005
Expression (RPM)
UT
DE
D
Browser track of example coordinated expression of gene and proximal repeat element
chr11:5,144,803-5,166,734
21 kb
RNA-seq
plutipotent
2 day EC
1 day DE
siControl
siERH
[0 - 22]
OR52A1
LTR7
HERVH-int
LTR7
E
Browser track of example coordinated expression of gene and proximal L1PA3 element
chr3:112,594,207-112,760,744
166 kb
chr3:112,711,088-112,743,281
32 kb
RNA-seq
plutipotent
2 day EC
1 day DE
siControl
siERH
[0 - 12]
[0 - 2.55]
CCDC80
L1PA3
L1PA3
F
R=0.27
p-value<2.2e-16
L1 element
log2 fold change (siERH/siControl)
Nearest genes to L1s
log2 fold change (siERH/siControl)
G
H3K9me3 levels in pluripotent H1 and after 2 days endoderm differentiation at endoderm differentiation ERH repressed genes
H1 hESC
Day 2 Endoderm
H3K9me3 signal
-1.0
TSS
TES
1.0Kb
Endoderm repressed genes upregulated in siERH vs siControl
H3K9me3 signal

Figure S5

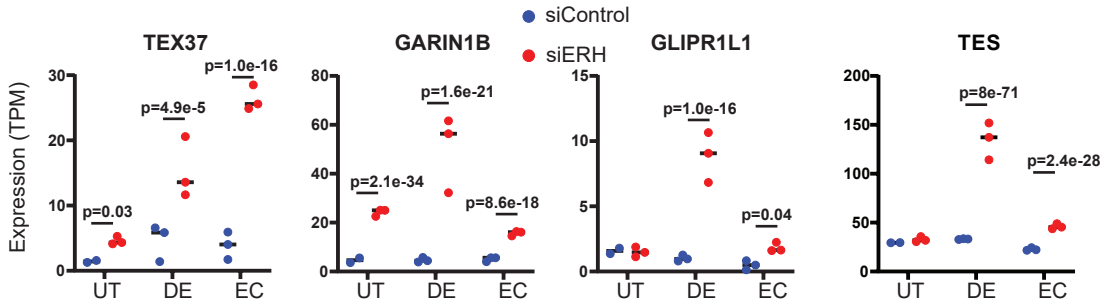
A

GO analysis for statisticaly enrichment of genes upregulated in siERH ectoderm



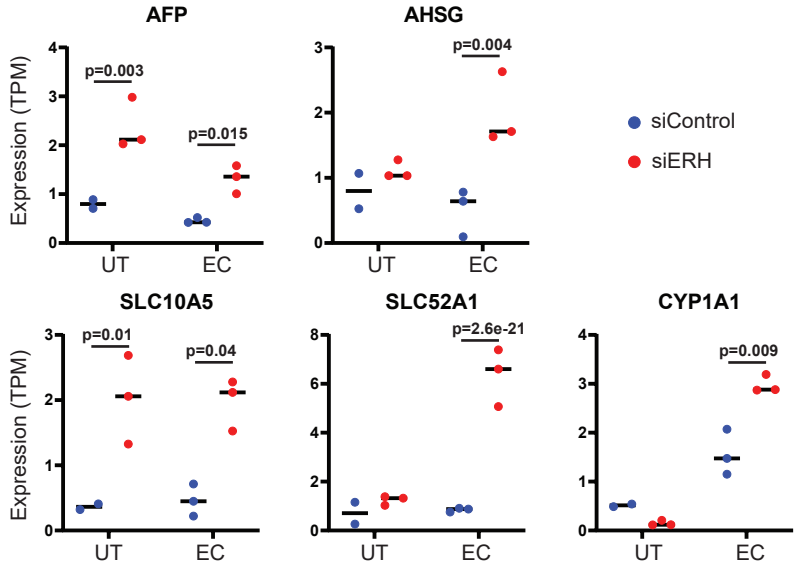

B

Expression of example gametogenic genes

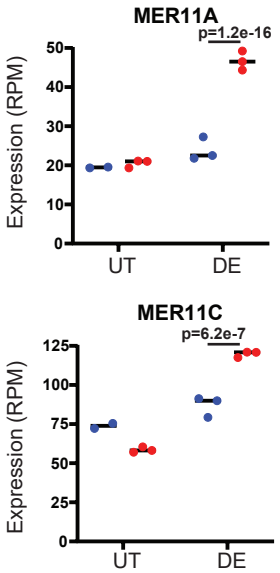



C

Expression of example liver genes in siControl and siERH H1 cells during pluripotency and ectoderm differentiation




D Expression of MER11 repeats




E

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

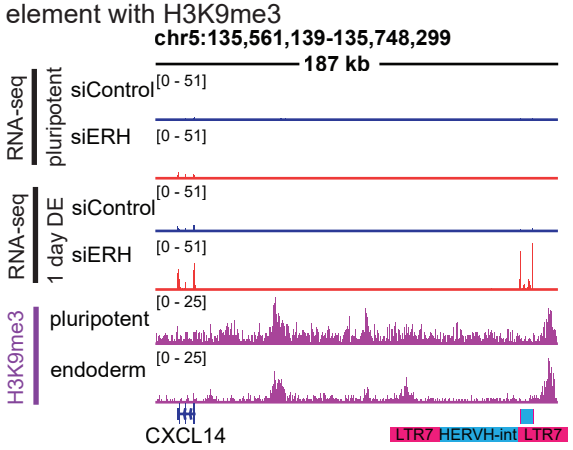
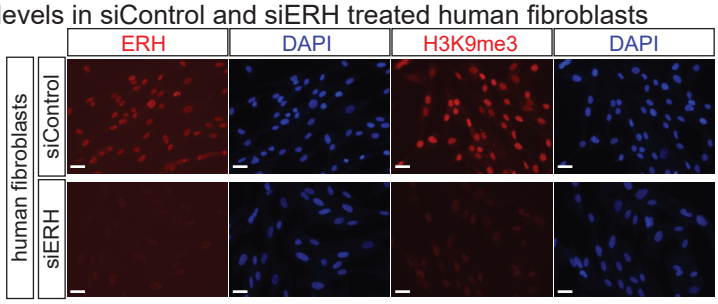
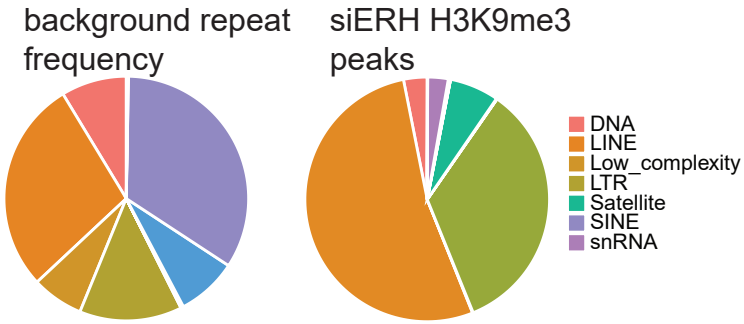


Figure S6

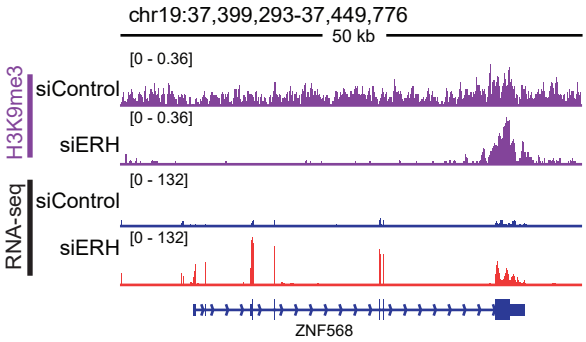
A Immunofluorescence images showing ERH and H3K9me3 levels in siControl and siERH treated human fibroblasts



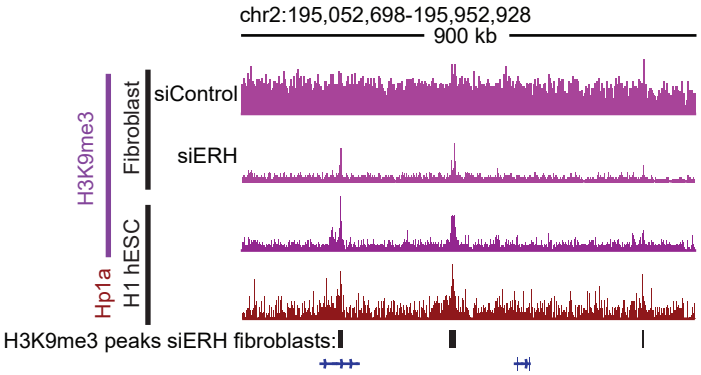

B Proportion of repeat class at siERH fibroblast H3K9me3 peaks vs background abundance




C Browser track of H3K9me3 and expression change at example ZNF gene




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




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

| | p-adj | Motif |
|---|---|---|
| ZNF136 | $<10^{-5}$ | |
| SOX10 | $<10^{-5}$ | |
| ZFX | $<10^{-5}$ | |
| SOX21 | $<10^{-5}$ | |
| SOX3 | $<10^{-5}$ | |
| SOX4 | $<10^{-5}$ | |
| ZNF711 | $<10^{-5}$ | |
| ZNF317 | $<10^{-5}$ | |
| SOX15 | $<10^{-5}$ | |
| SOX9 | $<10^{-5}$ | |

# Figure S7

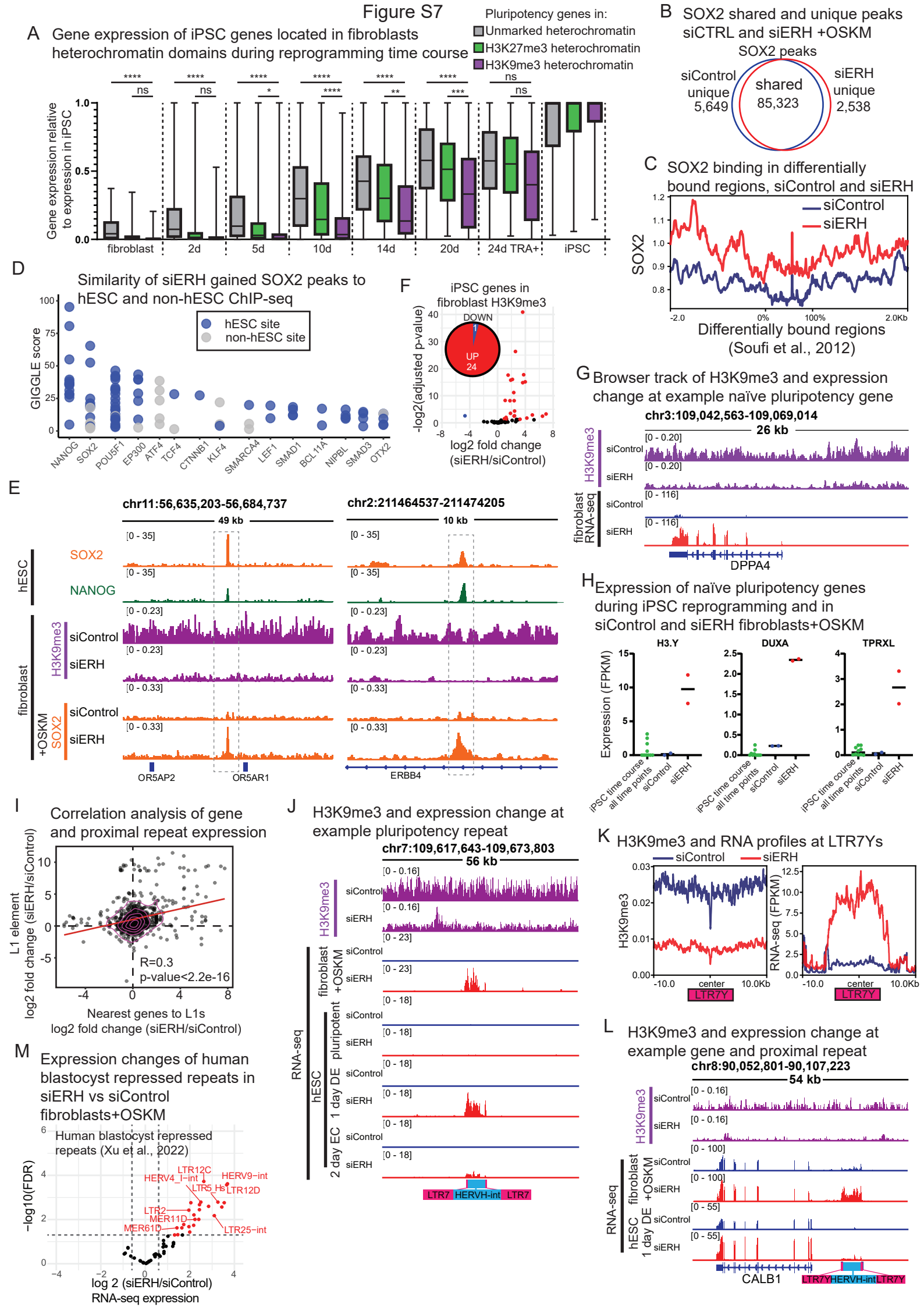
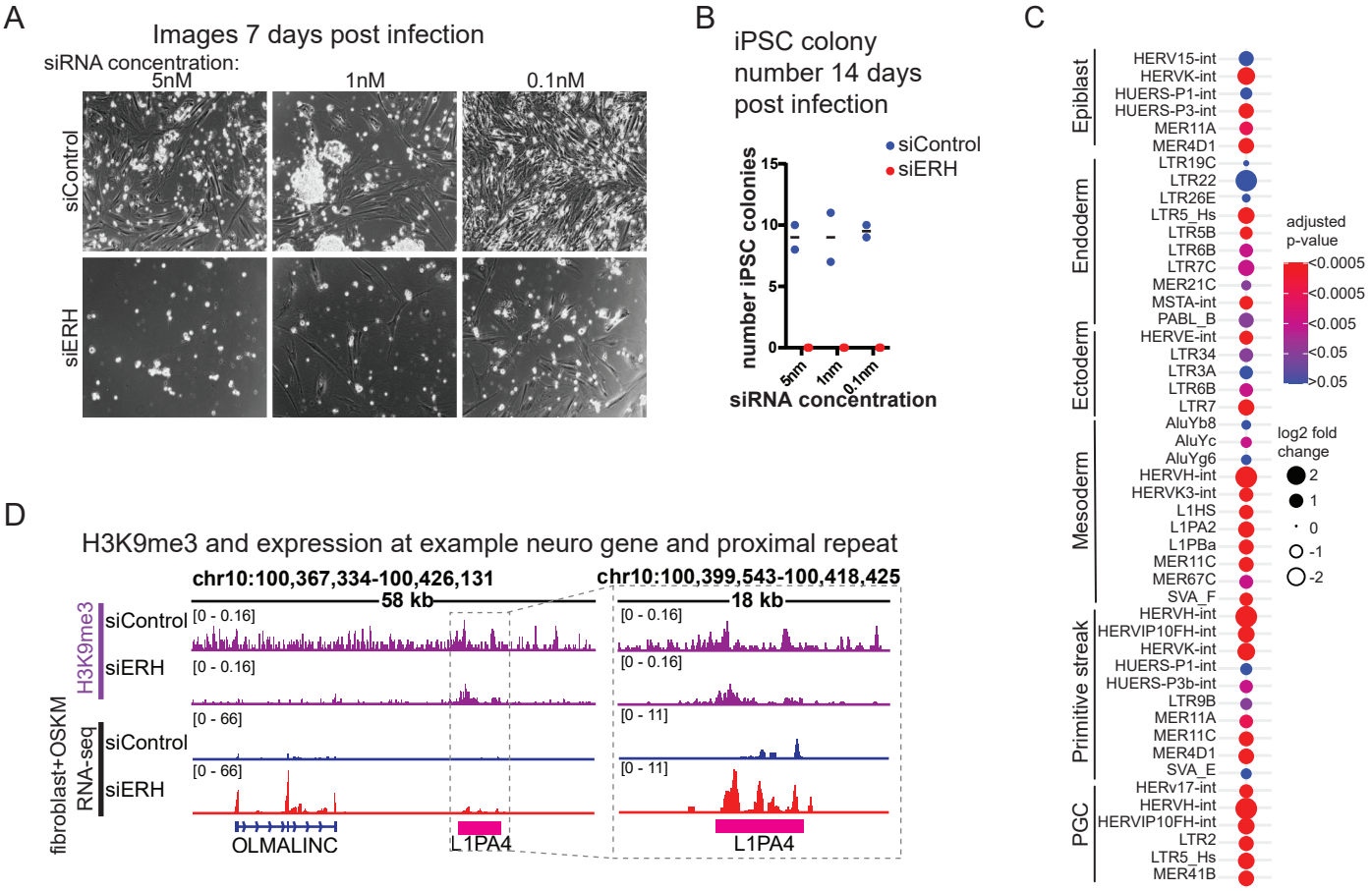

A
Gene expression of iPSC genes located in fibroblasts heterochromatin domains during reprogramming time course
Pluripotency genes in:
Unmarked heterochromatin
H3K27me3 heterochromatin
H3K9me3 heterochromatin
Gene expression relative to expression in iPSC
fibroblast
2d
5d
10d
14d
20d
24d TRA+
iPSC
B
SOX2 shared and unique peaks siCTRL and siERH +OSKM
SOX2 peaks
siControl unique 5,649
shared 85,323
siERH unique 2,538
C
SOX2 binding in differentially bound regions, siControl and siERH
siControl
siERH
SOX2
Differentially bound regions (Soufi et al., 2012)
D
Similarity of siERH gained SOX2 peaks to hESC and non-hESC ChIP-seq
hESC site
non-hESC site
GIGGLE score
NANOG
SOX2
POU5F1
EP300
ATF4
TCF4
CTNNB1
KLF4
SMARCA4
LEF1
SMAD1
BCL11A
NIPBL
SMAD3
OTX2
F
iPSC genes in fibroblast H3K9me3
DOWN
UP 24
-log2(adjusted p-value)
log2 fold change (siERH/siControl)
G
Browser track of H3K9me3 and expression change at example naïve pluripotency gene
chr3:109,042,563-109,069,014
26 kb
H3K9me3
siControl
siERH
fibroblast RNA-seq
DPPA4
E
chr11:56,635,203-56,684,737
49 kb
chr2:211464537-211474205
10 kb
hESC
SOX2
NANOG
fibroblast
H3K9me3
siControl
siERH
+OSKM
SOX2
OR5AP2
OR5AR1
ERBB4
H
Expression of naïve pluripotency genes during iPSC reprogramming and in siControl and siERH fibroblasts+OSKM
H3.Y
DUXA
TPRXL
Expression (FPKM)
iPSC time course
all time points
siControl
siERH
I
Correlation analysis of gene and proximal repeat expression
L1 element log2 fold change (siERH/siControl)
R=0.3
p-value<2.2e-16
Nearest genes to L1s log2 fold change (siERH/siControl)
J
H3K9me3 and expression change at example pluripotency repeat
chr7:109,617,643-109,673,803
56 kb
H3K9me3
siControl
siERH
RNA-seq
fibroblast +OSKM
hESC
pluripotent
1 day DE
2 day EC
LTR7
HERVH-int
LTR7
K
H3K9me3 and RNA profiles at LTR7Ys
siControl
siERH
H3K9me3
RNA-seq (FPKM)
center
LTR7Y
M
Expression changes of human blastocyst repressed repeats in siERH vs siControl fibroblasts+OSKM
Human blastocyst repressed repeats (Xu et al., 2022)
−log10(FDR)
LTR12C
HERV4_I-int
HERV9-int
LTR5_Hs
LTR12D
LTR2
MER11D
MER61D
LTR25-int
log 2 (siERH/siControl) RNA-seq expression
L
H3K9me3 and expression change at example gene and proximal repeat
chr8:90,052,801-90,107,223
54 kb
H3K9me3
siControl
siERH
RNA-seq
fibroblast +OSKM
hESC 1 day DE
CALB1
LTR7YHERVH-intLTR7Y

Figure S8


A
Images 7 days post infection
siRNA concentration:
5nM
1nM
0.1nM
siControl
siERH
B
iPSC colony number 14 days post infection
siControl
siERH
number iPSC colonies
15
10
5
0
5nm
1nm
0.1nm
siRNA concentration
C
Epiblast
Endoderm
Ectoderm
Mesoderm
Primitive streak
PGC
HERV15-int
HERVK-int
HUERS-P1-int
HUERS-P3-int
MER11A
MER4D1
LTR19C
LTR22
LTR26E
LTR5_Hs
LTR5B
LTR6B
LTR7C
MER21C
MSTA-int
PABL_B
HERVE-int
LTR34
LTR3A
LTR6B
LTR7
AluYb8
AluYc
AluYg6
HERVH-int
HERVK3-int
L1HS
L1PA2
L1PBa
MER11C
MER67C
SVA_F
HERVH-int
HERVIP10FH-int
HERVK-int
HUERS-P1-int
HUERS-P3b-int
LTR9B
MER11A
MER11C
MER4D1
SVA_E
HERv17-int
HERVH-int
HERVIP10FH-int
LTR2
LTR5_Hs
MER41B
adjusted p-value
<0.0005
<0.0005
<0.005
<0.05
>0.05
log2 fold change
2
1
0
-1
-2
D
H3K9me3 and expression at example neuro gene and proximal repeat
chr10:100,367,334-100,426,131
58 kb
chr10:100,399,543-100,418,425
18 kb
H3K9me3
siControl
siERH
fibroblast+OSKM RNA-seq
siControl
siERH
[0 - 0.16]
[0 - 0.16]
[0 - 0.16]
[0 - 0.16]
[0 - 66]
[0 - 66]
[0 - 11]
[0 - 11]
OLMALINC
L1PA4
L1PA4