Global hyperactivation of enhancers stabilizes human and mouse naïve pluripotency through inhibition of CDK8/19 Mediator kinases

Cian J. Lynch[1,2], Raquel Bernad[1,2], Ana Martínez-Val[3], Marta N. Shahbazi[4], Sandrina Nóbrega-Pereira[5], Isabel Calvo[1,2], Carmen Blanco-Aparicio[6], Carolina Tarantino[7], Elena Garreta[7], Laia Richart-Ginés[8], Noelia Alcazar[1,2], Osvaldo Graña-Castro[9], Gonzalo Gómez-Lopez[9], Irene Aksoy[10], Maribel Muñoz-Martín[1,2], Sonia Martinez[6], Sagrario Ortega[11], Susana Prieto[12], Elisabeth Simboeck[12], Alain Camasses[12], Camille Stephan-Otto Attolini[13], Agustin F. Fernandez[14], Marta I. Sierra[14], Mario F. Fraga[14], Joaquin Pastor[6], Daniel Fisher[12], Nuria Montserrat[7,15], Pierre Savatier[10], Javier Muñoz[3], Magdalena Zernicka-Goetz[4] and Manuel Serrano[1,2,16,*]

[1] Tumour Suppression Group,
Spanish National Cancer Research Centre (CNIO),
Madrid 28029, Spain.

[2] Cellular Plasticity and Disease Group,
Institute for Research in Biomedicine (IRB Barcelona),
Barcelona Institute of Science and Technology (BIST),
Barcelona 08028, Spain.

[3] ProteoRed-ISCIII Proteomics Unit,
Spanish National Cancer Research Centre (CNIO),
Madrid 28029, Spain.

[4] Zernicka-Goetz Lab,
Mammalian Embryo and Stem Cell Group,
University of Cambridge,
Department of Physiology, Development and Neuroscience,
Downing Street,
Cambridge CB2 3EG, UK.

[5] Department of Medical Sciences and Institute of Biomedicine iBiMED,
University of Aveiro,
3810-193 Aveiro,
Portugal.

[6] Experimental Therapeutics Program,

Spanish National Cancer Research Centre (CNIO),
Madrid 28029, Spain.

[7] Pluripotency for organ regeneration,
Institute for Bioengineering of Catalonia (IBEC),
The Barcelona Institute for Science and Technology (BIST),
Barcelona, Spain.

[8] Maintenance of Transcriptional Repression by Polycomb Proteins,
Institut Curie,
13 Rue Pierre et Marie Curie,
75005 Paris, France.

[9] Bioinformatics Unit,
Spanish National Cancer Research Centre (CNIO),
Madrid 28029, Spain.

[10] Univ Lyon, Université Lyon 1, INSERM U1208,
Stem Cell and Brain Research Institute,
69500 Bron, France.

[11] Transgenic Mice Unit,
Spanish National Cancer Research Centre (CNIO),
Madrid 28029, Spain.

[12] Laboratory of Nuclear Control of Cell Proliferation,
Institute of Molecular Genetics of Montpellier,
1919 Route de Mende,
34293 Montpellier, Cedex 5, France.

[13] Bioinformatics-Biostatistics Unit,
Institute for Research in Biomedicine (IRB Barcelona),
Barcelona Institute of Science and Technology (BIST),
Barcelona 08028, Spain.

[14] Cancer Epigenetics and Nanomedicine Laboratory,
Nanomaterials and Nanotechnology Research Center (CINN-CSIC),
Institute of Oncology of Asturias (IUOPA), ISPA-Hospital Universitario
Central de Asturias (HUCA), Universidad de Oviedo, Oviedo, Spain.

[15] Centro de Investigación Biomédica en Red en Bioingeniería,
Biomateriales y Nanomedicina, Barcelona, Spain.

[16] Catalan Institution for Research and Advanced Studies (ICREA)
Barcelona 08010, Spain

* *Correspondence*: manuel.serrano@irbbarcelona.org

Abstract

Pluripotent stem cells (PSCs) transition between cell states *in vitro* and reflect developmental changes in the early embryo. PSCs can be stabilized in the naïve state by blocking extracellular differentiation stimuli, particularly FGF-MEK signaling. Here, we report that multiple features of the naïve state in human and mouse PSCs can be recapitulated without affecting FGF-MEK-signaling or global DNA methylation. Mechanistically, chemical inhibition of CDK8 and CDK19 kinases removes their ability to repress the Mediator complex at enhancers. Thus CDK8/19 inhibition increases Mediator-driven recruitment of RNA Pol II to promoters and enhancers. This efficiently stabilizes the naïve transcriptional program and confers resistance to enhancer perturbation by BRD4 inhibition. Moreover, naïve pluripotency during embryonic development coincides with a reduction in CDK8/19. We conclude that global hyperactivation of enhancers drives naïve pluripotency, and this can be achieved *in vitro* by inhibiting CDK8/19 kinase activity. These principles may apply to other contexts of cellular plasticity.

Introduction

Each cell-type contains a unique repertoire of active enhancer complexes at specific DNA regions, which arise by high concentration of lineage-specific transcription factors and signaling pathways[1–3]. The Mediator complex is enriched at enhancers, where it integrates multiple upstream signals and recruits RNA polymerase II (RNA Pol II) to nearby and distant promoters[4–6]. A large fraction of Mediator and other transcriptional machinery clusters within a small number of unusually long multipartite enhancers, known as super-enhancers (SEs)[2,3,7]. SEs drive high expression of the master transcription factors which maintain cell identity, yet SEs also possess vulnerability to perturbation via sharp transitions in their phase separation[2,3,7]. Given Mediator´s central role in enhancer-driven transcription, modulation of its activity may influence cellular identity and plasticity[7,8]. Indeed, global enhancer activation was identified across multiple human cancers[9]. The kinase CDK8 and its similar but poorly studied paralog CDK19 represent the only enzymatic activity of the 30-subunit Mediator complex[4–6]. This CDK8/19 sub-module can regulate RNA Pol II recruitment via its kinase activity, although it may also sterically hinder the association between Mediator and RNA Pol II[5,6,10–13]. Accordingly, chemical inhibition of CDK8/19 results in global hyperactivation of enhancer function within cancer cells[14]. Additional layers of complexity include that CDK8 can phosphorylate multiple Mediator subunits, the RNA Pol II C-terminal regulatory domain, chromatin regulators and transcription factors[4–6,11–17].

Pluripotent stem cells (PSC) provide a prototypical model of cellular plasticity, whose transcriptional program can be stabilized, extinguished or re-captured[18–23]. While human PSCs offer great therapeutic promise, successful clinical applications remain few, as human pluripotency is less characterized and less stable in vitro, than in mice[21,22,24]. Chemical inhibition of MEK and GSK3 kinases with a two-inhibitor cocktail known as “2i” shields mouse PSCs cells from extra-cellular differentiation in a state known as naïve pluripotency[25]. Mouse PSCs cultured in 2i (referred to as “2i-naïve” cells) phenocopy the stable and homogenous state of undifferentiated naïve pluripotency that exists transiently in the E4.5 pre-implantation embryo epiblast[18,19,25]. In contrast, culture of PSCs

without 2i shifts cell identity towards post-implantation epiblast ~E6.5, also known as primed pluripotency[18,19,23]. Enhancer destabilization by chemical blockade of BRD4, a key component of enhancers and SEs, triggers loss of Mediator-driven gene expression in many cell types, and induces differentiation in primed PSCs[26–28]. Remarkably, 2i-naïve PSCs are highly resistant to enhancer destabilization[28], indicating an association between naïve pluripotency and enhancer stability/resilience. MEK-inhibition has been implicated upstream of potent and rapid reconfiguration of the transcriptome, proteome, and DNA methylome, within embryonic or 2i-naïve pluripotency[18–20,23]. However, the molecular mediators of 2i that are responsible for enhancer stabilization remain unclear.

Here, we assess the impact of Mediator-CDK8/19-kinase inhibition on PSC identity, in order to elucidate the transcriptional basis of their plasticity. In summary, stimulating Mediator through its kinase module represses differentiation, favors self-renewal, and upregulates pre-implantation naïve epiblast gene expression in mouse and in human.

RESULTS

Inhibition of Mediator kinase stabilizes mouse naïve pluripotency

GFP knock-in reporters at key stem cell marker genes such as *Nanog* represent well-established and precise indicators of the naïve (GFP^{high}) and primed states (GFP^{low})[18,22,29]. For example, in 2i-naïve state, *Nanog* promoter activity is enhanced, yielding a characteristically homogenous *Nanog*-GFP^{high} cell expression pattern and uniform dome-shaped colonies (**Fig.1A-C, and Extended Data Fig.1A**). In contrast, the *Nanog* promoter is metastable in primed state PSCs, reversibly oscillating between high and low activity, presenting a heterogeneous *Nanog*-GFP expression pattern and flattened diffuse colonies, indicative of a general underlying switch in transcriptional program[18,20,23,29,30]. The BRD4 inhibitor JQ1 destabilizes enhancers and resulted in colony flattening and GFP^{low} status (**Fig.1A**), as reported[26–28]. In this experimental setting, we tested the effect of manipulating the transcriptional cyclin-dependent kinases (CDK7, CDK8/19 and CDK9) with a panel of small molecule inhibitors. Several potent

and structurally-unrelated CDK8/19 inhibitors had a positive effect, inducing the formation of homogenous dome-shaped colonies, and upregulating both the *Nanog*-GFP reporter and endogenous *Nanog* expression, similar to PSC in the 2i-naïve state (**Fig.1A-E; Extended Data Fig.1A; Supplementary Table 1**), while inhibition of CDK7 or CDK9 did not. Potency and selectivity of CDK8/19-inhibitors, commercially available or developed in-house, were assessed by multiple methods: (i) selectivity was suggested by a KinomeScan panel of 456 kinases; (ii) Lanthascreen assays demonstrated inhibitory activity at nanomolar concentrations against pure recombinant CDK8/CCNC and CDK19/CCNC; (iii) luciferase reporter cell assays (TOP-FLASH); and (iv) potent inhibition of STAT1-Ser727 phosphorylation in human PSCs, a well-documented CDK8 target site[11,14,16,31] (**Fig.1F; Extended Data Fig.1B; Supplementary Table 1**; also: **Supplementary Information** file). Based on these data, we focused on the CNIO molecule CDK8/19i-ETP-47799, which was the most effective at improving mouse PSCs, and we will refer to it simply as CDK8/19i (**Fig.1A,B and Extended Data Fig.1A**; for the structure and characterization of this inhibitor, and comparison with other inhibitors used in this study: **Supplementary Table 1; Supplementary Information**). In addition to the improvements in *Nanog*-GFP profile and colony morphology mentioned before, the effect of CDK8/19i on mouse PSCs resembled 2i in three additional ways: (i) it was observed in serum- and serum-free based media (**Fig.1A and Extended Data Fig.1A**); (ii) it was reversible upon CDK8/19i-withdrawal with kinetics similar to that of 2i-removal (**Extended Data Fig.1C**); and, (iii) upon removal of LIF or inhibition of LIF-signaling with a JAK inhibitor, the presence of CDK8/19i delayed downregulation of *Nanog*-GFP expression (**Extended Data Fig.1D,E**). We conclude that inhibition of Mediator kinase CDK8/19 shifts mouse PSC morphology and *Nanog* expression towards their characteristic status in the naïve state[18,23,29].

Exploring genetic validation, depletion via shRNA-knockdown of CDK8, CDK19, but most successfully, their regulatory partner cyclin C (CCNC; essential for full kinase activity[8]), led to upregulation of *Nanog* expression and naïve-like colony morphology (**Fig.1G and Extended Data Fig.1F,G**). In another genetic approach, CDK8 and CDK19 double-knockout (dKO) mouse PSCs were

generated (**Extended Data Fig.1H-K**). CDK8/19-dKO PSCs could self-renew indefinitely, but did not acquire naïve morphological features or *Nanog* upregulation, and no longer responded to CDK8/19-inhibitors (**Extended Data Fig.1L,M**). This suggested that the beneficial effects observed may require the physical presence of the inactive-kinase. In agreement, we found that CDK8/19-dKO PSCs reconstituted with exogenous CDK8 rescued the ability to respond to the CDK8/19 inhibitor molecule, observed by naïve morphological features and *Nanog, Klf4,* and *Oct4* upregulation (**Extended Data Fig.1L,M**). Moreover, CDK8/19-dKO PSCs reconstituted with a CDK8-kinase dead mutant (CDK8-KD; D173A) displayed homogenous naïve colony morphology, high expression of naïve-state markers (**Fig.1H-M**), and downregulation of *Fgf5*, a key marker of the primed state[18,19,23] (**Fig.1J**); all without the need of any chemical inhibitor and despite maintaining active MEK-ERK signaling (**Fig.1K**). Thus, CDK8/19-dKO cells expressing CDK8-KD phenocopy the effects of chemical inhibition of CDK8/19. Lastly, post-implantation epiblast stem cells (EpiSC; cultured with FGF2/Activin) are a more developmentally advanced primed state than mouse PSCs in serum/LIF[18,19]. Interestingly, EpiSC exogenously expressing CDK8-KD lost *Fgf5*, upregulated *Nanog, Rex1,* and *Klf4*, and formed dome-shaped colonies with high alkaline phosphatase staining, altogether characteristic of conversion to the naïve state (**Fig.2A and Extended Data Fig.2A**). In summary, CDK8 kinase inhibition is sufficient to promote key characteristics of naïve pluripotency, despite the continued presence of MEK-ERK signaling.

Long-term culture of mouse PSCs (>10 passages) in CDK8/19i maintained their upregulation of naïve features, including colony morphology, high alkaline phosphatase, *Nanog*-GFP^{high}, high endogenous *Nanog*, high ICAM1 cell surface expression, and nuclear localization of TFE3[21,24,32–34] (**Fig.2B-D and Extended Data Fig. 2B,C**). Long-term CDK8/19i-adapted PSCs displayed typical developmental capacity following inhibitor withdrawal, specifically, retinoic-acid-induced differentiation, embryoid body cardiac centre formation, spheroid polarization and lumenogenesis[35], generation of teratomas containing three germ layers, and robust chimera contribution after morula aggregation and blastocyst micro-injection assays (traced by constitutive GFP or RFP) evaluated at E4.5,

E7.5, E14.5, and fully-developed adults which subsequently completed germline transmission (**Fig.2D-I and Extended Data Fig.2D-F**). Of note, continued presence of CDK8/19i impaired the early developmental events[35] of polarization and lumenogenesis *in vitro* (**Fig.2E**), an observation discussed further below. Thus, PSCs long-term adapted to CDK8/19i maintain their upregulation of naïve features, their self-renewal, and their developmental capacity.

CDK8/19i induces and stabilizes the naïve state in human PSCs

We tested CDK8/19i on human stem cell identity. STAT3 overexpression plus 2i/LIF induces the human naïve state[36], and we observed that CDK8/19i could replace 2i in this system (**Fig.2J**). Even in the absence of STAT3 overexpression, other transgenes, or chemicals, CDK8/19i treatment progressively converted human iPS colonies from flat and primed-like, to dome-shaped naïve-like bi-refringent morphology. This was observed for a total of 7 human PSC lines treated with 0.4μM or 1.1μM CDK8/19i/LIF for 2-3 weeks (**Fig.2K and Extended Data Fig.2G**), including human iPSCs carrying a specific *HERVH*-GFP reporter insertion that marks human naïve cell identity[37] (**Fig.2K and Extended Data Fig.2H**). A 2i-based chemical cocktail combined with selection by cell-sorting (abbreviated as 2i p38iJNKi) induced naïve colony morphology, as expected[33,37], with homogeneous high *HERVH*-GFP (**Extended Data Fig.2H**). Interestingly, treatment with CDK8/19 inhibitors (CNIO-47799 or SnxA) also produced morphological conversion and increased GFP, similar to 2i p38iJNKi (**Fig.2L and Extended Data Fig.2H**). The changes induced by CDK8/19 inhibition were gradual, required no selection upon passage (sorting or manual picking), required no additional supplements except rhLIF, and were stable in the continuous presence of the inhibitor. In contrast, CDK7 inhibition failed to change colony morphology or GFP fluorescence, and produced cell death (**Extended Data Fig.2H**). Culture of human PSCs in CDK8/19i increased their clonogenicity, alkaline phosphatase intensity, and pluripotency markers[32–34,38] NANOG, OCT4, SSEA4, TRA1-81, TFCP2L1, and KLF17 (**Fig.2M, 3A, and Extended Data Fig.2I and 3A-C**). MYC, known to be reduced in naïve cells[25,32], was also reduced in cells maintained in CDK8/19i (**Fig.3A**). Therefore, similar to

observations in mouse PSCs above, treatment of human PSCs with CDK8/19i establishes features characteristic of the naïve state.

Developmental potential of CDK8/19i-adapted human PSCs

Chemical induction of the human naïve state can trigger genomic instability, severely impairing developmental potential[24,39]. We found that CDK8/19i-adapted human PSCs (five lines) had normal karyotype over >16 passages (**Extended Data Fig.3D**) and, upon inhibitor withdrawal, maintained the capacity to contribute towards all three embryonic germ layers by embryoid body differentiation *in vitro* and by teratoma assay *in vivo* (**Fig.3B-D, and Supplementary Table 1**), comparable to control primed cells. Preimplantation interspecies chimerism tests for naïve-specific properties, namely, capacity for clonal survival in a host embryo[40,41]. We tested CDK8/19i-adapted human iPSCs carrying a constitutive Tomato-red marker for human-rabbit interspecies chimerism by micro-injecting them into E2.5 rabbit morulae. Interestingly, the presence of human cells (Tomato-positive) was detected 72h later in up to 50% of the injected rabbit blastocysts (**Fig.3E**). In contrast, human PSC in the primed state were unable to integrate or survive in rabbit embryos (0/24 rabbit embryos), similar to previous reports for primed state human PSCs within the embryos of mice, pigs, and cattle[40,41]. In summary, long-term adaptation of human PSCs to CDK8/19i stabilizes naïve pluripotency while preserving their developmental potential. We conclude that the role of CDK8/19 in pluripotency is conserved in mouse and humans, and therefore presumably, across mammals.

CDK8/19i resets the transcriptome and proteome similar to 2i

Using RNA-seq, we compared global gene expression in mouse PSCs long-term adapted to CDK8/19i versus 2i. Overall, CDK8/19i altered gene expression with a magnitude similar to 2i conditions, and with a highly significant overlap in the identity and biological functions of genes up- or down-regulated, in both serum-containing and serum-free media (**Fig.4A, Extended Data Fig.3E,F, and Supplementary Table 2**). Compared to control primed conditions, naïve pluripotency markers were enhanced in CDK8/19i and 2i (**Fig.4B and Extended**

329 **Data Fig.3G**), while differentiation markers were globally downregulated in
330 CDK8/19i and 2i (**Supplementary Table 2**).
331
332 Endogenous retrovirus (ERV) expression is highly stage-specific during
333 mammalian pre-implantation and precisely defines naïve and primed PSC
334 identity[42–46]. The transcriptomic overlap between CDK8/19i or 2i treatments
335 extended to ERVs, with similar viral families significantly up- or downregulated in
336 mouse PSCs (**Fig.4C and Supplementary Table 2**). In particular, LINE L1
337 families, each with thousands of copies across the genome, were regulated in
338 close parallel, displaying highly similar alteration in CDK8/19i and 2i-naïve states
339 (**Extended Data Fig.3G,H**). Another aspect of the plasticity of mouse PSCs is
340 their ability to transition to a 2-cell-like (2C) state, specifically marked by hyper-
341 activation of the MERVL family of ERVs and by *Zscan4c* expression[46,47].
342 Stabilization of the naïve state with 2i impairs the 2C-like fluctuation[46,47]. This was
343 also observed in CDK8/19i-treated PSCs, using multiple 2C-markers, including
344 *MERVL* and *Zscan4c*, and *MERVL*-Tomato and *Zscan4c*-eGFP 2C-reporter
345 models (**Fig.4D, and Extended Data Fig.3I-M and 4A-C**). Finally, our CDK8/19i
346 and 2i transcriptomic data correlated with published transcriptomes[48–51] from
347 independent studies of 2i-naïve mouse PSCs and the transcriptome of E4.5
348 epiblast single-cells[52] (**Fig. 4E, and Extended Data Fig.4D,E**).
349
350 RNA-seq analyses of human PSCs adapted to CDK8/19i, or a 2i-based
351 naïve cocktail, overlapped significantly (**Fig.4F; Supplementary Table 3**).
352 Markers of human and primate pre-implantation epiblasts and *in vitro* naïve
353 human PSCs[32–39] were upregulated by CDK8/19i, while differentiation markers
354 were repressed[53–59] (**Fig.4G, Extended Data Fig.4F-J, and Supplementary**
355 **Table 3**). Moreover, the global human ERV transcriptomes of CDK8/19i- or 2i-
356 adapted cells overlapped extensively, including upregulation of the SVA, LTR7,
357 and HERV families (**Fig.4H-J and Supplementary Table 3**), consistent with
358 reports of ERV expression in human and primate naïve PSCs and pre-
359 implantation epiblast[43–45]. Lastly, we observed high correlation between
360 CDK8/19i-adapted human PSCs and RNA expression datasets from seven

independent studies in human and primate PSCs in the *in vitro* naïve state and in embryo naïve epiblast single-cell analyses[39,53–60] (**Fig.4K**).

While PSC plasticity has been explored in terms of RNA expression, its proteome remains relatively ill-defined. We analyzed the proteome of mouse PSCs in control serum/LIF versus 2i-naïve or CDK8/19i-adapted conditions. Across five mouse PSC lines, CDK8/19i altered the expression levels of 465 proteins, 159 (34%) of which changed in the same direction in 2i conditions (**Fig.4L,M; Extended Data Fig.5A,B; Supplementary Table 4**). Importantly, among the overlapping changes in both 2i-naïve and CDK8/19i conditions, we noted key pluripotency regulators such as KLF4, and metabolic pathways such as oxidative phosphorylation, featured amongst the most-upregulated; while LIN28A, MYC-target genes, and differentiation markers, were down-regulated (**Fig.4M; Extended Data Fig.5C; Supplementary Table 4**). In addition, proteomic changes in 2i and CDK8/19i significantly correlated with the transcriptomic changes observed (**Extended Data Fig.5D,E**).

In summary, CDK8/19i upregulates pluripotency markers, reshapes the endogenous retroviral transcriptome, and represses differentiation markers, in a manner similar to the transcriptomic and proteomic resetting observed in previous studies of naïve pluripotency, *in vitro* and *in vivo*, in mouse and in human.

CDK8/19i does not reset global DNA methylation levels

Many 2i-based chemical cocktails induce global DNA hypomethylation, both in mouse and human PSCs[21]. This has been attributed to MEK-dependent stabilization of UHRF1, a critical factor for the recruitment of DNMT1 to DNA[61]. Importantly, the pattern of demethylation induced by 2i diverges significantly compared to the pre-implantation naïve epiblast state, and is associated with PSCs exhibiting genomic instability, chromosomal defects and loss of pluripotency[24,39,62,63]. Recent 2i-variant cocktails (with partial MEK-inhibition) offer the advantage of largely preserving global DNA methylation[62–64]. Importantly, neither mouse nor human CDK8/19i-adapted PSCs showed evidence of global

DNA hypomethylation (**Fig.5A,B**). In support of global DNA hypomethylation, 2i or MEK-inhibition-alone, induced demethylation of LINE L1 repeat regions (**Fig.5C**), and major satellite regions (**Extended Data Fig.5F**), but had no effect on methylation of IAP repeats (**Extended Data Fig.5G**), all as previously reported[65]. In contrast, CDK8/19i did not reduce methylation at any of these mouse repeat elements (**Fig.5C; Extended Data Fig.5F,G**), or UHRF1 levels (**Supplementary Table 4**). Thus, CDK8/19i induces naïve features in the absence of global DNA hypomethylation, and this is likely due to its lack of MEK inhibition (**Fig.2**), or Uhrf1 downregulation. By not recapitulating the partial demethylation of the naïve epiblast, CDK8/19i has the advantage of preserving chromosomal stability and pluripotency upon cell expansion (**Fig.2,3**), which is particularly relevant for naïve human PSCs. This is in line with variant media cocktails based on minimizing MEK-inhibition both in mouse and human naïve PSCs[62–64].

X-chromosome reactivation status is another molecular signature reported in human naïve pluripotency during MEK-inhibition[21,66,67], which may be inferred by assessing XIST RNA expression in female cells. However, analysis by qPCR revealed very low XIST expression in our primed human PSCs (**Extended Data Fig.5H**), suggesting that erosion of X-silencing may have already occurred in the parental cells under primed conditions, as observed previously[67]. Notably, some 2i-based cocktails reactivate XIST expression even in X-eroded primed human PSCs[66,67], but this was not the case of our CDK8/19i-adapted cells (**Extended Data Fig.5H**). In summary, CDK8/19i treatment does not recapitulate the reactivation of XIST in X-eroded primed cells, indicating another distinction with most human media cocktails based on MEK-inhibition.

CDK8/19i induces phospho-changes similar to 2i

We assessed the phospho-proteome of mouse PSCs just 15min after exposure to CDK8/19i or 2i, to explain their phenotypic similarity. Strikingly, out of 622 phospho-sites altered, 495 (79.6%) were similarly regulated by CDK8/19i and 2i (**Fig.5D,E**). The co-regulated phospho-sites occurred on proteins heavily enriched for functions in transcriptional regulation and key stem cell signaling

pathways (**Fig.5E; Extended Data Fig.5I; Supplementary Table 5**). Note, CDK8/19i did not inhibit the kinase activity of purified recombinant GSK3 or MEK (**Supplementary Table 1**), and CDK8/19i did not reduce the relative levels of phospho-ERK (**Fig.1K; Fig.5F,G; Extended Data Fig.5J**). However, 2i-treatment reduced CDK8/19 kinase activity (**Fig.5F; Extended Data Fig.5K**), and moderately downregulated CDK8 protein levels (**Extended Data Fig.5L**). These data suggest that CDK8/19 inhibition occurs downstream of 2i, such that both treatments result in highly overlapping phospho-site changes.

CDK8/19i resets global RNA Pol II loading similar to 2i

2i and CDK8/19i induce similar phospho-proteomic changes (converging on transcriptional machinery) and similar transcriptomic changes. Thus, to understand how CDK8/19-inhibition phenocopies the transcriptome of 2i-induced naïve pluripotency, we investigated their global regulation of RNA Pol II abundance on chromatin by ChIP-seq in mouse PSCs. Overall, total- and Ser5-phosphorylated-RNA Pol II genomic distribution was consistent with published resources[23,68,69] (ENCODE: https://www.encodeproject.org/). However, we observed that 2i globally increases RNA Pol II binding to promoters (**Fig.5H,I; Extended Data Fig.6A-C**), which was confirmed by re-analyzing independent data[23,69]. Remarkably, this global effect of 2i was phenocopied by CDK8/19i, regarding both total- and Ser5-phosphorylated-RNA Pol II (**Fig.5H,I; Extended Data Fig.6A-C**). We measured RNA Pol II abundance in the promoter, gene body and transcription termination site (TTS) for each gene (**Fig.5J; Extended Data Fig.6D; Supplementary Table 6**). Consistent with previous analyses in mouse PSCs[23,68], most genes (90%) possessed a promoter to gene body loading ratio >2.0 (**Fig.5J; Supplementary Table 6**). Comparison of RNA Pol II ratios between the promoter, gene body, or termination sub-regions of each gene indicated that 2i induces an increase in RNA Pol II binding selectively to the promoter region (**Fig.5J; Extended Data Fig.6D**). Importantly, this was recapitulated by CDK8/19i, increasing RNA Pol II binding to promoters at a similar magnitude to that observed in 2i-induced naïve pluripotency, following a gene-specific pattern (**Fig.5J; Fig6A-F; Supplementary Table 6**). Therefore, 2i- and CDK8/19i-induced naïve pluripotency are accompanied by widespread

accumulation of RNA Pol II abundance at promoters. We also observed a correlation between changes in RNA Pol II promoter abundance in 2i or in CDK8/19i conditions, and changes in mRNA expression for each gene (**Extended Data Fig.6E-I**). In summary, gene-specific changes in RNA Pol II promoter loading may explain a significant proportion of the mRNA expression profile characteristic of 2i- or CDK8/19i-induced naïve pluripotency.

CDK8/19i and 2i trigger activation of super-enhancers

The primary role of Mediator is at enhancers, regulating RNA Pol II recruitment to promoters[4–6]. Using published ChIP-seq datasets[2,3] (**Supplementary Table 7; Source Data**), we confirmed that CDK8/19 was enriched at promoter, typical enhancer (TE) and super-enhancer (SE) regions as previously defined in mouse PSCs[2] (**Fig.6G; Extended Data Fig.7A-C**). A strong correlation existed between the abundance of CDK8/19, Mediator subunits, and other factors critical for enhancer activity[1,6] (such as, p300, CBP, Pol II, or BRD4) (**Extended Data Fig.7D**); the highest levels of CDK8/19 occurred within SE-regions (**Fig.6G; Supplementary Table 7; Source Data**); and lastly, putative target genes proximal to genomic CDK8/19 binding loci were highly enriched in preimplantation functions characteristic of pluripotent cell identity (**Extended Data Fig.8A-C**). Thus, we hypothesized that in mouse PSCs, CDK8/19-inhibition might act through Mediator to trigger changes in enhancer activity, explaining the observed increase of RNA Pol II loading at promoters and regulation of pluripotent states. Since CDK8/19 protein was particularly enriched at SE-regions, we examined the impact of CDK8/19i on SE function. Enhancers contain RNA Pol II which transcribes enhancer-RNAs (eRNAs), a process that faithfully reflects enhancer activity[4,8,70]. Therefore, we measured the effect of CDK8/19i or 2i on the levels of RNA Pol II and eRNAs at SEs. Importantly, the abundance of RNA Pol II was selectively increased at CDK8/19 binding sites and, accordingly, RNA Pol II recruitment was also preferentially increased at SEs compared to TEs (**Fig.6H,I; Extended Data Fig.8D**). Consistent with this, mouse PSCs treated with 2i or CDK8/19i displayed elevated enhancer-derived eRNA levels, and RNA Pol II abundance, within enhancers specific for the naïve-state[71] (**Fig.7A; Extended Data Fig.8E,F**). Induction of naïve-specific eRNAs and naïve marker

493 genes was an early event, occurring within 48h of adding 2i or CDK8/19i, and it
494 was rapidly reversible (**Fig.7A; Extended Data Fig.8G**). Lastly, consistent with
495 naïve-specific enhancer activation, the expression levels of SE target-genes were
496 preferentially upregulated in both 2i and CDK8/19i (**Fig.7B,C; Extended Data**
497 **Fig.8H**). We conclude that in PSCs, CDK8/19i and 2i hyper-activate existing SE,
498 and upregulate SE target-genes, in a manner which reinforces naïve
499 pluripotency.
500

501 **CDK8/19 inhibition compensates BRD4 inhibition**

502 Loss of Mediator function preferentially decreases expression of enhancer target
503 genes across multiple cell types[4–6,26–28]. In particular, BRD4-inhibition in primed
504 state PSCs decreases the ability of Mediator to recruit RNA Pol II, and this
505 results in loss of Mediator-driven transcription, collapse of pluripotency gene
506 expression, and differentiation[27,28] (**Fig.1A**). Compared to primed PSCs, naïve
507 PSCs are highly resistant to the decreased Mediator activity and enhancer
508 destabilization induced by BRD4-inhibition[28]. Interestingly, mouse PSCs lacking
509 endogenous CDK8 and CDK19 and reconstituted with kinase-dead CDK8 were
510 resistant to enhancer destabilization by BRD4-inhibition for 10 passages (>3
511 weeks), maintaining naïve morphology, and high expression of alkaline
512 phosphatase, naïve-specific pluripotency markers, and naïve-specific eRNAs,
513 similar to 2i-naïve PSCs (**Fig.7D,E; Extended Data Fig.8I,J**). Thus, PSCs
514 expressing kinase-dead CDK8 phenocopy the robust resistance to enhancer
515 destabilization characteristic of 2i-naïve PSCs.
516

517 **Roles of CDK8/19 during early embryonic development**

518 Given our observations that CDK8/19 inhibition stabilizes naïve pluripotency, we
519 investigated CDK8/19 function during early embryonic development. We focused
520 on CDK8, which we found highly expressed compared to CDK19, both in mouse
521 and human PSCs (**Extended Data Fig.1K,9A**). Using a CDK8-specific antibody
522 (**Extended Data Fig.1J**), we detected CDK8 protein from the mouse zygote to
523 morula (**Extended Data Fig.9B**). Consistent with this, CDK8-knockout zygotes
524 cannot progress beyond 4-8 cell stage[72], and we observed that CDK8/19i

525 impaired the progression of 1-cell zygotes to the 2-cell stage (**Extended Data**
526 **Fig.9C**). Therefore, CDK8 activity is essential for the zygote to morula transition.
527
528 Next, we investigated the role of CDK8 post-morula. CDK8 mRNA
529 expression declines until blastocyst stage, both in mouse and human pre-
530 implantation embryos (**Extended Data Fig.9D-F**). CDK8 protein expression per
531 cell was homogenous in the mouse inner cell mass (ICM) at E3.5 (**Fig.8A,B**).
532 Interestingly, at E4.5, when the ICM segregates into the naïve epiblast and the
533 primitive endoderm (PE), CDK8 protein levels diverged, with lower levels in
534 epiblast compared to PE (**Fig.8A,B; Extended Data Fig.10A**). This pattern was
535 transient, and it reversed in post-implantation epiblast at E5.5 (**Fig.8A,B;**
536 **Extended Data Fig.10A**). To further document that CDK8 levels are reduced in
537 the naïve epiblast, embryos were cultured from E3.5-E4.5 with MEK inhibitor
538 (MEKi), which blocks PE formation, and permits only development of naïve
539 epiblast[73,74]. As expected, in MEKi, the ICM contained only naïve epiblast cells
540 and not PE, facilitating observation of their reduced CDK8 expression (**Fig.8C,D**).
541 Also, the CDK8 binding partner and essential activating subunit cyclin C altered
542 its nuclear-cytoplasmic ratio during this developmental window. Specifically, E4.5
543 epiblast contained significantly less nuclear cyclin C than E5.5 *in vivo* (**Extended**
544 **Data Fig.10B,C**), and a similar pattern was observed comparing 2i-naïve versus
545 primed state PSCs *in vitro* (**Extended Data Fig.10D**). In summary, the
546 emergence of naïve pluripotency during embryo development, at E4.5, coincides
547 with decreased CDK8 expression and decreased availability of its essential
548 subunit cyclin C. This parallels the effect of MEKi on CDK8 expression, and
549 stabilization of naïve epiblast identity in PSCs *in vitro* (**Extended Data Fig.5**).
550
551 We wondered whether inhibition of CDK8/19 affects the emergence of
552 naïve pluripotency. Similar to MEKi, CDK8/19i treatment during E3.5-E4.5 did not
553 interfere with embryo naïve epiblast development (**Fig.8E-G**) and allowed the
554 derivation of PSC lines. In contrast to MEKi, CDK8/19i permitted PE formation
555 (**Fig.8E-G**). This suggests that the critical roles of MEK for PE segregation are
556 independent of CDK8/19, and corresponds with our observation that MEK activity
557 is unaffected by CDK8/19i (**Fig.5F,G**).

Finally, we examined the importance of CDK8/19 activity during pre-implantation to post-implantation epiblast developmental progression. We focused on lumen formation within the epiblast, which marks the initiation of morphogenesis downstream of naïve pluripotency exit[35]. We found that CDK8/19i treatment during E4.5-E5.5 impaired embryo epiblast lumenogenesis (**Fig.8H; Fig.2E** for spheroids). This indicates a requirement for CDK8/19 activity to support epiblast development, from the naïve pre-implantation to primed post-implantation embryonic stages, consistent with the significant elevation we observed in CDK8 expression at this time (**Fig.8A,B; Extended Data Fig.9E**).

These data suggest that CDK8/19 expression in early embryonic development mirrors its function: the transition from zygote to morula, and the formation of the post-implantation epiblast require CDK8/19 activity; while the intervening naïve ICM has low CDK8 expression and reduced nuclear cyclin C (summary: **Extended Data Fig.10E**).

DISCUSSION

We uncover a role for the Mediator-kinases CDK8/19 in defining the equilibrium between naïve and primed pluripotent states, in both mouse and human pluripotent cells. Collectively, our data point towards the following model: 2i and CDK8/19i rapidly induce a highly overlapping set of phospho-changes focused on the transcriptional machinery, triggering enhancer hyperactivation, global increase in RNA Pol II recruitment to promoters, and resetting of gene expression. This includes the upregulation of enhancer-derived RNAs (eRNAs), and the resetting of endogenous retroviral and repeat element expression, to coordinate this cell identity conversion. Further evidence supporting transcriptional stabilization of naïve pluripotency includes repression of the 2C fluctuation in PSC identity, similar to 2i. Thus, the ability of 2i and CDK8/19i to induce naïve features appears to originate from their common effect on Mediator and RNA Pol II transcriptional activity. In support, super-enhancers interact with more target promoters[75], engage in more long-range interactions[75], and display

591 increased H3K27ac[76], in the naïve state versus primed. Our model agrees with
592 the concept that transitions in cell identity are driven by early reconfiguration of
593 the active enhancer network, which resets the transcriptional machinery to the
594 new program[70,71,77].
595
596 Evidence presented here suggests a signaling hierarchy; in particular,
597 MEK inhibition results in CDK8/19 inhibition, while inhibition of CDK8/19 does not
598 affect MEK activity. Accordingly, we observe that (i) the ability of MEK/GSK
599 inhibition (2i) to induce naïve features in PSCs *in vitro* is recapitulated by
600 CDK8/19i; (ii) 2i and CDK8/19i exhibit a 79.6% overlap in downstream phospho-
601 changes; (iii), CDK8/19 down-regulation coincides with the emergence of naïve
602 state *in vivo*, when MEK-ERK signaling is decreased[18,19]; and, (iv) both MEK-
603 ERK activation[18,19,30,73,74] and CDK8/19[78,79] drive post-implantation epiblast
604 differentiation, a process we found impaired by CDK8/19i, and a period where
605 CDK8 was up-regulated. Thus, we propose CDK8/19-inhibition is a common
606 downstream feature of naïve-inducing media cocktails. Further studies will
607 elucidate how MEK-ERK signaling regulates CDK8/19-Mediator activity in PSCs.
608 Intriguingly, Mediator hyperactivation via CDK8/19-inhibition triggers cancer cell
609 death[14], while we find a similar approach reinforces naive pluripotent identity.
610 Cancer cells commonly develop novel oncogenic SEs, becoming addicted to a
611 defined range of enhancer-driven transcription that appears sensitive to
612 perturbation[9,80]. This provides an interesting parallel with MEK inhibition, which is
613 also detrimental to many cancer cells, but beneficial to pluripotency.
614
615 Stabilization of the human naïve pluripotent state in vitro is challenging
616 and remains to be optimized[21,24]. Our understanding of stem cell identity
617 indicates a continuum of molecular changes along a spectrum from naïve to
618 primed states, which also reflects the developmental path in early
619 embryos[18,19,21,22]. Where does CDK8/19i position PSCs along this gradient? We
620 find that CDK8/19 inhibition recapitulates the majority of molecular characteristics
621 associated with primed to naïve transition. However, other molecular features
622 associated with the more-naïve end of this spectrum are not recapitulated by
623 CDK8/19 inhibition, particularly, global DNA hypo-methylation, X-chromosome

624 reactivation[66,67], and SSEA4 down-regulation[24,39]. Achieving these last features of
625 naïve pluripotency appears to come at a price. Naïve-inducing media cocktails
626 dependent on MEK-inhibition generate harmful side effects, specifically acute
627 chromosomal instability and imprinting erasure[24,39,62,63]. Other cocktails do not
628 downregulate SSEA4, produce modest DNA demethylation, and are not
629 associated with genomic instability[24,33,64]. Similarly, CDK8/19i installs many naïve
630 features in human cells while maintaining SSEA4, DNA global methylation and
631 genomic stability. CDK8/19i-treated cells retain normal karyotype after prolonged
632 culture. We suggest these important differences are due to CDK8/19i not
633 impinging directly on MEK-signaling.
634
635 In summary, CDK8/19i stimulates Mediator recruitment of RNA Pol II. This
636 effectively hyper-activates enhancers and stabilizes the transcriptional program
637 of naïve pluripotent cell identity. Thus, chemical inhibition of CDK8/19 may help
638 to solve remaining challenges in unstable human naïve PSC culture. Similarly,
639 these principles of stabilizing cellular identity may apply to other contexts of
640 cellular plasticity.

ACKNOWLEDGMENTS

We are grateful to Austin Smith, Todd MacFarlan, Zsuzsanna Izsvák, Minoru Ko, Jacob Hanna, Damien Grégoire and Ula Hibner for generous gifts of reagents. We thank assistance from Neus Prats and Luis César Fernández, as well as CNIO and IRB core facilities. M.N.S was funded by a Leverhulme Trust early career fellowship. Work in the lab of M.Z-G was funded by Wellcome Trust (098287/Z/12/Z) and ERC (669198). I.C was funded by Secretaria d'Universitats i Recerca de la Generalitat de Catalunya and European Social Fund. I.A. and P.S. were supported by Fondation pour la Recherche Medicale (DEQ20170336757), Infrastructure Nationale en Biologie et Santé INGESTEM (ANR-11-INBS-0009), IHU-B CESAME (ANR-10-IBHU-003), LabEx REVIVE (ANR-10-LABX-73), LabEx "DEVweCAN" (ANR-10-LABX-0061) and LabEx "CORTEX" (ANR-11-LABX-0042) of University of Lyon within the program 'Investissements d'Avenir' (ANR-11-IDEX-0007). Work in the laboroatory of N.M. was funded by the European Research Council (ERC), under the European Union Horizon 2020 research and innovation program (StG-2014-640525_REGMAMKID), the Spanish Association Against Cancer (AECC/LABAE16006), Carlos III Health Institute (Red TerCel, CardioCel, RD16/0011/0027), Ministry of Economy and Competitiveness (MINECO) projects SAF2017-89782-R, SAF2015-72617-EXP, and RYC-2014-16242, and the CERCA/Government of Catalonia (2017 SGR 1306). Work in the laboratory of S.O. was funded by SAF2013-44866-R from MINECO Spain. Work in the laboratory of M.F.F. was funded by Plan Nacional de I+D+I 2013-2016/FEDER (PI15/00892 to M.F.F. and A.F.F.); the ISCIII-Subdireccion General de Evaluación y Fomento de la Investigación, and Plan Nacional de I+D+I 2008-2011/FEDER (CP11/00131 to A.F.F.); IUOPA (M.I.S.); the Asturias Regional Government (GRUPIN14-052 to M.F.F.). The IUOPA is supported by the Obra Social Liberbank-Cajastur, Spain. Work in the laboratory of M.S. was funded by the CNIO, the IRB, and by grants from Spanish Ministry of Economy co-funded by European Regional Development Fund (ERDF) (SAF2013-48256-R), European Research Council (ERC-2014-AdG/669622), Botin Foundation, Banco Santander (Santander Universities Global Division), "laCaixa" Foundation, and,

673 Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement of
674 Catalonia (Grup de Recerca consolidat 2017 SGR 282).

675

676 **AUTHOR CONTRIBUTIONS**

677 C.J.L. designed and performed most experiments with mouse cells and embryos,
678 contributed to bioinformatic data analysis, and co-wrote the manuscript. R.B.
679 designed and performed most experiments with human cells, and provided
680 general experimental support. A.M.-V performed proteomic and bioinformatic
681 analysis. M.N.S. performed embryo experiments, immunofluorescence, and data
682 analysis; S.N.-P., I.C., L.R.-G., N.A. and M.M.-M. contributed to experimental
683 work and data analysis; C.T. and E.G. contributed to work with human PSCs and
684 performed differentiation, immunofluorescence and confocal analysis of these
685 experiments, supervised by N.M. O.G.-C., G.G.-L., and C. S-O. A., contributed to
686 bioinformatic analyses; C.B-A., S.M., and J.P. selected, synthesized, and
687 characterized small molecule inhibitors; S.O. provided reagents, contributed to
688 experimental design and supervised mouse embryo work; I.A. and P.S.
689 performed human-rabbit interspecies chimera and STAT3 assays; S.P., E.S.,
690 A.C., and D.F. generated the CDK8-knockout mouse, provided reagents, and
691 performed additional inhibitor analyses; A.F.F., M.I.S. and M.F.F. performed DNA
692 methylation analysis; P.S., D.F., J.M., and M.Z-G. provided reagents, discussion,
693 and revisions. M.S. designed and supervised the study, secured funding,
694 analyzed the data, and co-wrote the manuscript. All authors discussed the results
695 and commented on the manuscript.

696

697 **DECLARATION OF INTERESTS**

698 The authors declare no competing interests. The funders had no role in study
699 design, data collection and analysis, decision to publish, or preparation of the
700 manuscript.

REFERENCES

1. Heinz, S., Romanoski, C. E., Benner, C. & Glass, C. K. The selection and function of cell type-specific enhancers. *Nat. Rev. Mol. Cell Biol.* **16**, 144–154 (2015).
2. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
3. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
4. Allen, B. L. & Taatjes, D. J. The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 155–166 (2015).
5. Jeronimo, C. & Robert, F. The Mediator complex: at the nexus of RNA Polymerase II transcription. *Trends Cell Biol.* **27**, 765–783 (2017).
6. Soutourina, J. Transcription regulation by the Mediator complex. *Nat. Rev. Mol. Cell Biol.* **4**, 262–274 (2017).
7. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A Phase Separation Model for Transcriptional Control. *Cell* **169**, 13–23 (2017).
8. Fant, C. B. & Taatjes, D. J. Regulatory functions of the Mediator kinases CDK8 and CDK19. *Transcription* **2**, 76–90 (2019).
9. Chen, H. *et al.* A Pan-Cancer Analysis of Enhancer Expression in Nearly 9000 Patient Samples. *Cell* **173**, 386-399.e12 (2018).
10. Clark, A. D., Oldenbroek, M. & Boyer, T. G. Mediator kinase module and human tumorigenesis. *Crit. Rev. Biochem. Mol. Biol.* **50**, 393–426 (2015).
11. Poss, Z. C. *et al.* Identification of Mediator Kinase Substrates in Human Cells using Cortistatin A and Quantitative Phosphoproteomics. *Cell Rep.* **15**, 436–450 (2016).
12. Knuesel, M. T., Meyer, K. D., Bernecky, C. & Taatjes, D. J. The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. *Genes Dev.* **23**, 439–451 (2009).
13. van de Peppel, J. *et al.* Mediator Expression Profiling Epistasis Reveals a

Signal Transduction Pathway with Antagonistic Submodules and Highly Specific Downstream Targets. *Mol. Cell* **19**, 511–522 (2005).

14. Pelish, H. E. *et al.* Mediator kinase inhibition further activates super-enhancer-associated genes in AML. *Nature* **526**, 273–276 (2015).

15. Gonzalez, D. *et al.* Suppression of Mediator is regulated by Cdk8-dependent Grr1 turnover of the Med3 coactivator. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 2500–5 (2014).

16. Bancerek, J. *et al.* CDK8 Kinase Phosphorylates Transcription Factor STAT1 to Selectively Regulate the Interferon Response. *Immunity* **38**, 250–262 (2013).

17. Galbraith, M. D., Donner, A. J. & Espinosa, J. M. CDK8: a positive regulator of transcription. *Transcription* **1**, 4–12 (2010).

18. Hackett, J. A. & Surani, M. A. Regulatory principles of pluripotency: From the ground state up. *Cell Stem Cell* **15**, 416–430 (2014).

19. Nichols, J. & Smith, A. Naive and Primed Pluripotent States. *Cell Stem Cell* **4**, 487–492 (2009).

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

21. Weinberger, L., Ayyash, M., Novershtern, N. & Hanna, J. H. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* **17**, 155–169 (2016).

22. Wu, J. & Izpisua Belmonte, J. C. Dynamic Pluripotent Stem Cell States and Their Applications. *Cell Stem Cell* **17**, 509–525 (2015).

23. Marks, H. *et al.* The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**, 590–604 (2012).

24. Liu, X. *et al.* Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. *Nat. Methods* **11**, 1055–1062 (2017).

761 25. Ying, Q.-L. *et al.* The ground state of embryonic stem cell self-renewal.
762 *Nature* **453**, 519–523 (2008).

763 26. Bhagwat, A. S. *et al.* BET Bromodomain Inhibition Releases the Mediator
764 Complex from Select cis-Regulatory Elements. *Cell Rep.* **15**, 519–530
765 (2016).

766 27. Di Micco, R. *et al.* Control of embryonic stem cell identity by brd4-
767 dependent transcriptional elongation of super-enhancer-associated
768 pluripotency genes. *Cell Rep.* **9**, 234–247 (2014).

769 28. Finley, L. W. S. *et al.* Pluripotency transcription factors and Tet1/2 maintain
770 Brd4-independent stem cell identity. *Nat. Cell Biol.* **5**, 565-574 (2018).

771 29. Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline
772 development. *Nature* **450**, 1230–1234 (2007).

773 30. Ficz, G. *et al.* FGF signaling inhibition in ESCs drives rapid genome-wide
774 demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell*
775 **13**, 351–359 (2013).

776 31. Dale, T. *et al.* A selective chemical probe for exploring the role of CDK8
777 and CDK19 in human disease. *Nat. Chem. Biol.* **11**, 973–980 (2015).

778 32. Takashima, Y. *et al.* Resetting Transcription Factor Control Circuitry toward
779 Ground-State Pluripotency in Human. *Cell* **158**, 1254–1269 (2014).

780 33. Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent
781 stem cells. *Nature* **504**, 282–286 (2013).

782 34. Theunissen, T. W. *et al.* Systematic identification of culture conditions for
783 induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**,
784 471–487 (2014).

785 35. Shahbazi, M. N. *et al.* Pluripotent state transitions coordinate
786 morphogenesis in mouse and human embryos. *Nature* **552**, 239–243
787 (2017).

788 36. Chen, H. *et al.* Reinforcement of STAT3 activity reprogrammes human
789 embryonic stem cells to naive-like pluripotency. *Nat. Commun.* **6**, 7095
790 (2015).

37. Wang, J. *et al.* Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature* **516**, 405–409 (2014).

38. Stirparo, G. G. *et al.* Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. *Development* **145**, dev158501 (2018).

39. Pastor, W. A. *et al.* Naive Human Pluripotent Cells Feature a Methylation Landscape Devoid of Blastocyst or Germline Memory. *Cell Stem Cell* **18**, 323–329 (2016).

40. Wu, J. *et al.* Interspecies Chimerism with Mammalian Pluripotent Stem Cells. *Cell* **168**, 473-486.e15 (2017).

41. Wu, J. *et al.* Stem cells and interspecies chimaeras. *Nature* **540**, 51–59 (2016).

42. Gifford, W. D., Pfaff, S. L. & MacFarlan, T. S. Transposable elements as genetic regulatory substrates in early development. *Trends Cell Biol.* **23**, 218–226 (2013).

43. Göke, J. *et al.* Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells. *Cell Stem Cell* **16**, 135–141 (2015).

44. Grow, E. J. *et al.* Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature* **522**, 221–225 (2015).

45. Theunissen, T. W. *et al.* Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem Cell* **19**, 502–515 (2016).

46. Macfarlan, T. S. *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57-63 (2012).

47. Eckersley-Maslin, M. A. *et al.* MERVL/Zscan4 Network Activation Results in Transient Genome-wide DNA Demethylation of mESCs. *Cell Rep.* **17**, 179–192 (2016).

48. Kolodziejczyk, A. A. *et al.* Single Cell RNA-Sequencing of Pluripotent States Unlocks Modular Transcriptional Variation. *Cell Stem Cell* **17**, 471–485 (2015).

821 49. Fidalgo, M. *et al.* Zfp281 Coordinates Opposing Functions of Tet1 and Tet2
822 in Pluripotent States. *Cell Stem Cell* **19**, 355–369 (2016).

823 50. Buecker, C. *et al.* Reorganization of enhancer patterns in transition from
824 naive to primed pluripotency. *Cell Stem Cell* **14**, 838–853 (2014).

825 51. Bulut-Karslioglu, A. *et al.* Inhibition of mTOR induces a paused pluripotent
826 state. *Nature* **540**, 119–123 (2016).

827 52. Boroviak, T. *et al.* Lineage-Specific Profiling Delineates the Emergence and
828 Progression of Naive Pluripotency in Mammalian Embryogenesis. *Dev. Cell*
829 **35**, 366–382 (2015).

830 53. Blakeley, P. *et al.* Defining the three cell lineages of the human blastocyst
831 by single-cell RNA-seq. *Development* **142**, 3613–3613 (2015).

832 54. Nakamura, T. *et al.* A developmental coordinate of pluripotency among
833 mice, monkeys and humans. *Nature* **537**, 57–62 (2016).

834 55. Yan, L. *et al.* Single-cell RNA-Seq profiling of human preimplantation
835 embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1131–1139
836 (2013).

837 56. Xue, Z. *et al.* Genetic programs in human and mouse early embryos
838 revealed by single-cell RNA sequencing. *Nature* **500**, 593–597 (2013).

839 57. Petropoulos, S. *et al.* Single-Cell RNA-Seq Reveals Lineage and X
840 Chromosome Dynamics in Human Preimplantation Embryos. *Cell* **165**,
841 1012–1026 (2016).

842 58. Guo, G. *et al.* Naive Pluripotent Stem Cells Derived Directly from Isolated
843 Cells of the Human Inner Cell Mass. *Stem Cell Reports* **6**, 437–446 (2016).

844 59. Chan, Y. S. *et al.* Induction of a human pluripotent state with distinct
845 regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell*
846 **13**, 663–675 (2013).

847 60. Sahakyan, A. *et al.* Human Naive Pluripotent Stem Cells Model X
848 Chromosome Dampening and X Inactivation. *Cell Stem Cell* **20**, 87-101
849 (2017).

850 61. von Meyenn, F. *et al.* Impairment of DNA Methylation Maintenance Is the

Main Cause of Global Demethylation in Naive Embryonic Stem Cells. *Mol. Cell* **62**, 848–861 (2016).

62. Yagi, M. *et al.* Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation. *Nature* **548**, 224–227 (2017).

63. Choi, J. *et al.* Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. *Nature* **548**, 219–223 (2017).

64. Di Stefano, B. *et al.* Reduced MEK inhibition preserves genomic stability in naive human embryonic stem cells. *Nat. Methods* **15**, 732–740 (2018).

65. Leitch, H. G. *et al.* Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* **20**, 311–316 (2013).

66. Sahakyan, A. *et al.* Human Naive Pluripotent Stem Cells Model X Chromosome Dampening and X Inactivation Human Naive Pluripotent Stem Cells Model. *Stem Cell* 1–15 (2017). doi:10.1016/j.stem.2016.10.006

67. Khan, S. A., Audergon, P. N. C. B. & Payer, B. X-chromosome activity in naive human pluripotent stem cells—are we there yet? *Stem Cell Investig.* **4**, 54–54 (2017).

68. Rahl, P. B. *et al.* C-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).

69. Williams, L. H. *et al.* Pausing of RNA Polymerase II Regulates Mammalian Developmental Potential through Control of Signaling Networks. *Mol. Cell* **58**, 311–322 (2015).

70. Arner, E. *et al.* Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. *Science* **347**, 1010–1014 (2015).

71. Respuela, P. *et al.* Foxd3 Promotes Exit from Naive Pluripotency through Enhancer Decommissioning and Inhibits Germline Specification. *Cell Stem Cell* **18**, 118–133 (2016).

72. Westerling, T., Kuuluvainen, E. & Makela, T. P. Cdk8 Is Essential for Preimplantation Mouse Development. *Mol. Cell. Biol.* **27**, 6177–6182 (2007).

881 73. Yamanaka, Y., Lanner, F. & Rossant, J. FGF signal-dependent segregation
882 of primitive endoderm and epiblast in the mouse blastocyst. *Development*
883 **137**, 715–724 (2010).

884 74. Nichols, J., Silva, J., Roode, M. & Smith, A. Suppression of Erk signalling
885 promotes ground state pluripotency in the mouse embryo. *Development*
886 **136**, 3215–3222 (2009).

887 75. Lopes Novo, C. *et al.* Long-Range Enhancer Interactions Are Prevalent in
888 Mouse Embryonic Stem Cells and Are Reorganized upon Pluripotent State
889 Transition. *Cell Rep.* **22**, 2615–2627 (2018).

890 76. Di Stefano, B. *et al.* The RNA Helicase DDX6 Controls Cellular Plasticity by
891 Modulating P-Body Homeostasis. *Cell Stem Cell* **25**, 622-638.e13 (2019).

892 77. Factor, D. C. *et al.* Epigenomic comparison reveals activation of 'seed'
893 enhancers during transition from naive to primed pluripotency. *Cell Stem*
894 *Cell* **14**, 854–863 (2014).

895 78. Postlmayr, A., Dumeau, C. E. & Wutz, A. Cdk8 is required for
896 establishment of H3K27me3 and gene repression by Xist and mouse
897 development. *Development* (2020). doi:10.1242/dev.175141

898 79. Feldmann, A., Dimitrova, E., Kenney, A., Lastuvkova, A. & Klose, R. J.
899 CDK-Mediator and FBXL19 prime developmental genes for activation by
900 promoting atypical regulatory interactions. *Nucleic Acids Res.* **48**, 2942–
901 2955 (2020).

902 80. Bradner, J. E., Hnisz, D. & Young, R. A. Transcriptional Addiction in
903 Cancer. *Cell* **168**, 629–643 (2017).

904

905

FIGURE LEGENDS

Figure 1:

An inhibitor screen for promotion of PSC naïve state identifies a distinct role for Mediator kinase activity.

(**A**) FACS: effect of indicated treatments on Nanog expression (*Nanog*-GFP^{high}) per cell, using mouse *Nanog*-GFP knockin reporter PSC cell line[29] in standard serum/LIF base media. Mean+/-SD of n=4 independent experiments.

(**B**) FACS histogram of *Nanog*-GFP expression. *Nanog*-GFP^{Low} and *Nanog*-GFP^{High} cell populations in serum/LIF population (grey). Dotted line: threshold where >95% cells are *Nanog*-GFP^{High} in 2i-naïve culture condition. Data representative of 3 experiments.

(**C**) PSC colony morphology in indicated treatments. Brightfield and *Nanog*-GFP shown. Data representative 6 experiments.

(**D,E**) Endogenous *Nanog* mRNA (**D**), or protein (**E**), expression levels in mouse PSCs adapted to indicated conditions. Data representative of n=3 experiments, Mean+/-SD, *P < 0,05, t-test, unpaired, two-tailed.

(**F**) Phosphorylation levels of CDK8-target STAT1 Ser727. HERVH human iPS PSCs treated with CDK8/19i concentrations for 3h +/- simultaneous induction of STAT1-Ser727P by γ-interferon for 3h. Data representative of 2 experiments.

(**G**) Cell morphology and qRT-PCR of mouse PSCs following 7d of shRNA-mediated knockdown of CDK8, CDK19, CyclinC (CCNC). Data: Mean of 2 experiments.

(**H,I**) Cell morphology and alkaline phosphatase staining (**H**), FACS analysis of endogenous NANOG and OCT4 protein levels (**I**), in CDK8/19-dKO iPSCs stably expressing pMSCV-Empty or pMSCV-CDK8-Kinase Dead (CDK8-KD). Data representative of 3 independent clones.

(**J,K**) qRT-PCR (**J**; Data: Mean+/-SD from n=3 independent clones), and Western blot analysis of protein expression (K; Data: representative of 2

experiments) in WT iPSCs, or CDK8/19-dKO iPSCs stably expressing pMSCV-Empty or pMSCV-CDK8-KD, adapted to the indicated media conditions.

(**L**) Alkaline phosphatase staining. Cells fixed and stained at 14d after retroviral expression of pMSCV-Empty or pMSCV-CDK8-KD. Staining intensity was scored visually for each colony, using ten fields of view. Data represent n=3 experiments, Mean+/-SD.

(**M**) Immunofluorescence in CDK8/19-dKO iPSCs expressing pMSCV-CDK8-KD-puro-IRES-GFP. Data representative of 4 experiments.

Scale bars 100μm. See also **Supplementary Table 1** and **Supplementary Information** file for inhibitor characterization.

Figure 2:

Positive effect of long-term CDK8/19i on PSC self-renewal and pluripotency

(**A**) Morphology and mRNA expression of mouse EpiSC expressing pMSCV-Empty or pMSCV-CDK8-KD, then 7d in EpiSC media, or standard ES media serum/LIF. Plots: Mean+/-SD, n=3 experiments. Scale bar 100μm.

(**B,C**) Clonogenicity of mouse PSCs. *Nanog*-GFP PSCs were FACS-sorted 1-cell/well, cultured 7d, then Alkaline phosphatase stained (**B**), or scored for *Nanog*-GFP intensity (**C**), to assess pluripotent status/colony, in either standard media serum/LIF, 2i-naïve, or CDK8/19i conditions. Data: representative of 3 experiments.

(**D,E**) Differentiation capacity of mouse PSCs previously adapted to serum/LIF, 2i or CDK8/19i. (**D**), PSCs differentiated as indicated in 2D culture. PSC exit from pluripotency (*Nanog* down-regulation) and differentiation (*Nestin* upregulation) by qRT-PCR. Data: Mean of 2 experiments. (**E**) Pluripotency exit assessed by immunofluorescence after PSC culture in 3D-matrigel/LIF +/-CDK8/19i, to observe early epiblast development (rosette formation and lumenogenesis) in PSC-spheroids[35]. "Disorganized": differentiatiation failure. Data: representative of 3 experiments, n=30 spheroids/condition; T-test, unpaired, two-tailed, *P=0,0097. Scale bar 10μm.

(**F-I**) In vivo assays of developmental capacity. Mouse CDK8/19i-PSCs constitutively-labelled (ROSA26-GFP or Tg.CAG-Katushka) were aggregated with, or micro-injected into, host E2.5 morulae. Embryo chimerism was assessed visually: E4.5 blastocyst, n=10 (**F**); E6.5 egg cylinder, n=10 (**G**); E14.5, n=2 (**H**); and peri-natal E19.5, n=4 (**I**). In (**I**), 3 male adult chimeras (below left, percent chimerism by coat colour, indicated) displayed germline transmission, generating 3 litters (below right, coat colour confirmed germline transmission per litter). Scale bars, **F**: 25μm; **G**: 100μm; **H**: 1mm; **I**: 1mm.

(**J**) Induction of naïve colony morphology in human OSCAR ES cells. Tamoxifen-inducible consitutively-active STAT3/LIF/2i (TL2i)[36], or substituting CDK8/19i for 2i (TLCDK8/19i). Data: representative of 3 experiments. Scale bar 100μm.

(**K,L**) (**K**) Induction of naïve colony morphology in 3 human PSC lines; primed or +14d CDK8/19i. Scale bar 100μm. (**L**) Cytometry of *HERVH*-GFP intensity/cell in human PSCs; primed or 14d CDK8/19i culture. Panels **K** and **L**, representative of >5 experiments.

(**M**) Western blots of pluripotency markers in human PSCs; primed or cultured 14d with 2i-based or CDK8/19i-based media. SMC1, loading control.

Figure 3:

Positive effect of long-term CDK8/19i on human PSC self-renewal and pluripotency

(**A**) Immunofluorescence. Pluripotency markers; human PSCs cultured as indicated. Data representative of 2 experiments. Scale bar 100μm.

(**B**) Embryoid body differentiation assay using endoderm-directed, or caridac-directed protocols with human PSCs. Plots: mRNA expression of pluripotency or embryonic germ layer markers by qRT-PCR, Mean of n=4 technical replicates. Data for H9 cell line shown, representative of two experiments (human PSC lines H1 and H9: **Source Data** and **Supplementary Information** summarize all tested lineage markers, determined by qRT-PCR (17) or immunofluorecence (6).

(**C,D**) Human PSCs adapted to primed or CDK8/19i, then tested by teratoma differentiation assay. Data representative of n=3 three experiments (human PSC lines: H1, D2#2 and HERVH). (**C**) Immunofluorescence shows markers for 3 embryonic germ layers in H1 and D2#2 cell lines (indicated). Scale bar 50μm. (**D**) Summary of all tested lineage markers (6) determined by immunofluorecence in teratomas generated from the three human PSC lines in (C). "+" = detected; "0" = not detected; "n/d" = not determined. **Source Data Figure 3** summarizes all differentiation markers tested for all three cell lines in (**C,D**).

(**E**) Interspecies chimera assay in vivo to assess developmental capacity of human PSCs adapted to primed or CDK8/19i conditions. Constitutively-labelled human PSCs (td:tomato-red; HERVH iPS cell line) were introduced into host rabbit morulae of ~E2.5. Chimerism was assessed visually +72hrs later in ~E5.5 rabbit blastocyst. Below plot: indicated are the number of human cells introduced (5 or 10), and number of embryos (n) in each of the three experiments (Exp1-3). The bar chart quantifies the number of human cells observed in rabbit embryos; data from 3 independent experiments. A representative image shows immunofluorescence of E.5.5 rabbit blastocyst, with inner cell mass indicated (determined by Nanog staining, inset panel, below). Human PSCs adapted to CDK8/19i displayed moderate contribution to human-rabbit chimeras. Scale bar 20μm.

Figure 4:

Gene expression in mouse and human PSCs adapted to 2i or CDK8/19i

(**A,B**) Overlap and hypergeometric significance (**A**) of differentially expressed mRNAs in mouse PSCs in 2i-naïve or CDK8/19i, versus serum/LIF (RNA-seq; n=3 biological replicates; FDR<0.01). (**B**) Heatmap: changes in selected pluripotency regulators.

(**C,D**) Overlap and hypergeometric significance of differentially expressed ERV families (**C**), and overlap of 2-Cell (2C) fluctuation markers (**D**), in mouse PSCs in 2i-naïve or CDK8/19i, versus serum/LIF (n=3 biological replicates; FDR<0.05).

(**E**) Heatmap: normalized enrichment scores (NES) from GSEA comparison of mouse RNA-seq data in (**A**) (n=3 biological replicates), or five other studies[23,48–51], versus stage-specific marker genesets identified during mouse preimplantation development[52]. Significant comparisons: P<0.05 and FDR q<0.05, marked (*).

(**F,G**) Overlap and hypergeometric significance (**F**) of differentially expressed mRNAs in human PSCs cultured as indicated versus control primed cells (RNA-seq; n=3 biological replicates, FDR<0.05, threshold >2x fold-change). (**G**) Selected human naïve pluripotency markers (up), or post-implantation primed epiblast markers (down) in human PSCs cultured as indicated.

(**H,I**) Heatmaps: correlation in RNA expression changes for ERV families (**H**; n=1066 families, each row = Mean of family), or individual loci of selected ERV families (**I**; genomic loci n=indicated) in human PSCs cultured as indicated versus control primed cells.

(**J**) qRT-PCR: RNA expression in human PSCs cultured as indicated, as in (**F**) (n=3 biological replicates, Mean+/-SD).

(**K**) Heatmap: NES scores from GSEA comparison of our human PSC RNA-seq (n=3 biological replicates), or seven other studies (above heatmap)[32–34,39,58–60], verses stage-specific marker genesets identified during human preimplantation development (below each heatmap)[54–57]. Lower panel: comparison, our data

versus other studies. Significant comparisons: P<0.05 and FDR q<0.05, marked (*).

(**L,M**) Proteomic overview (**L**). Overlap and hypergeometric significance of differentially expressed proteins averaged across five mouse PSC lines in 2i-naïve or CDK8/19i, versus standard serum/LIF culture (n=5 biological replicates, FDR<0.05). Heatmap (**M**): protein changes in selected pluripotency regulators per cell line. Data per cell line: **Extended Data Fig.S5A,B**.

Supplementary Table 2 and **Source Data:** full gene lists, ERV lists, fold changes, statistical tests in (**A-I, K**). **Supplementary Table 4** list of differentially expressed proteins in (**L,M**).

Figure 5:

CDK8/19i regulates the phospho-proteome and RNA Pol II similar to 2i-naïve pluripotency, but not DNA methylation.

(**A**) Global DNA methylation (5-methyl-cytosine) changes in n=4 mouse PSC lines in 2i or CDK8/19i. Left panel, Mean of n=2 technical replicates for each cell line. Right panel, Mean+/-SD of n=4 cell lines, t-test, unpaired, two-tailed, P-values indicated.

(**B**) Global DNA methylation (5-methyl-cytosine) changes in two human PSC lines. Mean+/-SD of n=2-6 biological replicates per cell line, as indicated.

(**C**) Pyrosequencing for CpG methylation status of three specific loci in LINE L1 family of repeat elements in mouse PSC in treatments indicated in n=1 mouse PSC line (E14).

(**D,E**) Phosphoproteomic overview. Mouse ES lines in 2i or CDK8/19i for 15min. (**D**) Pie-charts summarize total and differential phosphosites detected. (**E**) Heatmap and Gene Ontology (GO) summary for proteins displaying significantly differential phosphorylation. (**D,E**) n=2 biological replicates, FDR<0.05. **Supplementary Table 5**: full protein lists.

(**F,G**) Western blots: protein phosphorylation during short-term exposure of PSCs to 2i or CDK8/19i, as indicated. (**F**) Left: ERK1/2 phosphorylation in mouse PSCs. Right: STAT1 Ser727 phosphorylation in human iPSCs. (**G**) Left: ERK1/2 phosphorylation in mouse ES cells. Right: plot shows relative ERK1/2 phospho-levels, normalized by total ERK1/2 levels. (**F,G**) Data representative of 2 biological experiements.

(**H,I**) ChIP-seq: RNA Pol II abundance at all Refseq Transcription Start Sites (TSS; n=28,441), in mouse PSCs, treated as indicated. (H) Heatmap, TSS +/-5Kb. (I) Metagene average +/-2Kb. ChIP-seq n=3 pooled replicates.

(**J**) Cumulative plots of RNA Pol II loading ratios in mouse PSCs adapted to culture conditions as in (**H,I**). RNA Pol II abundance was calculated in the indicated regions (schematic) for each gene: promoter, gene body, and transcription termination zone (**Methods**). The ratios of RNA Pol II abundance

between these regions were calculated for each gene, and displayed cumulatively in the three plots (n=12,072 genes). For Promoter Loading Index (PLI; on left), dotted line indicates 90% of genes have PLI > 2.0. **Supplementary Table 6**, and **Source Data**: full list of ChIP-seq data, genes, and abundance ratios.

Figure 6:

2i and CDK8/19i exert similar effects on Mediator, RNA Pol II loading, and enhancer activity.

(**A**) Change in Total- or Ser5P-RNA Pol II abundance at TSS by ChIP-seq in mouse PSCs in 2i or CDK8/19i versus serum/LIF control. Data: Mean+/-SD; Total Pol II, n=12,693TSS; Ser5P, n=4,470TSS.

(**B**) Overlap and hypergeometric significance (P-value indicated) of genes where RNA Pol II abundance at TSS increased >1.5xfold in 2i and CDK8/19i (as in **A**); TSS n=12,072.

(**C,D**) Fold-change in RNA Pol II abundance at TSS (**C**), and fold-change in RNA Pol II Promoter Loading Index (PLI) (**D**), on genes (n=12,693) where RNA Pol II was detected in mouse PSCs in 2i or CDK8/19i versus control serum/LIF.

(**E,F**) Genes with largest changes in ChIP-seq RNA Pol II abundance correlate in CDK8/19i or 2i-naïve conditions. (**E**), individual promoter-TSS with RNA Pol II loading altered in CDK8/19i by greatest (upper panel) or least amount (middle panel) versus serum/LIF control (top 200 TSS in each case), are compared against a ranked list (bottom panel) of differential Pol II loading on all TSS for 2i-naïve versus serum/LIF control. (**F**) Reverse comparison: top 200 altered TSS in 2i-naïve versus ranked list of Pol II changes in CDK8/19i.

(**G**) CDK8/19 abundance in mouse PSCs, defined by ChIP-seq[2,3], peak-calling, and grouped by genomic localization. SE/TE: super/typical enhancers. Promoter-TSS: TSS+/-1Kb. Gene Body: Exons, Introns, Transcription Termination Site TTS+/-1Kb. T-test, unpaired, two-tailed, ****P<0.001.

(**H**) RNA Pol II abundance in mouse PSC genomic regions without CDK8/19 binding (left, n=423), or with top 10% strongest CDK8/19 binding signals (right, n=464), as defined: **Extended Data Fig.8D**. T-test, unpaired, two-tailed, Welchs correction, ****P<0.0001.

(**I**) RNA Pol II abundance in mouse PSCs, at typical enhancers (TE; left, n=9981), or super-enhancers (SE; right, n=646), as defined[2,3]. T-test, unpaired, two-tailed, Welchs correction, ****P<0.0001.

(**A-F,H,I**): RNA Pol II abundance measured by ChIP-seq (3 pooled replicates). (**G,H,I**): Tukey box plot centre lines show median values, box limits represent upper and lower quartiles, whiskers show 1.5× interquartile range. Number of n regions per group, P-value calculations, list of defined loci, and HOMER functional annotations: **Supplementary Tables 3,7;** and **Source Data**.

Figure 7:

2i and CDK8/19i hyperactivate naïve-state enhancer activity, conferring resistance to enhancer destabilization.

(**A**) Left, qRT-PCR: pluripotency genes and naïve-specific enhancer RNA (eRNA)[71] abundance in mouse PSC at time intervals after exposure to 2i or CDK8/19i, relative to standard serum/LIF control. Data: Mean+/-SEM, n=3 independent experiments. Right: heatmap summarizing mean fold-change in expression of the replicate experiments determined by qRT-PCR in left panel, and **Extended Data Fig.8G** during inhibitor withdrawal.

(**B,C**) Selective up-regulation of SE-target genes. SE: super-enhancer. (**B**) Left: GSEA of SE-target gene mRNAs in mouse PSCs adapted to CDK8/19i (n=3 biological replicates). SE-targets were defined as the single nearest gene by GREAT analysis (**Methods**) and are significantly up-regulated (on left; 408 genes, FDR q-value<0.001). Right: expression-matched control genes (not predicted to be enhancer targets, but which are highly expressed similar to SE-target genes) show no significant change in levels (right panel; 464 genes, FDR q-value=1.0). (**C**) The relative specificity of SE-target gene up-regulation by CDK8/19i was determined by comparison with databases of many other genesets using GSEA. C5 Gene Ontology (GO) terms (n=3,844 genesets) and C2 KEGG (n=150 genesets) as defined by the Broad Institute (**Methods**). SE-target genes lie within the top 1% most-significantly up-regulated genesets realtive to these GO term or KEGG databases. (**B,C**) SE-target and expression-matched SE-non-target genes were defined as the single nearest gene by GREAT analysis (**Methods**). **Source Data** for SE-target gene list (408 genes) and expression-matched control geneset (464 genes).

(**D**) Brightfield images: typical cell morphology following treatment with 500nM BRD4i/JQ1 for 48h (left) or 7d (right), in WT iPS cells or in CDK8/19-double-knockout (CDK8/19-dKO) iPS cells stably expressing pMSCV-CDK8-Kinase Dead (CDK8-KD). Panels on right show brightfield image of colonies which have been fixed and stained for alkaline phosphatase at Day 21/passage 5 of treatment with 500nM BRD4i/JQ1. Data are representative of 3 independent cell experiments.

(**E**) qRT-PCR expression of naïve-specific eRNA[71] and marker genes following treatment with 500nM BRD4i/JQ1 for 48h. CDK8/19-dKO iPS +/- CDK8-KD were cultured +/- 2i or standard serum/LIF, as indicated. Mean+/-SD, n=3 clones.

Figure 8:

CDK8 expression in vivo and role of Mediator during mouse preimplantation development

(**A,B**) Immunofluorescence and quantification of CDK8 expression at indicated timepoints during early mouse embryo development. (**A**) Single Z-section shown. (**B**) CDK8 protein levels per nucleus quantified per timepoint relative to internal controls. OCT4/GATA6 co-expression marks all ICM cells at E3.5. OCT4/GATA6 segregation from E4.5 to E5.5, marks epiblast (Epi/OCT4+) and primitive endoderm (PE/GATA6+), which later forms visceral endoderm (VE) at E5.5. Embryo staging, CDK8 quantification and normalization (**Methods**). Data: 2 independent experiments (images representative), ICM, 6 embryos, n=64 nuclei; E4.5, 5 embryos, nuclei: n=51(Epi), 48(PE); E5.0, 5 embryos, nuclei: n=48(Epi), n=52(PE); E5.5, 6 embryos, nuclei: n=100(Epi), n=84(PE/VE) (see: **Source Data**). Plot: Mean+/-SD; significance assessed by Kruskal-Wallis test with Dunn's multiple comparisons.

(**C,D**) CDK8 expression is repressed by MEK-inhibition *in vivo*. Embryos were incubated 48h +/- MEKi (E3.5 morula to E4.5 blastocyst). (**C**) Immunofluorescence: CDK8 protein expression in E4.5 blastocyst. Data representative of 3 experiments. (**D**) CDK8 protein levels per cell quantified in ICM, or Trophectoderm, relative to internal controls, per z-slice/image. Left: Control, n=7 embryos, 55 images; MEKi, n=12 embryos, 44 images. Right: Control, n=22 embryos, 22 images; MEKi, n=27 embryos, 27 images. Tukey box plot centre lines show median values, box limits represent upper and lower quartiles, whiskers show 1.5× interquartile range; significance assessed by T-test, unpaired, two-tailed, ***P<0.001; ns = not significant.

(**E-G**) CDK8/19-inhibition does not prevent Epi/PE segregation. E3.5 embryos were incubated 24h +/- CDK8/19i, during Epi/PE segregation, then assessed by immunofluorescence (images representative; **E,G**). (**F**) bar charts: quantification of ICM cell number, and lineage allocation in ICM, defined as: Epi/NANOG+; PE/GATA6+; ICM: Nanog+ or Gata6+). Data: Mean+/-SEM; 2 experiments; significance by T-test, unpaired, two-tailed; ns not-significant; number of embryos: control: n=7; CDK8/19i: n=9.

(**H**) CDK8/19-inhibition interrupts pre- to post-implantation morphogenic events. Pre-implantation E4.5 embryos were cultured until E5.0 *in vitro*[35] +/- CDK8/19i. PODXL/F-ACTIN staining determines emergence of epiblast pro-amniotic cavity/lumen (outlined in representative images). Right: Morphogenesis quantified; significance by Chi-square test, *P<0.05. Data: 2 experiments; number of embryos: Control, n=15; CDK8/19i, n=16.

(**A-H**) Scale bars 20μm.

METHODS

For primers, antibodies, shRNAs, and CRISPR-Cas9 gRNAs: **Supplementary Table 8**

PLURIPOTENCY AND DIFFERENTIATION ASSAYS

Mouse and human work

Animal experimentation at the Spanish National Cancer Research Centre CNIO (Centro Nacional de Investigaciones Oncológicas) performed according to protocols approved by CNIO-ISCIII Ethics Committee for Research and Animal Welfare (CEIyBA). Animal experimentation at University of Cambridge was approved by the Home Office, performed according to Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012, and reviewed by University of Cambridge Animal Welfare and Ethical Review Body (AWERB). Cdk8-flox/flox RERT-Cre mice were generated by Daniel Fisher (IGMM, Montpellier).

Human pluripotent stem cell studies were ethically approved in CNIO, Madrid, by the Comisión de Garantías para la Donación y Utilización de Células y Tejidos Humanos, and signed by Director of Instituto de Salud Carlos III (Nuevas fronteras en la Reprogramación Celular: Explotando la plasticidad cellular; Ref: 303). Studies at the IRB Barcelona, were approved by Ethics Committee of CMRB, and by Comisión de Seguimiento y Control de la Donación de Células y Tejidos Humanos del Instituto de Salud Carlos III and Ministry of Health from Government of Catalonia (project numbers: 0336S/11730/2015; 0336S/11220/2016; 0336S/2473/2017; 0336/747/2018).

Mouse cells and culture conditions

Mouse Pluripotent Stem Cells (PSCs): E14Tg2a.4 (wild-type parental, 129/Ola background) from BayGenomics/MMRRC resource, University of California; Wild-type PSCs were derived at the Transgenic Mouse Unit of CNIO from E3.5 C57BL6 blastocysts, or mixed background C57BL6/129 blastocysts; Rosa26-GFP and Tg.CAG-Katushka-red PSC lines were derived from 129-*Gt(ROSA)26Sor*$^{tm1(CAG-EGFP)Luo}$/J mice (Jackson 006053) and from Tg.CAG-Katushka mice[81], respectively. *Nanog*-GFP knock-in mouse PSCs (TNGA, TON) as described[29] and shared by laboratory of Austin Smith. MERVL-td:Tomato mouse 2C-reporter PSCs were shared by laboratory of Todd Macfarlan[46]. ZS mouse 2C-reporter PSCs were shared by laboratory of Minoru Ko[82]. Mouse PSCs (ES cells and iPS cells), were routinely cultured on 0.1% gelatin-coated plates in a base media of either "Serum/LIF" (15% FBS), or Knockout Serum Replacement (KSR, Invitrogen) "KSR/LIF" (15% KSR), in DMEM (high glucose) basal media, with LIF (1000 Units/mL), non-essential amino acids, glutamax and β-mercaptoethanol plus antibiotics. Where used with mouse PSC, the "2i" two-inhibitor cocktail comprised 1μM MEK-inhibitor (PD0325901, Axon Medchem, #1408) plus 3μM GSK3β-inhibitor (CHIR 99021, Axon Medchem #1386) as described[25]. Cultures were routinely tested for mycoplasma. Primary mouse embryo fibroblasts (wild-type, MEFs, passage 2) were obtained at E13.5 from pure inbred C57BL6 background mice,

or from CDK8-flox/flox RERT-Cre mice. Human 293T cells were from ATCC. All above-mentioned cells were maintained in: DMEM, 10% FBS (Gibco), with antibiotics (penicillin/streptomycin 100U/ml). Reprogrammed iPS cells were initially derived and expanded on mitomycin-C inactivated feeder cells on gelatin-coated plates, before transfer to gelatin-only.

Human PSC cell resources

HERVH iPS PSCs were shared by laboratory of Zsussanna Izsvak (Max Delbruck Centre for Molecular Medicine)[37]. WIBR3 ES PSCs were shared by laboratory of Jacob Hanna (Weizmann Institute of Science). OSCAR ES PSCs carrying inducible STAT3 were shared by laboratory of Pierre Savatier (SBRI, Stem Cell and Brian Research Institute)[36]. H1 and H9 human ES PSCs, and CB5, D2#2, and D2#4 human iPS PSCs, were shared by laboratory of Nuria Montserrat (IBEC, Institute for Bioengineering).

Human PSC cell culture in primed state

Human PSC (H1, H9, WIBR3, HERVH, CB5, D2#2, D2#4, OSCAR) were maintained in conventional primed conditions as described[33,36,37], specifically, by culture on growth factor-reduced phenol red-free matrigel (BD Biosciences #356231) with mTeSR1 media (Stem Cell Technologies). Cultures were passage every 5-7 days manually using either 2mg/ml dispase (Gibco), 0,5µM EDTA/1xPBS, or accutase (Gibco).

Resetting human PSC from primed to naïve state using 2i-based media cocktail

The naïve human pluripotent state was obtained by two methods. OSCAR PSCs were reset to the naïve state with 2i (TL2i) or CDK8/19i (1,1µM or 0,4µM) plus rhLIF and STAT3 transgene induction, as described[36]. In a transgene-free approach, human PSCs were cultured in a 2i-based chemical cocktail[33] referred to here as "2i p38iJNKi". Cells were maintained on matrigel (BD Biosciences #356231) using mTeSR1 (Stem Cell Technologies). Media was supplemented with 20ng/ml of recombinant human LIF (Peprotech, as described[33], 1µM PD0325901 (MEKi, Axon Medchem), 1.5µM CHIR 99021 (GSK3i, Axon Medchem), 10µM SP600125 (JNKi, TOCRIS) plus 2µM BIRB796 (p38i, Axon Medchem). In order to obtain and maintain the naïve state using the 2i p38iJNKi media cocktail, cells were selected at each passage, by cytometry sorting for the top 10% HERVH-GFP levels, or by repeated manual picking for selection of colonies with dome-shaped morphology. Conversion of human PSC from primed to naïve required 3 passages/rounds of selection, over 14-18 days.

Resetting human PSC from primed to naïve pluripotent state using CDK8/19i

To adapt and maintain human PSC to CDK8/19i culture (CDK8/19i-adapted), cells were maintained on matrigel (BD Biosciences #356231) using mTeSR1 (Stem Cell Technologies). Media was supplemented with 20ng/ml of recombinant human LIF (Peprotech), as described[33], plus 0,4µM or 1.1µM of CNIO-CDK8/19 inhibitor, or plus 10µM of SenexinA-CDK8/19 inhibitor[83]. This adaption process was also successful with 10µM SP600125 (JNKi, TOCRIS) plus 2µM BIRB796 (p38i, Axon Medchem),

but these additives are not required with CDK8/19i. Following background cell death in the first passage, colonies gradually become dome-shaped within 10-14 days without additional selection, and could be expanded by passage, using 3–5min treatment with Accutase (Gibco) or 0,5μM EDTA/1xPBS to avoid confluency, usually every 5-7 days due to a slowdown in proliferation. The optimal CDK8/19i concentration was 1.1μM for HERVH-GFP human iPSC, while it was 0,4μM for all other human PSC lines.

Mouse ESC derivation

ES cell line derivation was by standard methods. 8-cell stage mouse embryos from oviducts of pregnant females were cultured in serum/LIF on mitomycin C-inactivated MEF feeders plus 2i or CDK8/19i (added fresh every 2 days) until emergence of colonies from hatched blastocysts. Feeders were not compatible with several days of CDK8/19i, therefore cells were passed every 2 days to fresh feeders, then transferred to 0,1% gelatin-only**.**

Mouse EpiSC derivation

PSCs in 2i/LIF cultured on gelatin were firstly induced to differentiate into Epiblast-Like Cells (EpiLCs) over 48h by seeding on fibronectin-coated plates (10ng/ml) and switching to media containing 1%KSR, N2B27, FGF2 (12ng/ml) and Activin A (20ng/ml)[50]. After 48h, the cells were in a flat EpiLC state, and media was switched to include 20% KSR, and expanded for 5 passages to stabilize the EpiSC primed state, confirmed by typical flat colony morphology and Fgf5 expression. EpiSC colonies were passaged as clumps.

Analysis of PSC self-renewal

Mouse or human PSC self-renewal and pluripotency was scored by colony morphology, cytometry (mouse: *Nanog*-GFP heterogeneity and overall intensity, and co-staining for ICAM1; human: HERVH-GFP intensity, and assessing expression of NANOG, OCT4, SSEA4, TRA1-81), by alkaline phosphatase staining (fixed cells; Promega #S3771), and by immunofluorescence and qRT-PCR (for pluripotency markers, indicated in each Figure). Alkaline phosphatase staining intensity was quantified by scoring colonies observed by brightfield microscopy in 10 random fields of view per well.

Mouse PSC differentiation with LIF-removal and retinoic acid

LIF was first removed for 24h by culture in LIF-free Differentiation medium (as for serum/LIF media except the LIF is omitted). Next, Retinoic Acid was added at 10μM from +24 to +72h, followed by LIF-free differentiation medium alone from +72 to +96h. Differentiation was also assessed by the same protocol of LIF-withdrawal except without adding Retinoic Acid.

Mouse PSC differentiation by hanging-drop and embryoid bodies

PSCs were transferred to LIF-free Differentiation medium (as above) and suspended in hanging drop culture at 1000-5000 cells/20μLs for 48h to form Embryoid Bodies (EBs),

followed by transfer to suspension culture in low-adherence petri-dishes. Fresh medium was added every 3d, and development of beating cells in cardiac centres was scored daily.

Mouse chimera assays : morula aggregation and blastocyst micro-injection

Following 10 passages in serum/LIF, 2i, or CDK8/19i, mouse PSCs labelled constitutively with Rosa26-GFP or Tg.CAG-Katushka[81] underwent morula aggregation at E2.5, or blastocyst micro-injection at E3.5, as described[25]. The extent of GFP+ or Katushka-red+ cell chimeric contribution was assessed by confocal fluorescence at E4.5, or embryos were introduced into CD1 pseudo-pregnant females for implantation, and harvested at post-implantation time-points: E6.5, E14.5, or E19.5. Chimeras developed to adulthood were assessed by coat colour contribution and capacity for germline transmission.

Cardiac and endoderm directed differentiation of EBs derived from hPSCs.

Human PSC colonies were dissociated and cultured in suspension for 3d to form EBs in DMEM/F12, 15% FBS, 2 mM L-glutamine, non-essential amino acids, and penicillin/streptomycin. To generate endoderm, EBs were transferred to 0.1% gelatin-coated plates for 2 weeks in differentiation medium (DMEM, 20% FBS, 2mM L-glutamine, 0.1mM 2-mercaptoethanol, non-essential amino acids, and penicillin-streptomycin). To generate cardiac tissue, differentiation medium was supplemented with 100μM ascorbic acid (Sigma). In all conditions, EBs spontaneously gave rise to neural cell clusters.

Teratoma assays

Mouse PSCs: 10^6 cells in 100μl were injected sub-cutaneously in nude mice. Human PSCs: $2x10^6$ cells in 30μl were injected into testis of male SCID beige mice.

Mouse embryo manipulation and analysis

Mouse embryo collection, culture for pre-implantation embryo development in vitro, and fixation for immunofluorescence, as described[74]. Pre- to post-implantation embryo development in vitro, immunofluorescence analysis of CDK8 and Cyclin C levels in pre-implantation mouse embryos, and lumenogenesis by mouse PSC embryoid formation in Matrigel, as described[35]; further details on request.

Viral production and iPS (induced pluripotent stem) cell reprogramming

As described[84]. Briefly, retroviral and lentiviral supernatants were produced in HEK-293T cells. Filtered supernatants were collected after 48h, and added to recipient cells in 4 infections. Retroviral supernatants delivered exogenous CDK8 expression constructs and iPS reprogramming vectors. Lentivirus supernatants delivered shRNA knockdown vectors and CRISPR-Cas9 vectors. Plasmids: **Supplementary Table 8**.

Interspecies chimera developmental potency

Primed human iPS PSCs were pre-cultured with ROCK-inhibitor 24h, prepared as a unicellular suspension, and electroporated (Neon Transfection System; Invitrogen; 1-pulse/1400V/20ms) with 10µg of DNA constructs for constitutive tdTomato expression (PB-Hygro-PGK-CAG-tdTomato and PBase pCMV-Transposase). Cells were subsequently plated on matrigel in mTeSR1 medium supplemented with ROCK inhibitor for 24h, then antibiotic selection with 20µg/ml hygromycin was applied for 12d, before cytometric sorting for tdTomato constitutively-labelled cells.

Sexually mature NZW female rabbits (HyPharm ; Roussay, France) were super-ovulated. 60h after artificial insemination. Fertilized embryos (8-cell stage ; E1.5) were flushed from explanted oviducts by using Euroflush® (IMV Technologies) and cultured in RDH medium (1/3 volume of DMEM-GlutaMAX, 1/3 volume of RPMI-GlutaMAX, and 1/3 volume of Ham's F10-GlutaMAX; Life Technologies) at 38°C, 5% CO_2.

Human PSCs were dissociated into single-cell suspension with trypsin, and 5–10 cells were micro-injected under the mucus coat and zona pellucida of morula 8-cell stage rabbit embryos, the day after collection. After microinjection, embryos were sequentially cultured in CDK8/19i media for 4h, followed by 20h incubation with 1:1 mixture of RDH:CDK8i media and finally in RDH medium for extended in vitro culturing. After 24h of in vitro culture, early blastocyst stage embryos (E3.5) were rinsed x3 in embryo-holding medium (IMV Technologies) and treated with 5mg/ml protease E (Sigma), 3min at 37°C, to digest mucus coat and weaken zona pellucida. Embryos were then rinsed x3 in 199 HEPES medium (Sigma) and cultured in RDH medium for 3d until late-blastocyst stage (E5.5). Rabbit embryos were fixed in 2% para-formaldehyde for 20min at room temperature, washed in PBS + 0.1% Tween-20, and permeabilized in PBS + 1% Tween-20 overnight at 4°C. After 1h blocking with 5% donkey serum, embryo immunofluorescence was as described[36]. Antibodies: **Supplementary Table 8.**

MOLECULAR METHODS

Transcriptional CDK inhibitors

Structure and characterization of the CNIO CDK8/19 inhibitor (CDK8/19i-47799), and notes on all other transcriptional CDK inhibitors used in this study, see: **Supplementary Table 1.**

Small molecule inhibitor characterization assays

In vitro enzyme-inhibitor quantitative assays with recombinant proteins were performed using LanthaScreen ™ Eu-Kinase Binding Assay (Invitrogen) for: CDK8/Cyclin C, CDK9/CyclinT; CDKs, DYRK1A, GSK3β, mTOR, PI3K, PIM1/2, FLT3, KDR, KIT, PDGR-α and SRC. Data from small molecule inhibitor characterization assays are summarized in **Supplementary Table 1**.

Generation of CDK8/19-double knockout iPS cells

To target mouse CDK19, we designed sgRNA against CDK19 exon1, targeting 76bp downstream of the ATG start of translation to generate indels (schematic: **Extended Data Fig.1I**). **Supplementary Table 8**, for sgRNA sequences, and plasmid details. Primary CDK8-flox/flox RERT-Cre MEFs of passage P+1 to P+4 were infected with lenti-CRISPR-Cas9 containing the CDK19 sgRNA (pLenti-CRISPRV2; Addgene #52961) followed by selection with puromycin (1μg/ml). CDK19-knockout was assessed by Western blot. The MEFs were reprogrammed to iPS, where single clones were picked, expanded and CRISPR-induced indels characterized by sequencing the CDK19 target region for frameshift mutations. CDK19-knockout iPS clones were compared versus iPS clones which retained wild-type CDK19 expression, and no effect of CDK19-knockout was observed in MEFs or in iPS cells. CDK8-knockout was induced by 6d culture with 0.5μM 4-hydroxy-tamoxifen to induce Cre-mediated deletion of CDK8 exon2 (schematic: **Extended Data Fig.1J**). CDK8-knockout was confirmed by allele-specific PCR (to demonstrate exon 2 deletion; **Extended Data Fig.1H**), and by western blot (to demonstrate complete loss of CDK8 protein; **Extended Data Fig.1J,K**).

Stable exogenous expression of CDK8

Wild-type CDK8 (CDK8-WT) and catalytically-inactivated kinase-dead CDK8 (CDK8-KD; D173A) were cloned into pMSCV-puro-IRES-GFP (Addgene #21654) using BglII and HpaI restriction enzymes, and confirmed by sequencing. Retroviral supernatants were generated in 293T cells with packaging plasmid pCl-Eco (Addgene#12371), followed by retroviral expression into CDK8/19-double knockout iPS cells. Two rounds of FACS-selection by GFP expression were used to enrich for expressing cells, and CDK8-WT or CDK8-KD protein expression was confirmed by Western blot (schematic and Western blot data in **Fig.1K** and **Extended Data Fig.1L**).

FACS Cytometry

FACS for SSEA1 or ICAM1 was performed with FlowJo 9.6.2 software as described[84]. Live cell analysis of the *Nanog*-GFP used 2i-adapted mouse PSCs to define the threshold (95% of cells) for the homogenous *Nanog*-GFP^{high} population, against which other treatment were compared (**Fig.1A,B**). Live cell sorting for human PSC carrying HERVH-GFP selected the top 10% GFP-expressing cells, as previously described[37]. **Extended Data Figure 9G**: FACS gating strategy for live/dead cell discrimination.

Cell lysis, fractionation, and Western blot

As described[84]. Antibodies: **Supplementary Table 8**. Nuclear/Cytoplasmic fractionation: using NE-PER kit (Thermofisher #78833).

G-banding Karyotype methodology

Subconfluent mouse and human PSC lines were arrested at metaphase stage by adding 0,02µg/ml KaryoMax colcemid (Gibco). 20 metaphase spreads were analysed per condition.

Mouse embryo immunohistochemistry and immunofluorescence

Mouse tissues were fixed in formalin at 4ºC, embedded in paraffin block, and sectioned at 5µm thickness. Staining by standard methods. Antibodies: **Supplementary Table 8.**

Mouse teratoma and embryoid body immunohistochemistry

As described[84]. Antibodies : **Supplementary Table 8.**

Cell immunofluorescence

PSCs were grown on chamber slides using culture conditions indicated in each experiment. Confocal immunofluorescence staining and microscopy, as described[35,84] using Leica SP5 microscope. Antibodies: **Supplementary Table 8**.

DNA methylation

Global DNA methylation status was quantified by mass spectrometry. CpG methylation status at individual CpG sites of repeat DNA regions assessed by DNA bisulphite-conversion and pyrosequencing. Primers for PCR amplification and sequencing: **Supplementary Table 8**.

Image analysis

All image analysis was done using Fiji software (http://fiji.sc).

PROTEOMICS

Full proteome quantitative analysis of 5 mouse ES cell lines

Five mouse ES cell lines (ZS, TNGA, TON, BL6 and V6.4) were cultured in serum/LIF (as control), or additionally, with either 2i or CDK8/19i for >2 weeks. Cell pellets were collected by trypsinization, washed with cold 1xPBS and preserved immediately at -80ºC for further analysis. Protein sample preparation for mass spectrometry, protein digestion, our scheme for isobaric labelling with iTRAQ8plex, detailed settings for high pH reverse phase fractionation, detailed settings for the whole proteome LC-MS/MS, and bioinformatic analyses with the whole Proteome data, as described[84].

Phospho-proteome analysis of mouse PSC lines after 15 min of inhibitor treatment

Two mouse ES PSC lines (TON and ZS) were cultured in serum/LIF (as control), or additionally, with either 2i or CDK8/19i. Inhibitor treatment of the cells was precisely 15min, after which, cells were collected rapidly by scraping in ice-cold PBS, washed with ice-cold PBS, snap-frozen on dry ice, and preserved at -80ºC for further analysis. Sample preparation for mass spectrometry, protein digestion, isobaric labelling, phosphopeptide enrichment, micro high pH reverse phase fractionation, settings used

for phospho-proteome LC-MS/MS, and bioinformatic analyses with phospho-proteomic data, as described[11,84].

TRANSCRIPTOMICS

RNA isolation and Quantitative real-time PCR (qRT-PCR)

Total RNA was extracted (on-column; RNeasy kit with DNA digestion; Qiagen #74104, #79254), and retro-transcribed into cDNA (Superscript Reverse Transcriptase; Biorad #170-889). Quantitative real time-PCR (qRT-PCR) was performed using GoTaq qPCR Master Mix (Promega A6002) in an ABI PRISM 7700 thermocycler (Applied Biosystem). Input normalization of all qRT-PCR data was by the $2^{-\Delta\Delta Ct}$ method using housekeeping genes β-Actin or Gapdh as indicated in each Figure. Primer list: **Supplementary Table 8**.

RNA-seq transcriptomic analyses

The complete set of reads has been deposited in GEO (GSE112208 and GSE127186). Complete list of meta-analyses expression comparisons between this study and multiple mouse and human published datasets, in vitro and in vivo, see: **Source Data** related to each Figure panel.

For RNA-seq in mouse, samples of 1µg of total RNA (RIN numbers: 9.8-10; Agilent 2100 Bioanalyzer), were used. PolyA+ fractions were processed using TruSeq Stranded mRNA Sample Preparation Kit (Agilent). Adapter-ligated library was completed by PCR with Illumina PE primers (8 cycles) and sequenced for 40 bases in a single-read format (Genome Analyzer IIx, Illumina).

For RNA-seq in human, samples of total RNA (RIN numbers: 9.0-10; Agilent 2100 Bioanalyzer), were used. For library construction 10ng of total RNA samples were processed with SMART-Seq v4 Ultra Low Input RNA Kit (Clontech) by following manufacturer instructions. Resulting cDNA was processed with the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB #E7645). Adapter-ligated libraries were completed by PCR (8 cycles), and sequenced for 50 bases in a single-read format, (Illumina HiSeq2500).

Reads were aligned to the reference mouse genome (GRCm38/mm10) or the human genome (GRCh37/hg19) with TopHat-2.0.4 (using Bowtie 0.12.7 and Samtools 0.1.16, allowing two mismatches and five multi-hits. Transcripts assembly, estimation of abundance, and differential expression, were calculated with Cufflinks 1.3.0. When comparing samples, total read numbers were normalized, and visualized using SeqMiner 1.3.3e or IGV (Integrated Genome Viewer) from the Broad Institute available at: http://software.broadinstitute.org/software/igv/

Functional analyses of differential gene expression

For differential gene expression lists, see: Supplementary Table 2 for mouse PSCs adapted to control serum/LIF, +2i, or +CDK8/19i: see, Supplementary Table 3 for

human PSCs adapted to control/primed, +2i, or +CDK8/19i. Genes were ranked using the FDR q-value statistic to identify significant genes (FDR<0.05), then by fold change in expression. Venn diagrams and hypergeometric testing were performed to assess any significant overlaps. Gene Set Enrichment Analysis (GSEA_Pre-ranked) was performed with MsigDB Hallmarks, C5-Gene Ontology (GO) terms, C2-Curated, KEGG, Reactome and NCI databases, with standard settings, and with 1000 permutations for Kolmogorov-Smirnoff correction for multiple testing. GSEA enrichment data were obtained and ranked according to FDR q-value (significance threshold: FDR q-value <0.25). Heatmaps of expression data were generated using Gene Pattern. RRHO (Rank Rank Hypergeometric Overlap) was performed using the ranked list of Log_2_fold-change in gene expression, or RNA Pol II abundance, using standard settings[85]. Colour intensity of RRHO heatmap indicates the -log10 p-value after Benjamini-Yekutieli correction of the hypergeometric overlap (http://systems.crump.ucla.edu/rankrank/rankranksimple.php)[85].

Analysis of Repeat sequences and Endogenous Retrovirus (ERV) expression was by using Repbase datasets for rodent or human repeat elements and featureCounts. In Extended Data Fig.3H, the total FPKM for RNA expression of LINE L1 sub-types was calculated by grouping and summing by family, and then arranged by evolutionary age[86]. Full list of 3 biological replicates for each viral sub-type and calculation for summary into each viral LINE L1 family: **Supplementary Table 2**, and **Source Data**.

Differential gene expression comparing published mouse/human studies

Gene expression changes have been comprehensively characterized in mouse, primate, and human PSCs in response to over-expression of transcription factors, upon culture in various media cocktails, or in vivo, during development of the mouse or human embryos[87,88] (see full list of datasets and references used here: **Supplementary Tables 2 and 3; Source Data**). We used the marker genesets for each developmental stage, to perform GSEA on the ranked list of genes up/down in the cellular studies of mouse and human. We also performed the analysis in reverse, comparing the genesets of significantly differentially expressed mRNAs up- or down-regulated in our cells, versus, the complete ranked list of differential gene expression in other studies. GSEA results are shown in Fig.4E (mouse) and Fig.4K (human). The readout is the Normalized Enrichment score (NES). Data with P<0.05 and FDR q<0.05 are considered significant and marked with an asterisk (*) in heatmaps of GSEA NES scores.

CHIP-SEQ AND GENOMIC ANALYSES

ChIP Methods and analyses

ChIP-qPCR was performed as described[84] with primers and antibodies listed in **Supplementary Table 8**. ChIP-seq, performed as described[2,3,68,84]. We performed 6 biological replicates for each condition (3: serum/LIF, 2i, CDK8/19i) and for each antibody (3: total RNA Pol II, Ser5P-RNA Pol II, control IgG). Three replicates were used for ChIP-qPCR validations, and the other three replicates were pooled for

sequencing. We note that our RNA Pol II ChIP-seq data for serum/LIF and 2i-naïve cells very closely match ChIP-seq involving the same comparison, that is, mouse PSCs in primed versus 2i-naïve states[23,69]: compare **Fig5H,I** versus **Extended Data Fig.6C**.
Definition of promoter and gene body regions and calculation of RNA Pol II total and Ser5P abundance along genes was based on methods of Young and colleagues[68,84] (see schematic: **Figure 5**; Pol II abundance data: **Supplementary Table 6; Source Data Figure 5**). RNA Pol II abundance was assessed by normalizing the total number of reads between treatments, and using featureCounts to calculate the background-subtracted Log_2_RPKM of RNA Pol II abundance in the indicated regions. Transcription Start Site (TSS) and Transcription Termination Zone (TTZ) were identified using Database of Transcriptional Start Sites (http://dbtss.hgc.jp). Metagenes were aligned +/-5Kb or +/-2Kb around the TSS, and visualized by SeqMiner.

The promoter, gene body and transcription termination zone (TTZ), and the ratios between these three regions for each gene (see schematic in **Fig.5J** and **Extended Data Fig.6D, Supplementary Table 6**), were defined similar to previous descriptions[68,84]. Total and Ser5P RNA Pol II abundance were quantified at promoter, gene body and TTZ for 31,167 Refseq gene loci where the transcription start and stop sites are known (**Supplementary Table 6**) in four steps, similar to previous reports[68]: (i) the number of reads per nucleotide was computed with BEDTools 'genomecov'; (ii), to extend this number to the number of reads per gene promoter or gene body, BEDTools 'map' was used; (iii), to correct for region size, the RNA Pol II abundance was calculated as: ((number of reads in region / region size)*scaling factor)*10^5, where scaling factor = (total number of reads in sample/genome length); (iv), for the analysis of Pol II abundance according to inhibitor treatment, genes were first filtered for high confidence Pol II detected at threshold of >3,000 units at the promoter, and detected in all three conditions (serum/LIF, 2i or CDK8/19i), yielding 12,072 genes (**Supplementary Table 6** and **Source Data**, for filtering and calculations). In **Fig.5H**, and **Extended Data 6A,C**, genes were arranged in rank by the abundance of RNA Pol II in the promoter region in the control serum/LIF condition.

CDK8/19 enrichment across the genome of wild-type mixed background V6.5 (C57BL/6-129) mouse PSCs was determined using published dataset: GSE44286, GSM1082346, as previously described[2,3], with peak calling by MACS v1.4.1 using standard settings and compared to input negative control. Note, the ChIP antibody for this ChIP-seq (Santa Cruz #sc-1521) is reported to bind to both CDK8 and CDK19[89]. Peak annotation within local genomic features, was done using HOMER and the enhancer regions previously defined as constituent regions of typical enhancers (n=9,981) or super-enhancers (n=646)[2,3], and of super-enhancer extended regions (n=231) as defined[2,3], where enhancers were defined by co-enrichment for Oct4, Sox2, Nanog, and Med1. For peak calls, CDK8/19 abundance at called peaks, and loci annotations, see: **Supplementary Table 7** and **Source Data**. Naïve-specific or primed-specific enhancer regions were defined by filtering the PREStige database of enhancers[77] which identifies enhancers by enrichment of H3K4me1 mono-methylation

in multiple tissues and lineages. Using the PREStige dataset, we identified enhancer regions with H3K4me1 enrichment >20 units, and specific only to pre-implantation naïve PSCs, or post-implantation EpiSC cells, versus all other tissue-specific enhancer regions listed in the database (~120,000), by subtraction of overlapping enhancers (1bp overhang threshold) as outlined in the schematic of **Extended Data Fig.8E**; see **Source Data**: for naïve ES-specific enhancers (n=1,424), or EpiSC-specific enhancers (n=1,005) lists. To identify the single-nearest target gene to each PSC super-enhancer and analyze their biological functions, GREAT v3.0.0[90] analysis was performed with standard settings, using the list of CDK8/19 peaks identified above (**Supplementary Table 7**); and for Gene Ontology analysis of target-gene functions, reporting the $-\log_{10}$ binomial P-value with conservative Bonferroni correction for multiple hypothesis testing[90]. Correlation matrix of ChIP-seq data in **Extended Data Fig.7D** was produced using Morpheus software, available from Broad Institute: https://software.broadinstitute.org/morpheus/.

For **Fig.7B,C**, GSEA was run with a geneset of the single nearest genes to super-enhancers (as identified by GREAT analysis above using standard settings (GREAT v3.0.0)[90], using the super-enhancers previously described in mouse PSCs[2,3], versus, the ranked list of differential gene expression determined by RNA-Seq for serum/LIF control compared to CDK8/19i-adapted mouse ES PSCs. See **Source Data**, for SE-target and expression-matched control gene lists.

For **Extended Data Fig.8H**, GREAT analysis using standard settings (GREAT v3.0.0)[90] was used to identify the set of single nearest genes (n=3,553 genes) to enhancer regions previously identified in mouse PSCs (n=10,627)[2,3]. Then the Log_2_fold-change in RNA expression of these genes from the current study was ranked high-to-low (S/L vs 2i; S/L vs CDK8/19i), and the extent of overlap with hypergeometric test of significance, of these two ranked lists is shown as heatmap in Extended Data Fig.8H, performed by RRHO (Rank Rank Hypergeometric Overlap)[85] using standard settings, available at: http://systems.crump.ucla.edu/rankrank/rankranksimple.php. Colour intensity of RRHO heatmap indicates the -log10 p-value after Benjamini-Yekutieli correction of the hypergeometric overlap.

STATISTICS AND DATA AVAILABILITY

Statistics and reproducibility

Unless otherwise specified, quantitative data are presented as Mean+/-SD and significance was assessed by the two-tailed Student's t test; *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. RRHO (Rank Rank Hypergeometric Overlap) was performed as described[85] using standard settings and after Benjamini-Yekutieli correction of the hypergeometric overlap. Except when annotated otherwise, each experiment shown was carried out three times, with similar outcomes. Statistical

analyses are described in detail for each panel. No statistical methods were used to pre-determine the sample size. Briefly, for differential gene expression by RNA-Seq, a threshold of FDR q-value of $q<0.05$, or $q<0.01$ was applied, as indicated in each case. In GSEA analysis, the standard threshold for significance was applied, where $p<0.05$ and FDR q-value<0.25. Genes that were differentially expressed in the RNA-Seq were called using DESeq2 or Cufflinks 1.3.0 (details above). Immunofluorescent image analysis is described in detail above in section on embryo analysis. Statistics were performed using MACS for peak calling of the ChIP-Seq experiments. Statistical analyses of ChIP-qPCR, quantitative RT-PCR, cell culture experiments used Prism GraphPad Software (version 7.03) or Microsoft Excel.

Data availability

RNA-seq and ChIP-seq data are available from the GEO database under accession numbers GSE112208 and GSE127186. The mass spectrometry proteomics data are available from the ProteomeXchange Consortium/PRIDE repository with the dataset identifier PXD009200. Published datasets included in this study: **Source Data Fig.4E,K**. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

SUPPLEMENTARY REFERENCES

81. Diéguez-Hurtado, R. et al. A Cre-reporter transgenic mouse expressing the far-red fluorescent protein Katushka. *Genesis* **49**, 36–45 (2011).
82. Zalzman, M. et al. Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature* **464**, 858–863 (2010).
83. Porter, D. C. et al. Cyclin-dependent kinase 8 mediates chemotherapy-induced tumor-promoting paracrine activities. *Proc. Natl. Acad. Sci.* **109**, 13799–13804 (2012).
84. Lynch, C. J. et al. The RNA Polymerase II Factor RPAP1 Is Critical for Mediator-Driven Transcription and Cell Identity. *Cell Rep.* **22**, 396–410 (2018).
85. Plaisier, S. B., Taschereau, R., Wong, J. A. & Graeber, T. G. Rank–rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures. *Nucleic Acids Res.* **38**, e169 (2010).
86. Castro-Diaz, N. et al. Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes Dev.* **28**, 1397–1409 (2014).
87. Papadopoulou, T., Kaymak, A., Sayols, S. & Richly, H. Dual role of Med12 in PRC1-dependent gene repression and ncRNA-mediated transcriptional activation. *Cell Cycle* **15**, 1479–1493 (2016).
88. Xiang, L. et al. A developmental landscape of 3D-cultured human pre-gastrulation embryos. *Nature* **577**, 537–542 (2020).
89. Broude, E. V et al. Expression of CDK8 and CDK8-interacting Genes as Potential Biomarkers in Breast Cancer. *Curr. Cancer Drug Targets* **15**, 739–49 (2015).

90. McLean, C. Y. et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat*. *Biotechnol*. **28**, 495–501 (2010).

Fig. 1

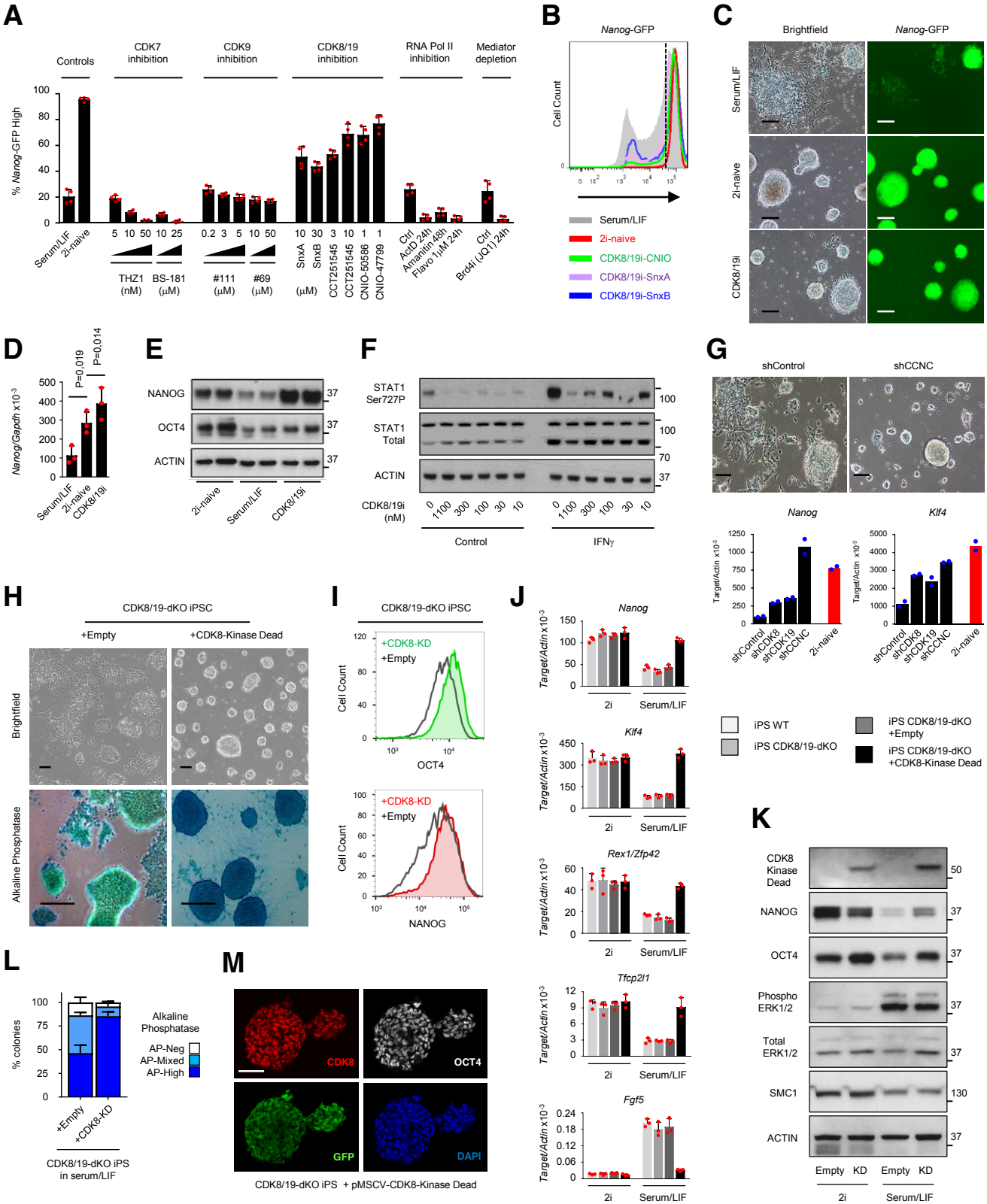

A
Controls
CDK7 inhibition
CDK9 inhibition
CDK8/19 inhibition
RNA Pol II inhibition
Mediator depletion
% Nanog-GFP High
Serum/LIF
2i-naive
THZ1 (nM)
BS-181 (μM)
#111 (μM)
#69 (μM)
SnxA (μM)
SnxB
CCT251545
CNIO-50586
CNIO-47799
Ctrl
ActD 24h
Amanitin 48h
Flavo 1μM 24h
Brd4i (JQ1) 24h
B
Nanog-GFP
Cell Count
Serum/LIF
2i-naive
CDK8/19i-CNIO
CDK8/19i-SnxA
CDK8/19i-SnxB
C
Brightfield
Nanog-GFP
Serum/LIF
2i-naive
CDK8/19i
D
Nanog/Gapdh x10-3
P=0,019
P=0,014
Serum/LIF
2i-naive
CDK8/19i
E
NANOG
OCT4
ACTIN
2i-naive
Serum/LIF
CDK8/19i
F
STAT1 Ser727P
STAT1 Total
ACTIN
CDK8/19i (nM)
Control
IFNγ
G
shControl
shCCNC
Nanog
Klf4
Target/Actin x10-3
shControl
shCDK8
shCDK19
shCCNC
2i-naive
iPS WT
iPS CDK8/19-dKO
iPS CDK8/19-dKO +Empty
iPS CDK8/19-dKO +CDK8-Kinase Dead
H
CDK8/19-dKO iPSC
+Empty
+CDK8-Kinase Dead
Brightfield
Alkaline Phosphatase
I
CDK8/19-dKO iPSC
+CDK8-KD
+Empty
Cell Count
OCT4
NANOG
J
Nanog
Klf4
Rex1/Zfp42
Tfcp2l1
Fgf5
Target/Actin x10-3
2i
Serum/LIF
K
CDK8 Kinase Dead
NANOG
OCT4
Phospho ERK1/2
Total ERK1/2
SMC1
ACTIN
Empty
KD
2i
Serum/LIF
L
% colonies
Alkaline Phosphatase
AP-Neg
AP-Mixed
AP-High
+Empty
+CDK8-KD
CDK8/19-dKO iPS in serum/LIF
M
CDK8
OCT4
GFP
DAPI
CDK8/19-dKO iPS + pMSCV-CDK8-Kinase Dead

Fig. 2

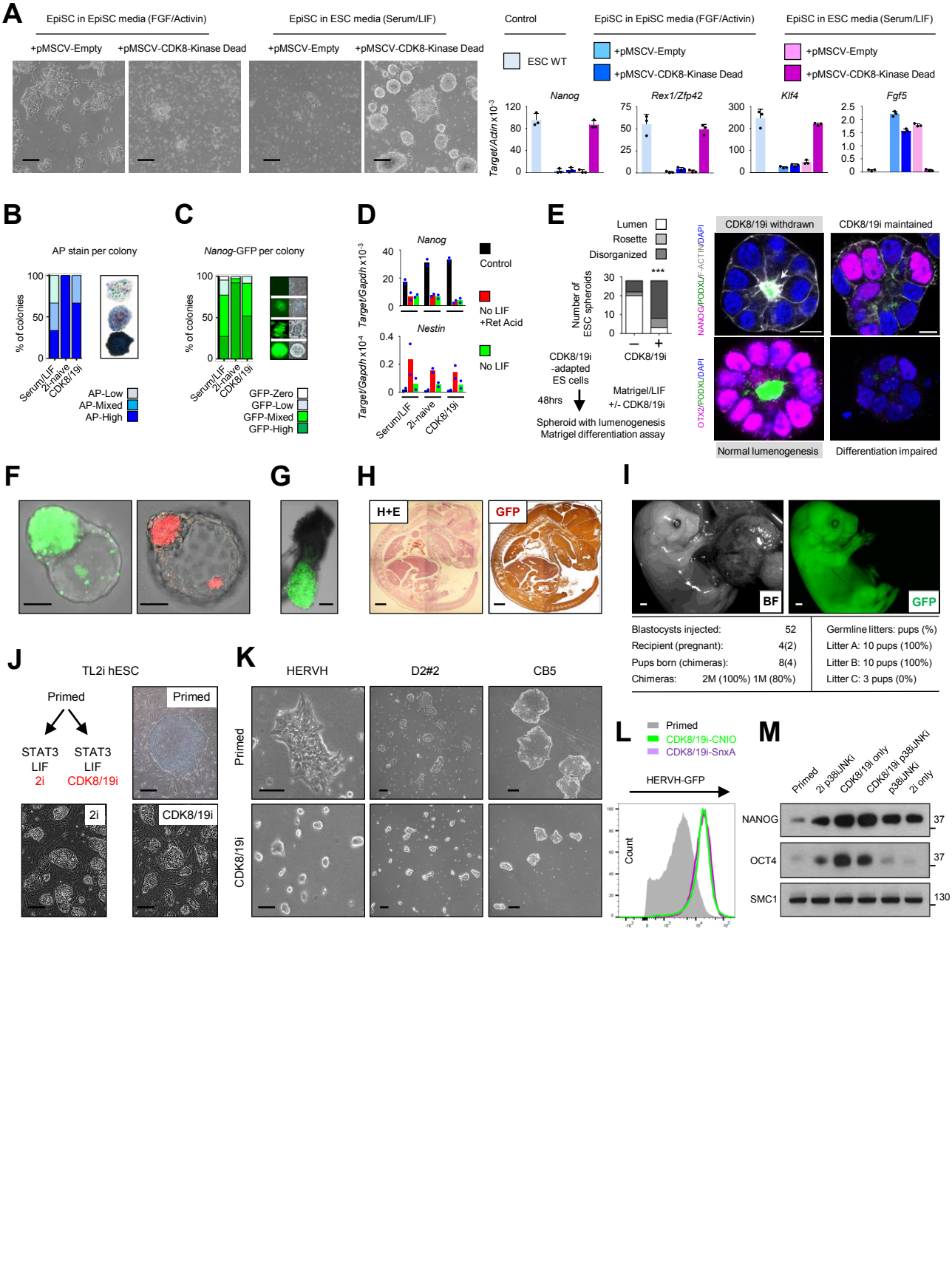

A
EpiSC in EpiSC media (FGF/Activin)
EpiSC in ESC media (Serum/LIF)
+pMSCV-Empty
+pMSCV-CDK8-Kinase Dead
Control
ESC WT
Nanog
Rex1/Zfp42
Klf4
Fgf5
Target/Actin x10-3
B
AP stain per colony
% of colonies
Serum/LIF
2i-naive
CDK8/19i
AP-Low
AP-Mixed
AP-High
C
Nanog-GFP per colony
GFP-Zero
GFP-Low
GFP-Mixed
GFP-High
D
Target/Gapdh x10-3
Target/Gapdh x10-4
Control
No LIF +Ret Acid
No LIF
E
Lumen
Rosette
Disorganized
Number of ESC spheroids
CDK8/19i -adapted ES cells
48hrs
Matrigel/LIF +/- CDK8/19i
Spheroid with lumenogenesis Matrigel differentiation assay
CDK8/19i withdrawn
CDK8/19i maintained
NANOG/PODXL/F-ACTIN/DAPI
OTX2/PODXL/DAPI
Normal lumenogenesis
Differentiation impaired
F
G
H
H+E
GFP
I
BF
GFP
Blastocysts injected: 52
Recipient (pregnant): 4(2)
Pups born (chimeras): 8(4)
Chimeras: 2M (100%) 1M (80%)
Germline litters: pups (%)
Litter A: 10 pups (100%)
Litter B: 10 pups (100%)
Litter C: 3 pups (0%)
J
TL2i hESC
Primed
STAT3 LIF 2i
STAT3 LIF CDK8/19i
2i
CDK8/19i
K
HERVH
D2#2
CB5
L
Primed
CDK8/19i-CNIO
CDK8/19i-SnxA
HERVH-GFP
Count
M
Primed
2i p38iJNKi
CDK8/19i only
CDK8/19i p38iJNKi
p38iJNKi
2i only
NANOG
OCT4
SMC1

Fig. 3

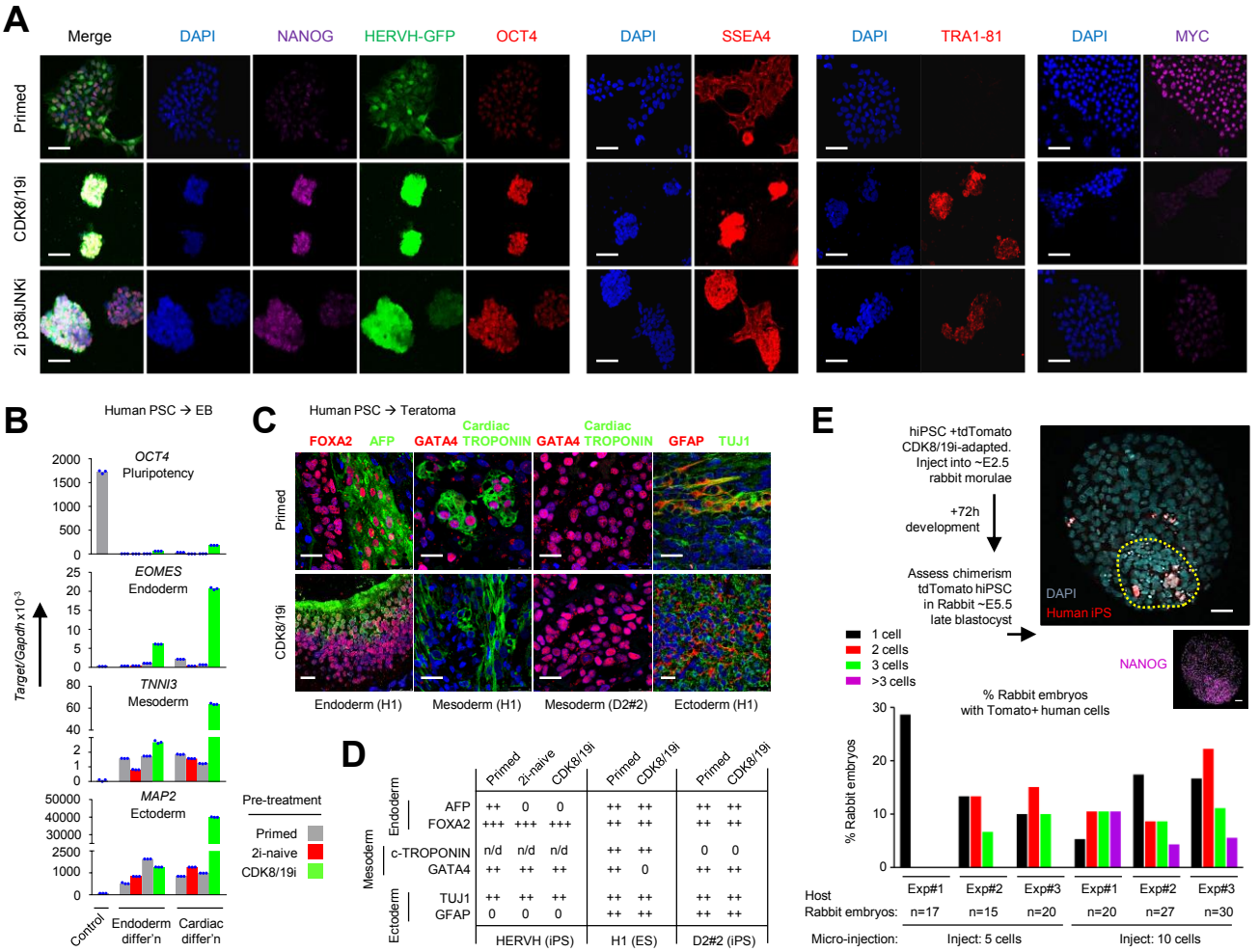

A
Merge
DAPI
NANOG
HERVH-GFP
OCT4
DAPI
SSEA4
DAPI
TRA1-81
DAPI
MYC
Primed
CDK8/19i
2i p38iJNKi
B
Human PSC → EB
OCT4
Pluripotency
EOMES
Endoderm
TNNI3
Mesoderm
MAP2
Ectoderm
Target/Gapdh x10-3
Control
Endoderm differ'n
Cardiac differ'n
Pre-treatment
Primed
2i-naive
CDK8/19i
C
Human PSC → Teratoma
FOXA2
AFP
GATA4
Cardiac TROPONIN
GATA4
Cardiac TROPONIN
GFAP
TUJ1
Primed
CDK8/19i
Endoderm (H1)
Mesoderm (H1)
Mesoderm (D2#2)
Ectoderm (H1)
D
Primed
2i-naive
CDK8/19i
Primed
CDK8/19i
Primed
CDK8/19i
Endoderm
AFP ++ 0 0 ++ ++ ++ ++
FOXA2 +++ +++ +++ ++ ++ ++ ++
Mesoderm
c-TROPONIN n/d n/d n/d ++ ++ 0 0
GATA4 ++ ++ ++ ++ 0 ++ ++
Ectoderm
TUJ1 ++ ++ ++ ++ ++ ++ ++
GFAP 0 0 0 ++ ++ ++ ++
HERVH (iPS)
H1 (ES)
D2#2 (iPS)
E
hiPSC +tdTomato
CDK8/19i-adapted.
Inject into ~E2.5
rabbit morulae
+72h
development
Assess chimerism
tdTomato hiPSC
in Rabbit ~E5.5
late blastocyst
DAPI
Human iPS
NANOG
1 cell
2 cells
3 cells
>3 cells
% Rabbit embryos
with Tomato+ human cells
% Rabbit embryos
Host
Rabbit embryos:
Exp#1 n=17
Exp#2 n=15
Exp#3 n=20
Exp#1 n=20
Exp#2 n=27
Exp#3 n=30
Micro-injection:
Inject: 5 cells
Inject: 10 cells

Fig. 4

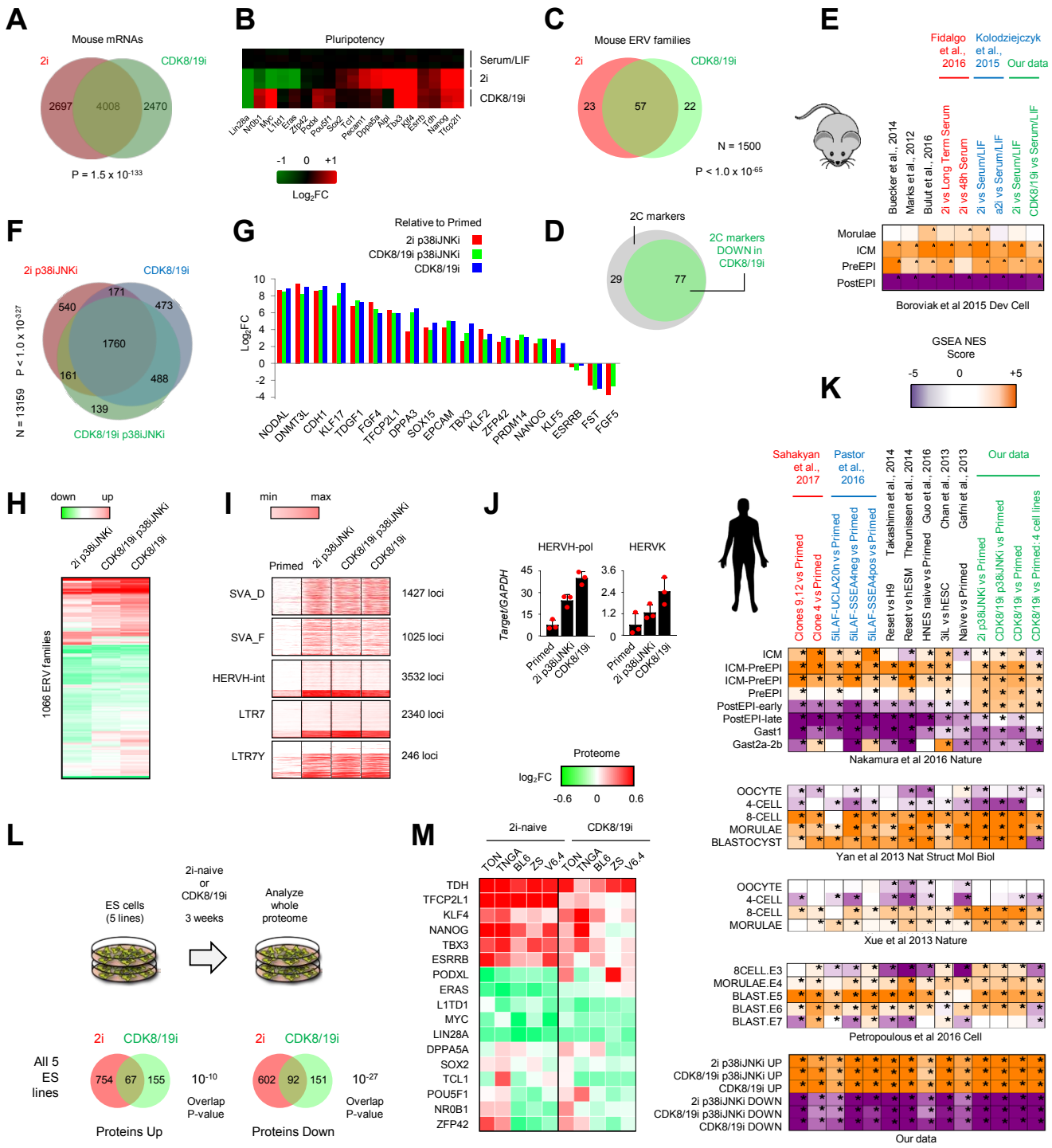

A
Mouse mRNAs
2i
CDK8/19i
2697
4008
2470
P = 1.5 x 10-133
B
Pluripotency
Serum/LIF
2i
CDK8/19i
Lin28a
Nr0b1
Myc
L1td1
Eras
Zfp42
Tbx3
Pou5f1
Sox2
Tcl1
Pecam1
Dppa5a
Akp1
Tbx3
Nr0b1
Esrrb
Tdh
Nanog
Tfcp2l1
-1 0 +1
Log2FC
C
Mouse ERV families
2i
CDK8/19i
23
57
22
N = 1500
P < 1.0 x 10-65
E
Fidalgo et al., 2016
Kolodziejczyk et al., 2015
Our data
Buecker et al., 2014
Marks et al., 2012
Bulut et al., 2016
2i vs Long Term Serum
2i vs 48h Serum
2i vs Serum/LIF
a2i vs Serum/LIF
2i vs Serum/LIF
CDK8/19i vs Serum/LIF
Morulae
ICM
PreEPI
PostEPI
Boroviak et al 2015 Dev Cell
F
2i p38iJNKi
CDK8/19i
171
540
473
1760
161
488
139
CDK8/19i p38iJNKi
N = 13159
P < 1.0 x 10-327
G
Relative to Primed
2i p38iJNKi
CDK8/19i p38iJNKi
CDK8/19i
Log2FC
NODAL
DNMT3L
CDH1
KLF17
TDGF1
FGF4
TFCP2L1
DPPA3
SOX15
EPCAM
TBX3
KLF2
ZFP42
PRDM14
NANOG
KLF5
ESRRB
FST
FGF5
D
2C markers
2C markers DOWN in CDK8/19i
29
77
K
GSEA NES Score
-5 0 +5
H
down
up
2i p38iJNKi
CDK8/19i p38iJNKi
CDK8/19i
1066 ERV families
I
min
max
Primed
2i p38iJNKi
CDK8/19i p38iJNKi
CDK8/19i
SVA_D
SVA_F
HERVH-int
LTR7
LTR7Y
1427 loci
1025 loci
3532 loci
2340 loci
246 loci
J
HERVH-pol
HERVK
Target/GAPDH
Primed
2i p38iJNKi
CDK8/19i
Proteome
log2FC
-0.6 0 0.6
Sahakyan et al., 2017
Pastor et al., 2016
Our data
Clones 9,12 vs Primed
Clone 4 vs Primed
5iLAF-UCLA20n vs Primed
5iLAF-SSEA4neg vs Primed
5iLAF-SSEA4pos vs Primed
Takashima et al., 2014
Theunissen et al., 2014
Guo et al., 2016
Chan et al., 2013
Gafni et al., 2013
2i p38iJNKi vs Primed
CDK8/19i p38iJNKi vs Primed
CDK8/19i vs Primed
CDK8/19i vs Primed: 4 cell lines
Reset vs H9
Reset vs hESM
HNES naive vs Primed
3iL vs hESC
Naïve vs Primed
ICM
ICM-PreEPI
ICM-PreEPI
PreEPI
PostEPI-early
PostEPI-late
Gast1
Gast2a-2b
Nakamura et al 2016 Nature
OOCYTE
4-CELL
8-CELL
MORULAE
BLASTOCYST
Yan et al 2013 Nat Struct Mol Biol
OOCYTE
4-CELL
8-CELL
MORULAE
Xue et al 2013 Nature
8CELL.E3
MORULAE.E4
BLAST.E5
BLAST.E6
BLAST.E7
Petropoulous et al 2016 Cell
2i p38iJNKi UP
CDK8/19i p38iJNKi UP
CDK8/19i UP
2i p38iJNKi DOWN
CDK8/19i p38iJNKi DOWN
CDK8/19i DOWN
Our data
L
ES cells (5 lines)
2i-naive or CDK8/19i
3 weeks
Analyze whole proteome
All 5 ES lines
2i
CDK8/19i
754
67
155
Proteins Up
10-10
Overlap P-value
602
92
151
Proteins Down
10-27
Overlap P-value
M
2i-naive
CDK8/19i
TON
TNGA
BL6
ZS
V6.4
TDH
TFCP2L1
KLF4
NANOG
TBX3
ESRRB
PODXL
ERAS
L1TD1
MYC
LIN28A
DPPA5A
SOX2
TCL1
POU5F1
NR0B1
ZFP42

Fig. 5

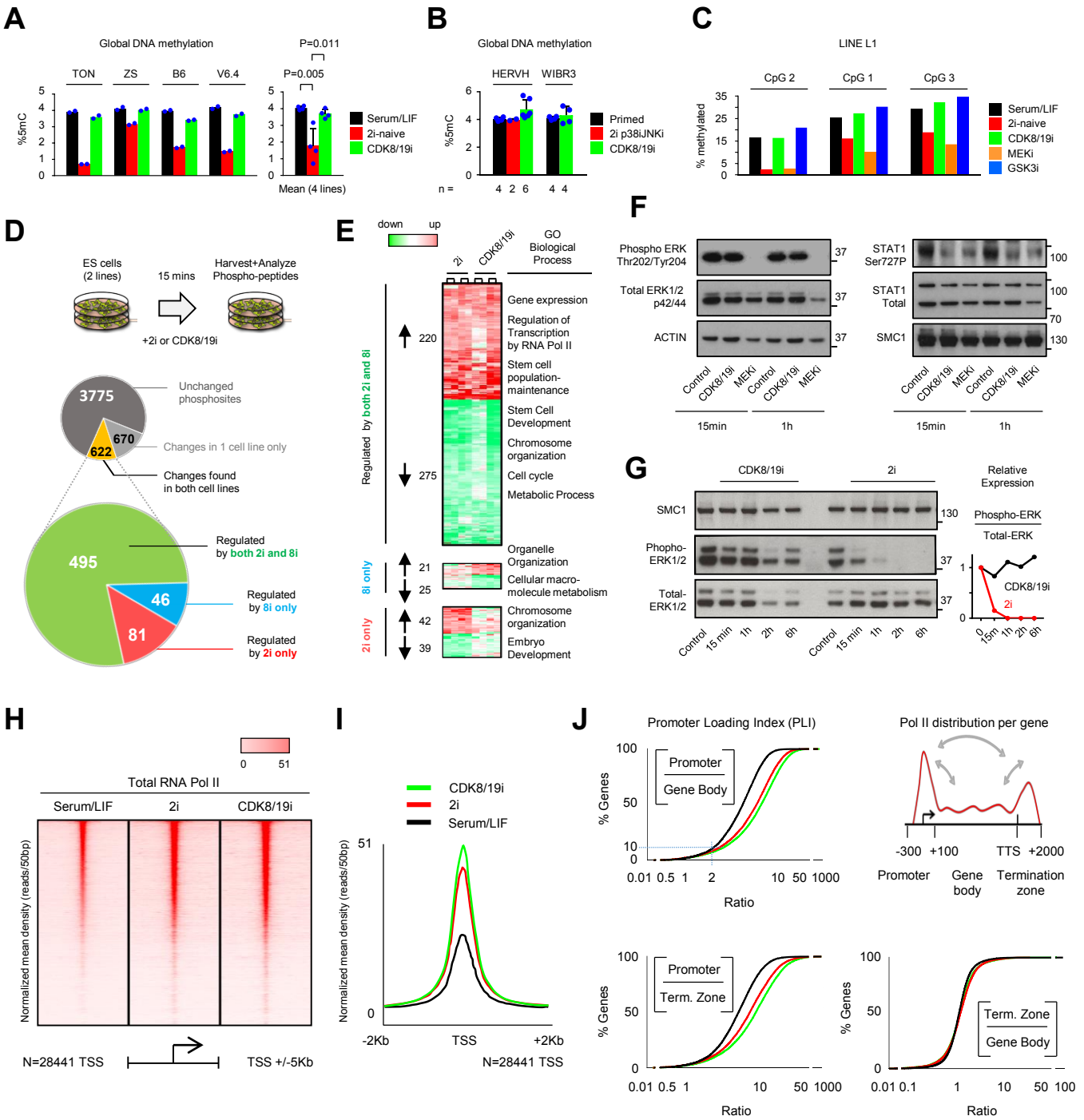

A
Global DNA methylation
TON
ZS
B6
V6.4
P=0.011
P=0.005
%5mC
Mean (4 lines)
Serum/LIF
2i-naive
CDK8/19i
B
Global DNA methylation
HERVH
WIBR3
n = 4 2 6 4 4
Primed
2i p38iJNKi
CDK8/19i
C
LINE L1
CpG 2
CpG 1
CpG 3
% methylated
Serum/LIF
2i-naive
CDK8/19i
MEKi
GSK3i
D
ES cells (2 lines)
15 mins
Harvest+Analyze Phospho-peptides
+2i or CDK8/19i
3775
670
622
Unchanged phosphosites
Changes in 1 cell line only
Changes found in both cell lines
495
46
81
Regulated by both 2i and 8i
Regulated by 8i only
Regulated by 2i only
E
down
up
2i
CDK8/19i
GO Biological Process
Regulated by both 2i and 8i
220
275
8i only
21
25
2i only
42
39
Gene expression
Regulation of Transcription by RNA Pol II
Stem cell population-maintenance
Stem Cell Development
Chromosome organization
Cell cycle
Metabolic Process
Organelle Organization
Cellular macro-molecule metabolism
Chromosome organization
Embryo Development
F
Phospho ERK Thr202/Tyr204
Total ERK1/2 p42/44
ACTIN
STAT1 Ser727P
STAT1 Total
SMC1
Control
CDK8/19i
MEKi
15min
1h
G
CDK8/19i
2i
Relative Expression
SMC1
Phopho-ERK1/2
Total-ERK1/2
Phospho-ERK
Total-ERK
Control
15 min
1h
2h
6h
H
Total RNA Pol II
Serum/LIF
2i
CDK8/19i
Normalized mean density (reads/50bp)
N=28441 TSS
TSS +/-5Kb
I
CDK8/19i
2i
Serum/LIF
Normalized mean density (reads/50bp)
-2Kb
TSS
+2Kb
N=28441 TSS
J
Promoter Loading Index (PLI)
Promoter
Gene Body
% Genes
Ratio
Pol II distribution per gene
-300
+100
TTS
+2000
Promoter
Gene body
Termination zone
Promoter
Term. Zone
Term. Zone
Gene Body

Fig. 6

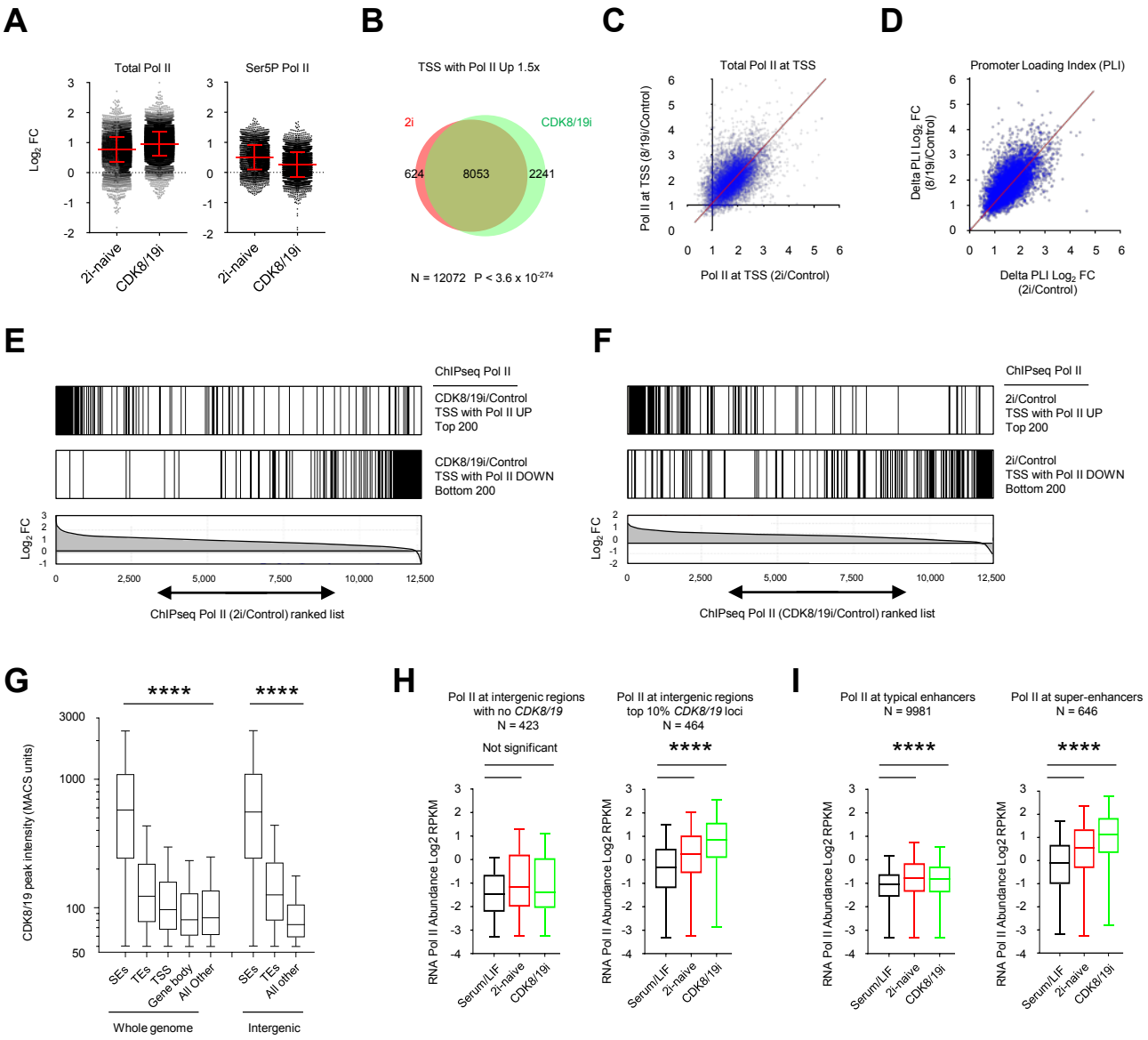

A
Total Pol II
Ser5P Pol II
Log2 FC
2i-naive
CDK8/19i
B
TSS with Pol II Up 1.5x
2i
CDK8/19i
624
8053
2241
N = 12072 P < 3.6 x 10-274
C
Total Pol II at TSS
Pol II at TSS (8/19i/Control)
Pol II at TSS (2i/Control)
D
Promoter Loading Index (PLI)
Delta PLI Log2 FC (8/19i/Control)
Delta PLI Log2 FC (2i/Control)
E
ChIPseq Pol II
CDK8/19i/Control
TSS with Pol II UP
Top 200
CDK8/19i/Control
TSS with Pol II DOWN
Bottom 200
Log2 FC
ChIPseq Pol II (2i/Control) ranked list
F
ChIPseq Pol II
2i/Control
TSS with Pol II UP
Top 200
2i/Control
TSS with Pol II DOWN
Bottom 200
Log2 FC
ChIPseq Pol II (CDK8/19i/Control) ranked list
G
****
****
CDK8/19 peak intensity (MACS units)
SEs
TEs
TSS
Gene body
All Other
SEs
TEs
All other
Whole genome
Intergenic
H
Pol II at intergenic regions with no CDK8/19 N = 423
Not significant
Pol II at intergenic regions top 10% CDK8/19 loci N = 464
****
RNA Pol II Abundance Log2 RPKM
Serum/LIF
2i-naive
CDK8/19i
I
Pol II at typical enhancers N = 9981
****
Pol II at super-enhancers N = 646
****
RNA Pol II Abundance Log2 RPKM
Serum/LIF
2i-naive
CDK8/19i

Fig. 7

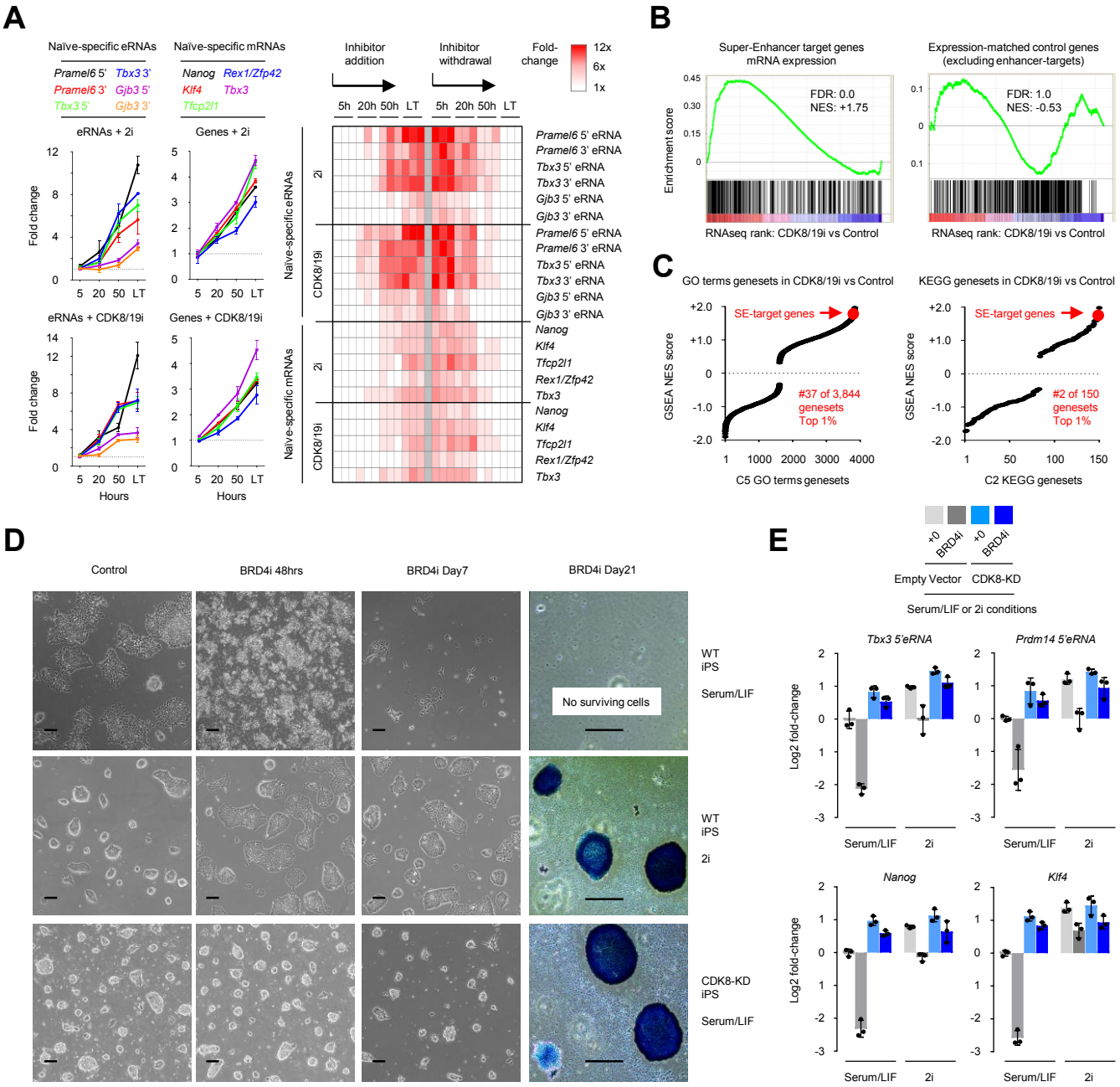
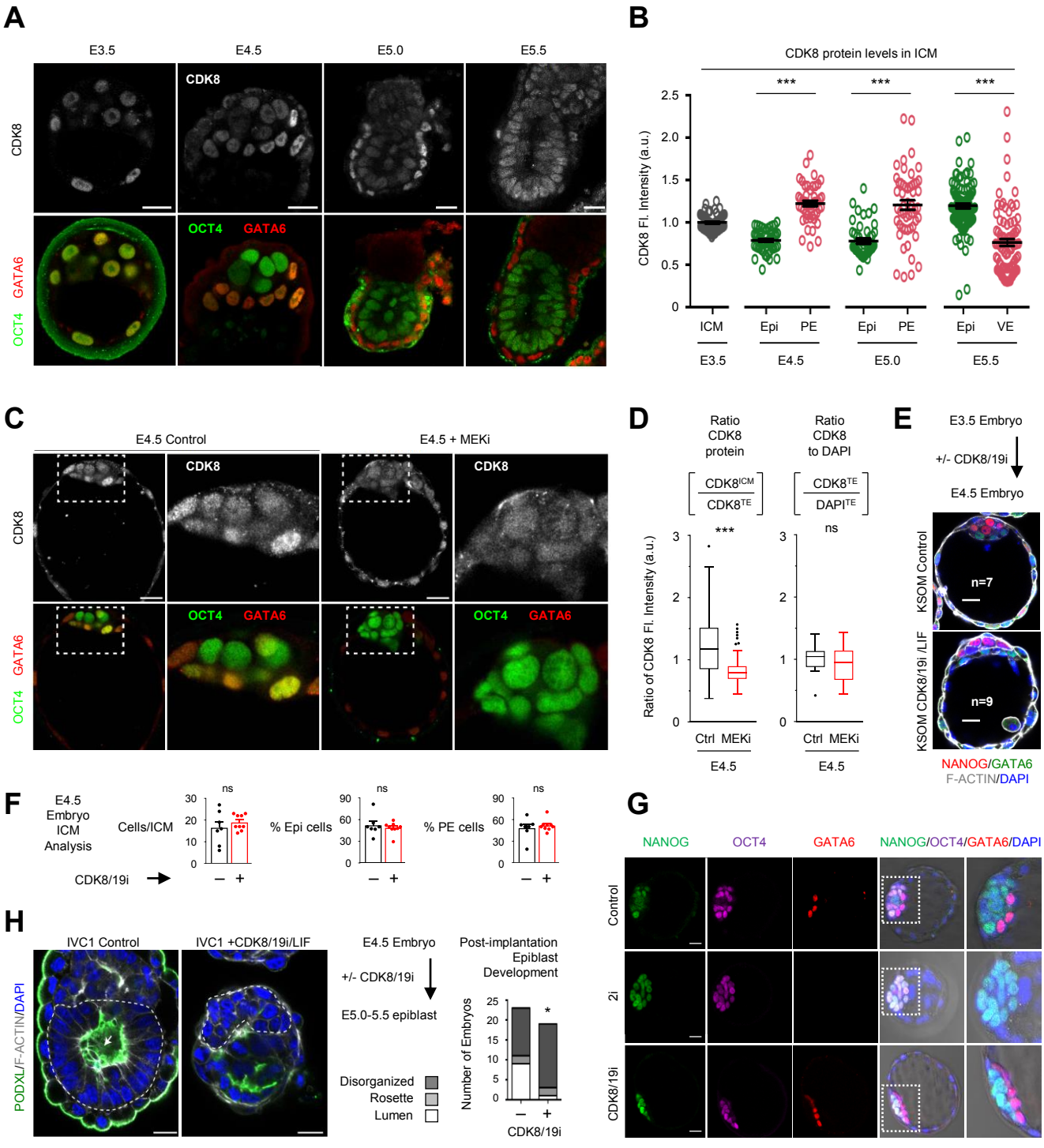
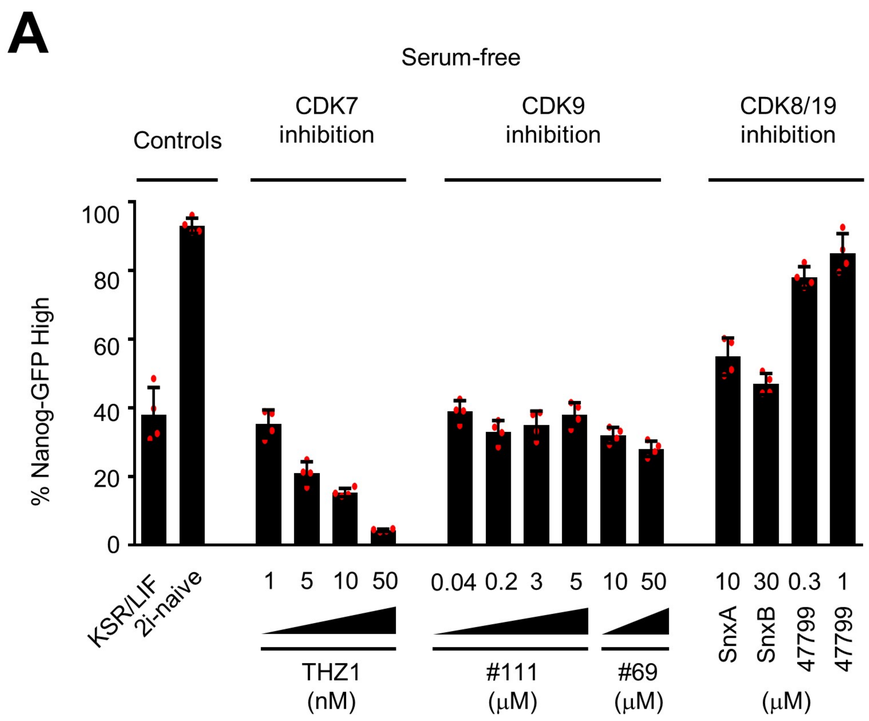
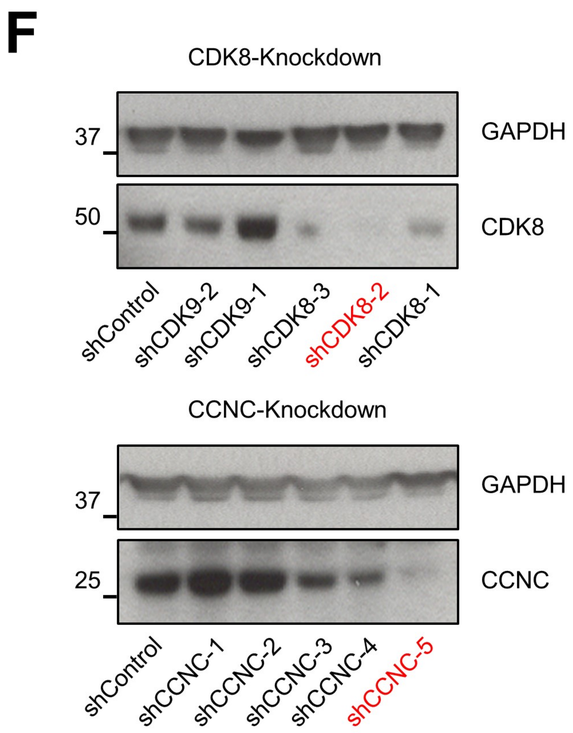

A
Naïve-specific eRNAs
Naïve-specific mRNAs
Pramel6 5'
Tbx3 3'
Pramel6 3'
Gjb3 5'
Tbx3 5'
Gjb3 3'
Nanog
Rex1/Zfp42
Klf4
Tbx3
Tfcp2l1
eRNAs + 2i
Genes + 2i
eRNAs + CDK8/19i
Genes + CDK8/19i
Fold change
Hours
Inhibitor addition
Inhibitor withdrawal
Fold-change
12x
6x
1x
5h 20h 50h LT
Naïve-specific eRNAs
Naïve-specific mRNAs
2i
CDK8/19i
Pramel6 5' eRNA
Pramel6 3' eRNA
Tbx3 5' eRNA
Tbx3 3' eRNA
Gjb3 5' eRNA
Gjb3 3' eRNA
Nanog
Klf4
Tfcp2l1
Rex1/Zfp42
Tbx3
B
Super-Enhancer target genes mRNA expression
Expression-matched control genes (excluding enhancer-targets)
Enrichment score
FDR: 0.0
NES: +1.75
FDR: 1.0
NES: -0.53
RNAseq rank: CDK8/19i vs Control
C
GO terms genesets in CDK8/19i vs Control
KEGG genesets in CDK8/19i vs Control
GSEA NES score
SE-target genes
#37 of 3,844 genesets Top 1%
#2 of 150 genesets Top 1%
C5 GO terms genesets
C2 KEGG genesets
D
Control
BRD4i 48hrs
BRD4i Day7
BRD4i Day21
No surviving cells
WT iPS Serum/LIF
WT iPS 2i
CDK8-KD iPS Serum/LIF
E
+0 BRD4i
Empty Vector
CDK8-KD
Serum/LIF or 2i conditions
Tbx3 5'eRNA
Prdm14 5'eRNA
Nanog
Klf4
Log2 fold-change
Serum/LIF
2i

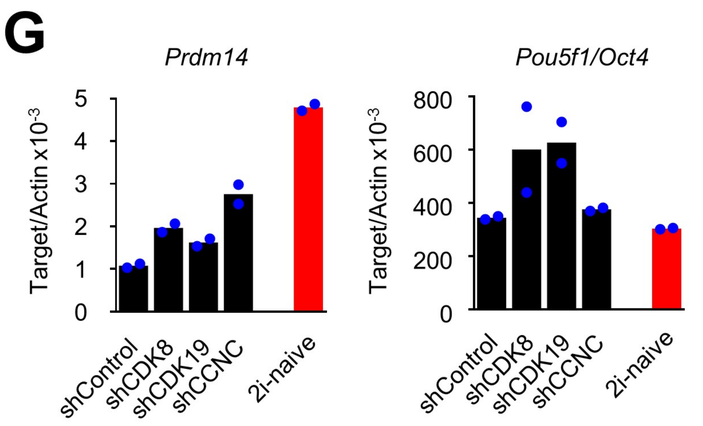
Fig. 8

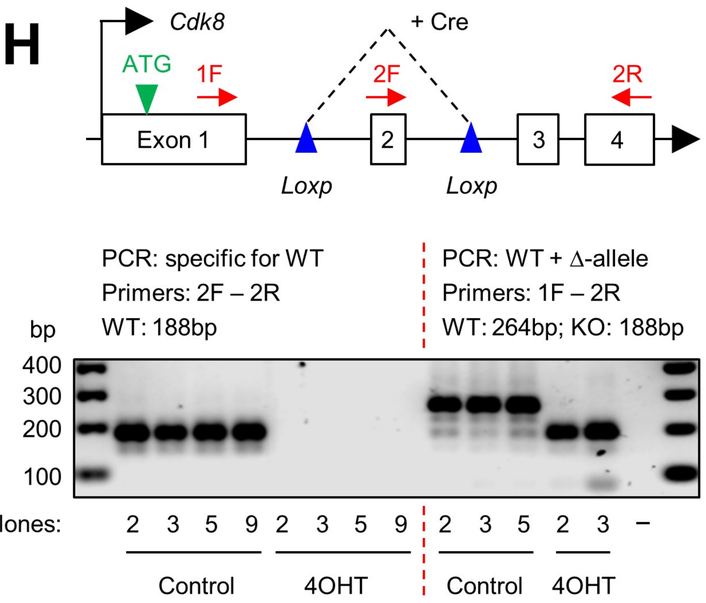

A
E3.5
E4.5
E5.0
E5.5
CDK8
OCT4
GATA6
B
CDK8 protein levels in ICM
CDK8 Fl. Intensity (a.u.)
ICM
Epi
PE
VE
E3.5
E4.5
E5.0
E5.5
C
E4.5 Control
E4.5 + MEKi
D
Ratio CDK8 protein
Ratio CDK8 to DAPI
Ratio of CDK8 Fl. Intensity (a.u.)
Ctrl
MEKi
E
E3.5 Embryo
+/- CDK8/19i
E4.5 Embryo
KSOM Control
KSOM CDK8/19i /LIF
n=7
n=9
NANOG/GATA6
F-ACTIN/DAPI
F
E4.5 Embryo ICM Analysis
Cells/ICM
% Epi cells
% PE cells
CDK8/19i
G
NANOG
OCT4
GATA6
NANOG/OCT4/GATA6/DAPI
Control
2i
CDK8/19i
H
IVC1 Control
IVC1 +CDK8/19i/LIF
PODXL/F-ACTIN/DAPI
E4.5 Embryo
+/- CDK8/19i
E5.0-5.5 epiblast
Disorganized
Rosette
Lumen
Post-implantation Epiblast Development
Number of Embryos
CDK8/19i

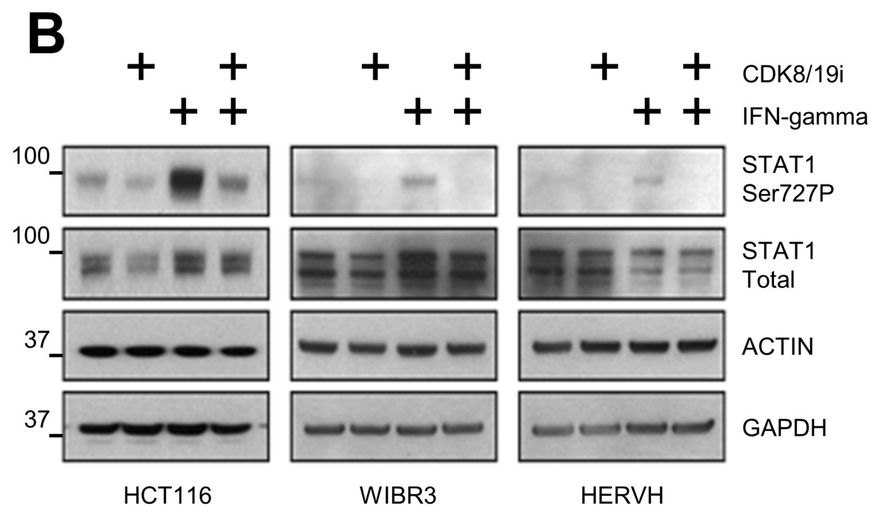
Extended Data Fig. 1

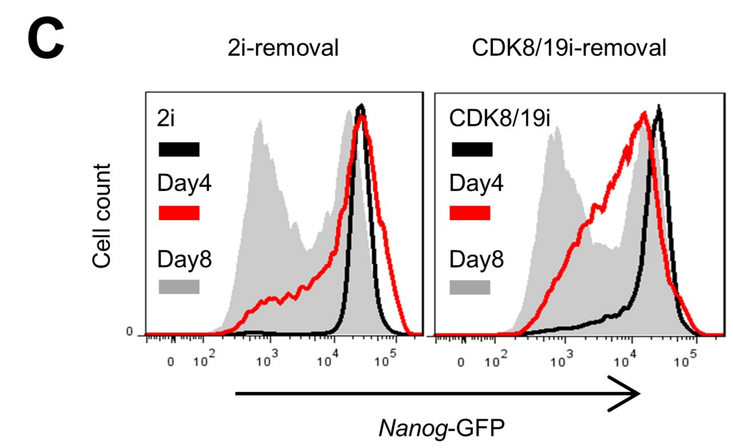


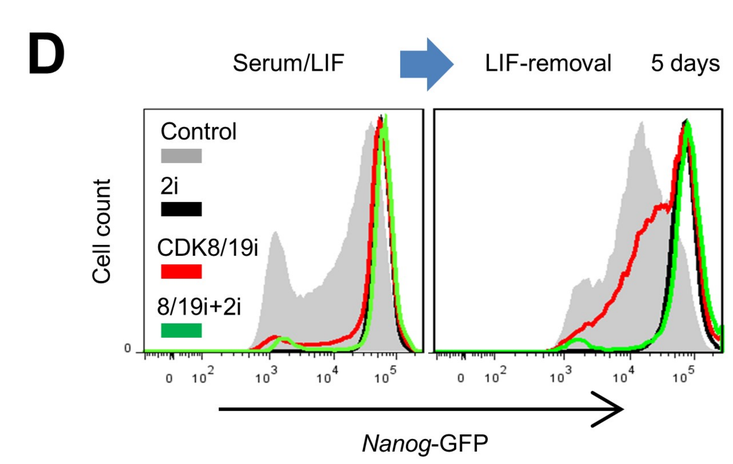


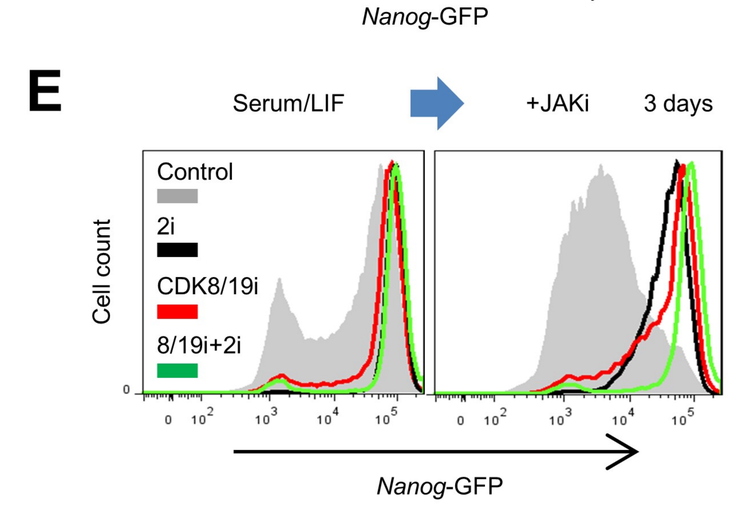


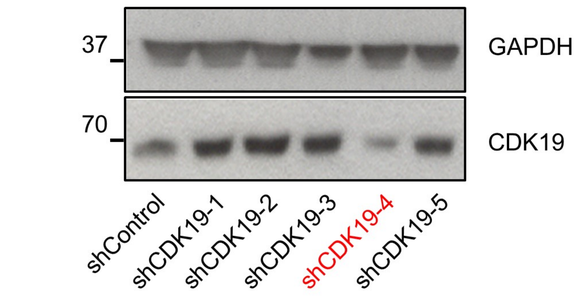


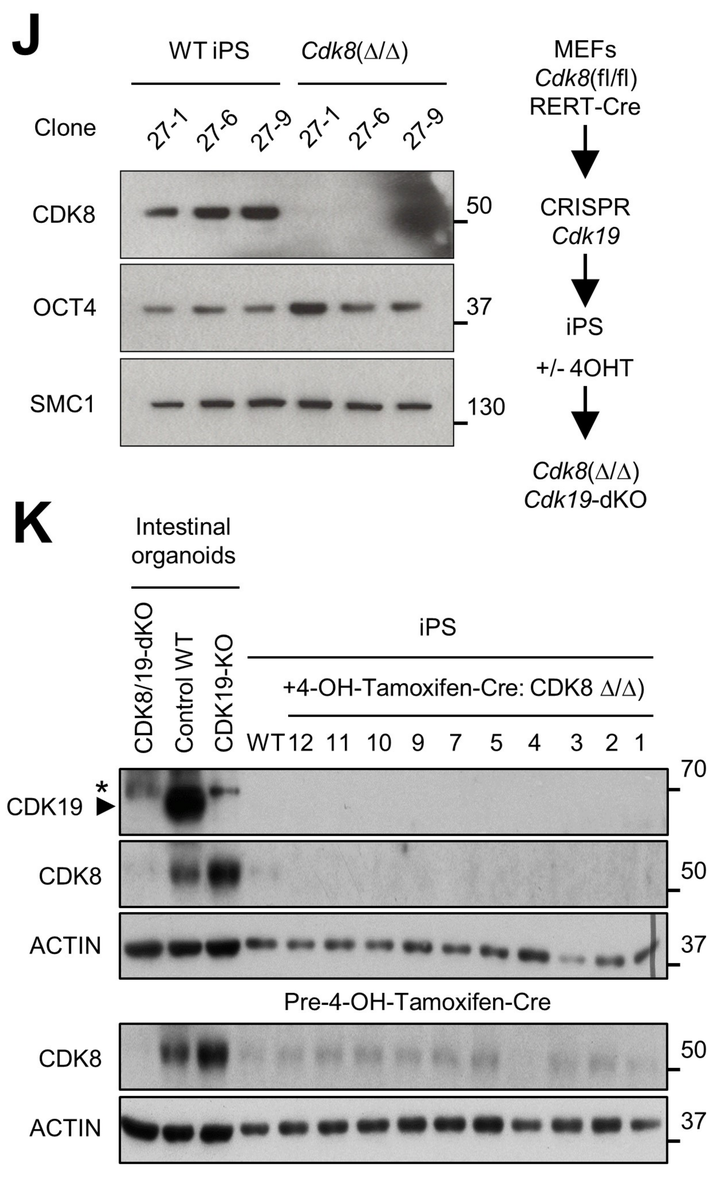


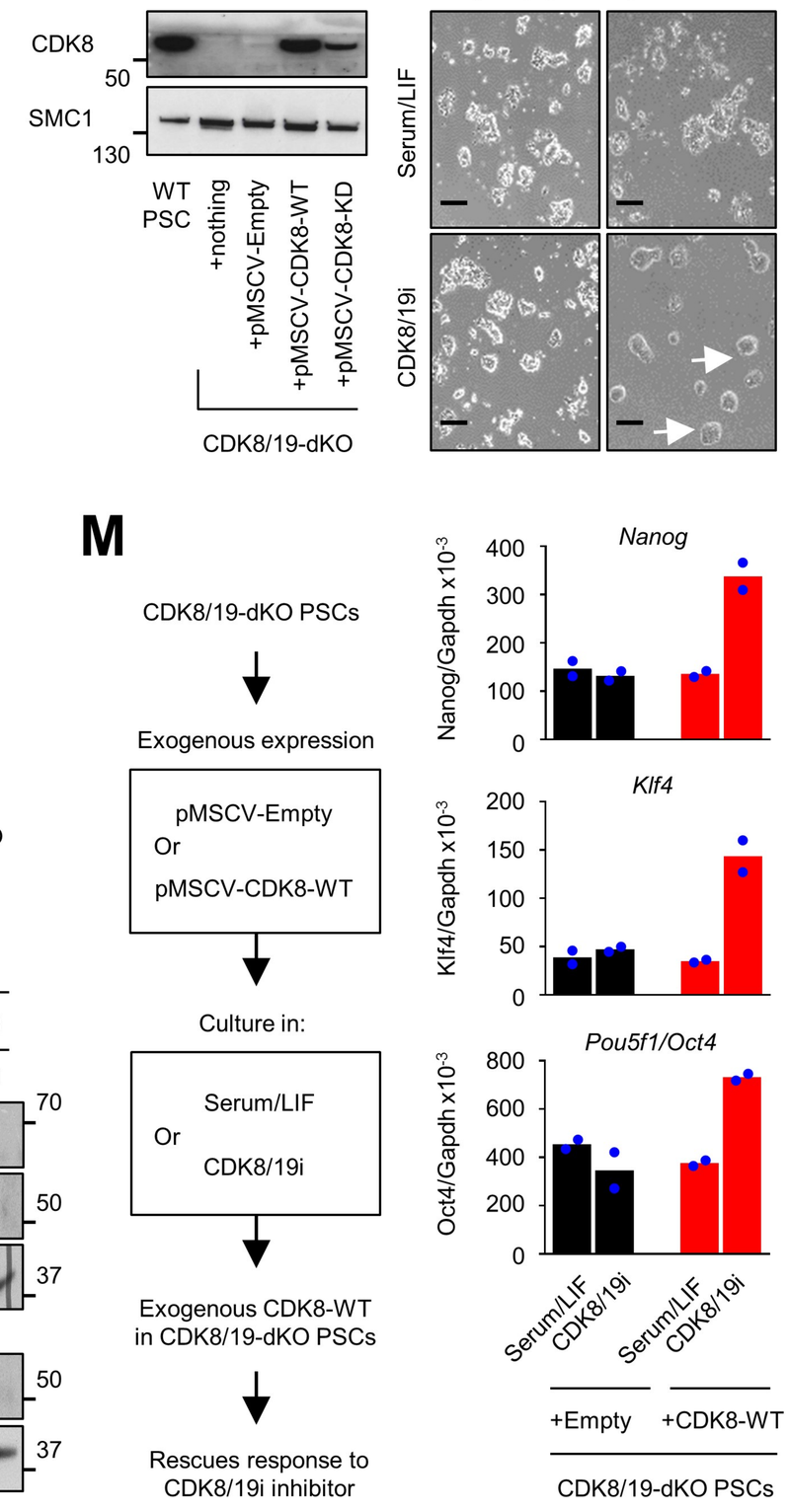




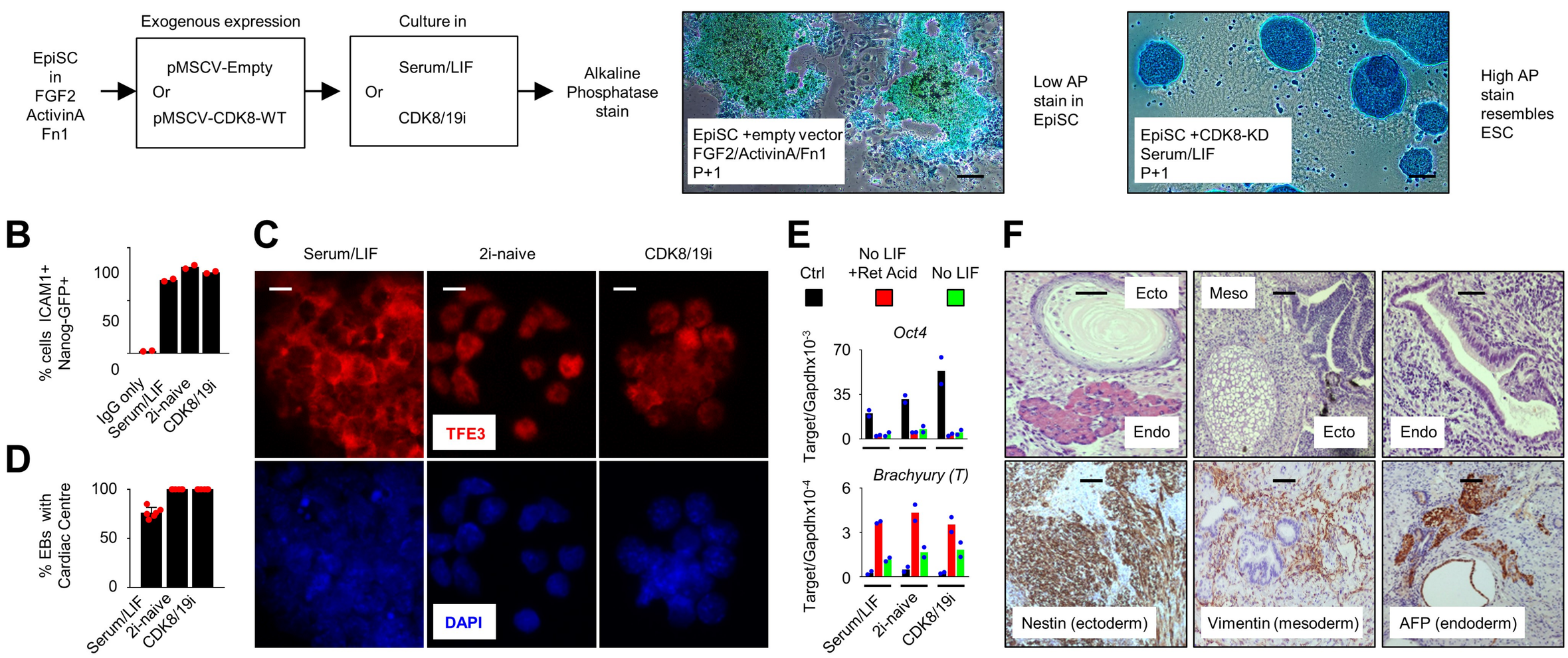




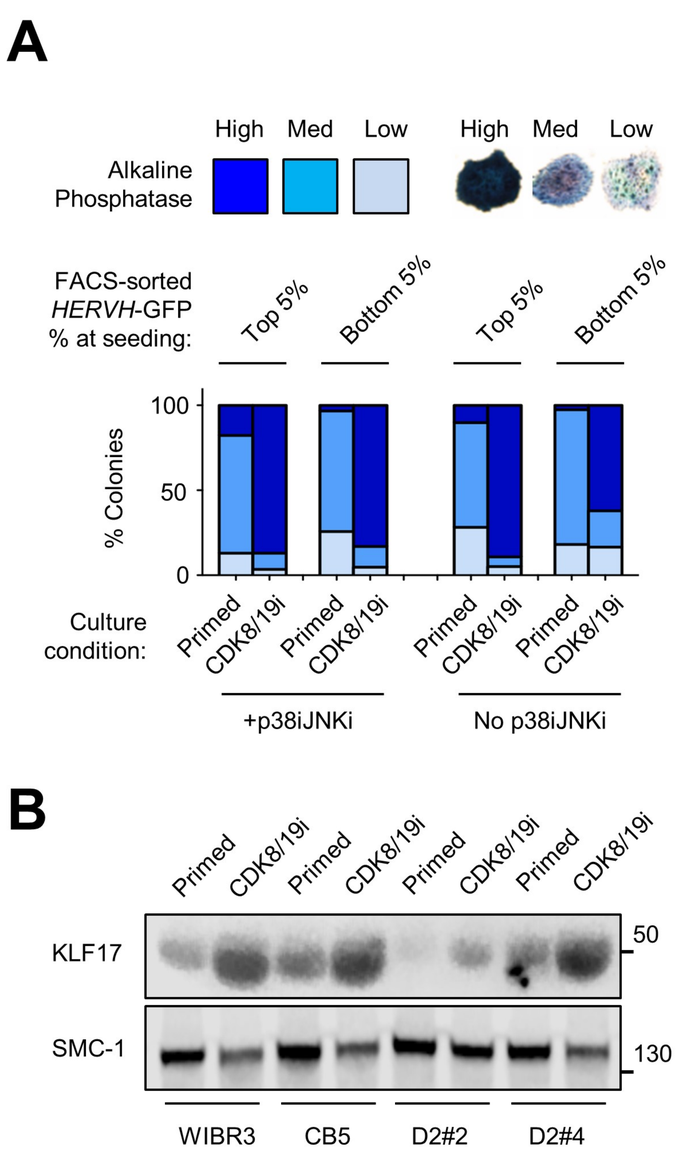


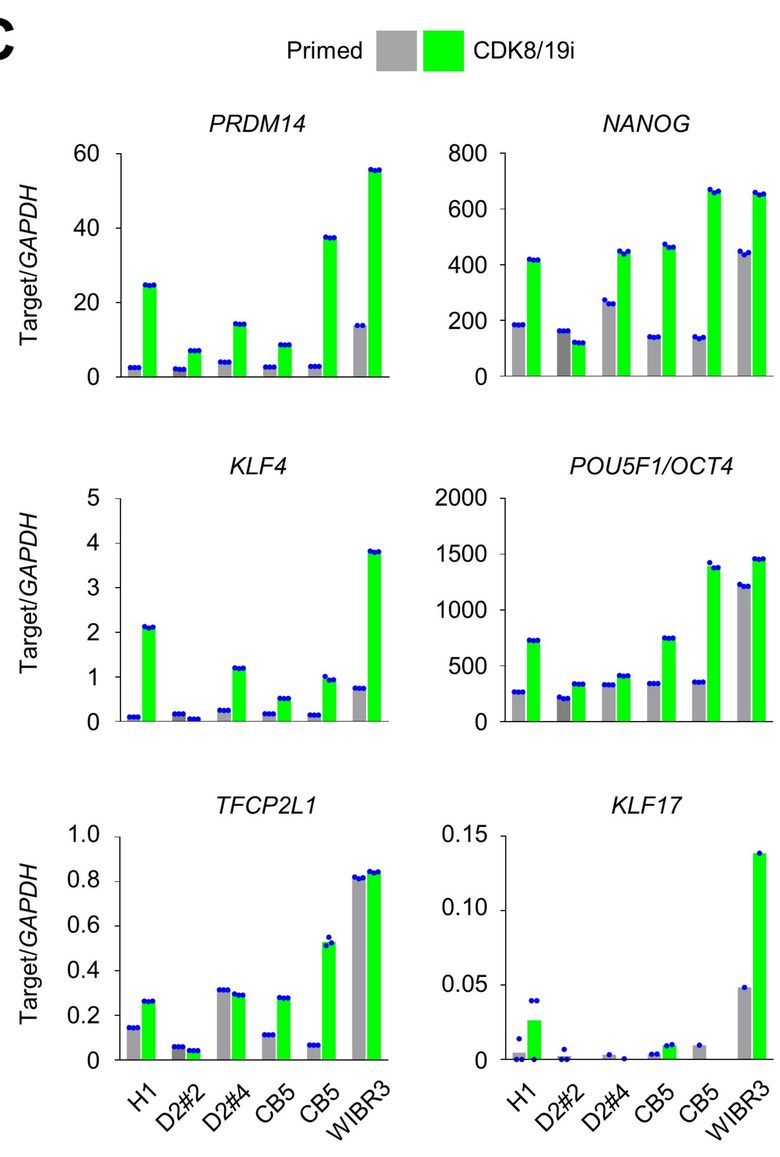


Extended Data Fig. 2



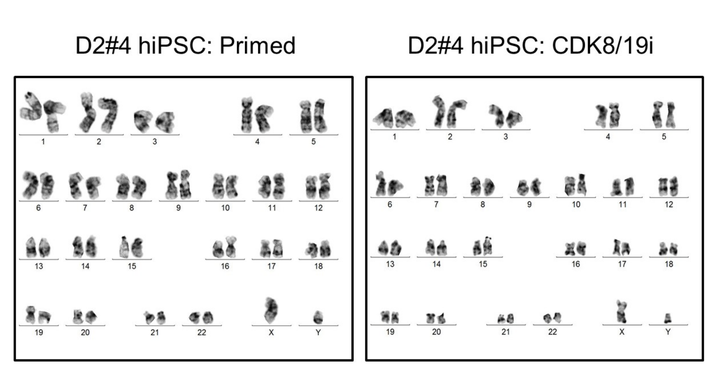
Extended Data Fig. 3

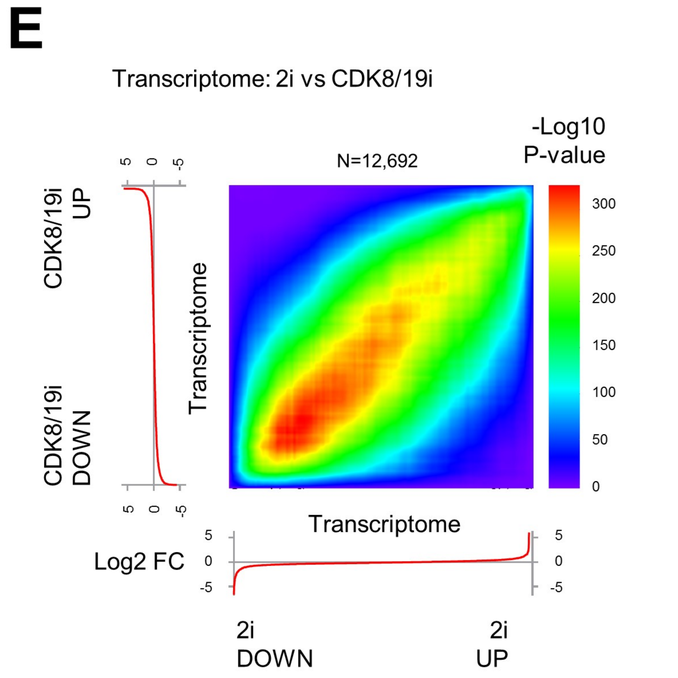


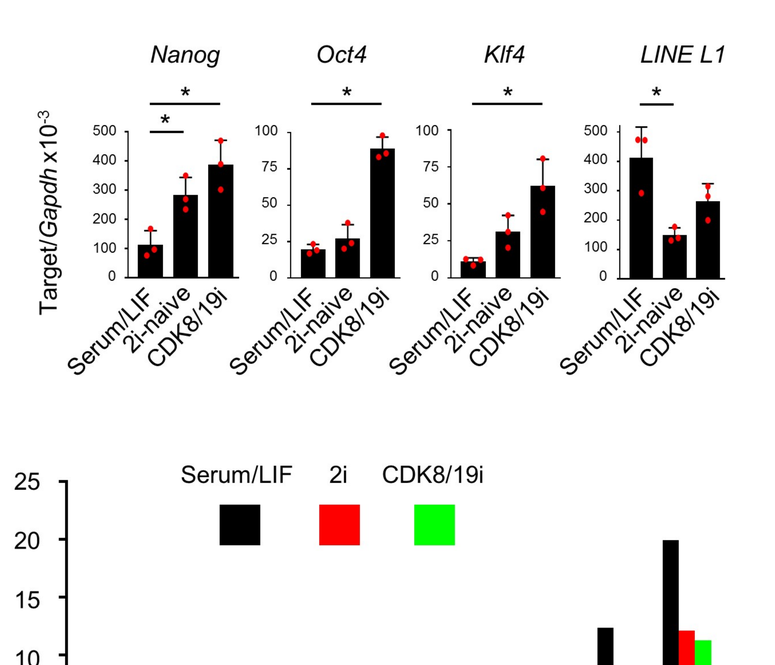


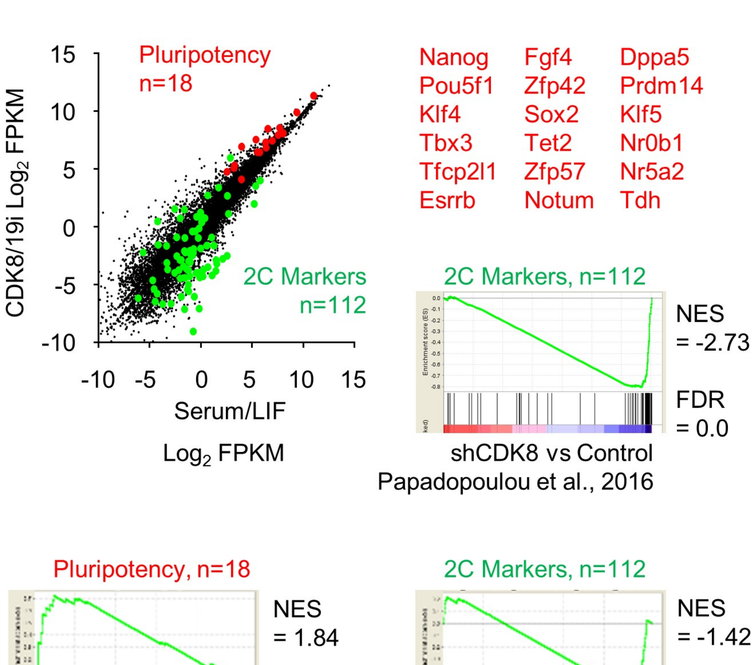
D

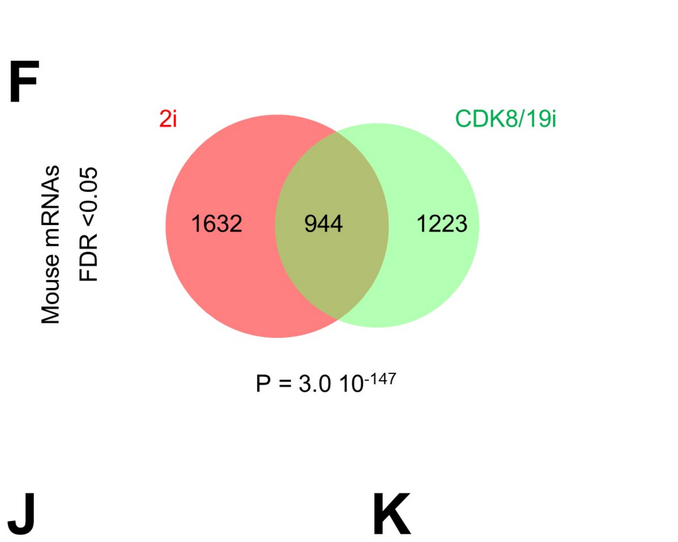
	Line	Treatment	Passage	Karyotype
Human	D2#4 hiPSC	Primed	P+19	46 XY
		CDK8/19i	P+19	46 XY
	H1 hESC	Primed	P+16	46 XY
		CDK8/19i	P+16	46 XY
	D2#2 hiPSC	Primed	P+18	46 XY
		CDK8/19i	P+18	46 XY
	WIBR3 hESC	Primed	P+16	46 XX
		CDK8/19i	P+16	46 XX
	CB5 hiPSC	Primed	P+18	46 XX
		CDK8/19i	P+17	46 XX
Mouse	TON mESC	S/L Primed	P+10	40 XY
		2i-naive	P+10	40 XY
		CDK8/19i	P+10	40 XY

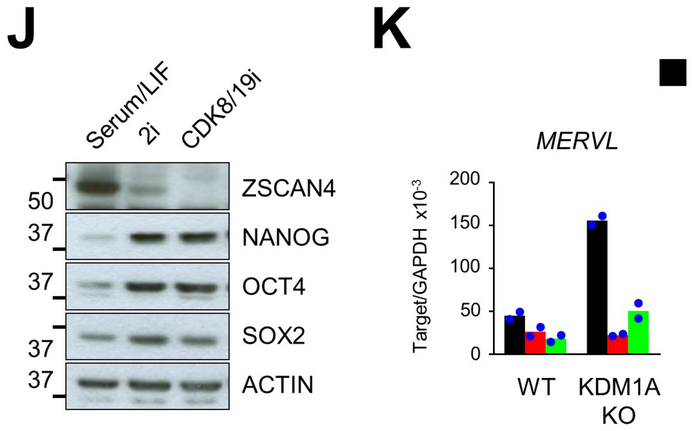


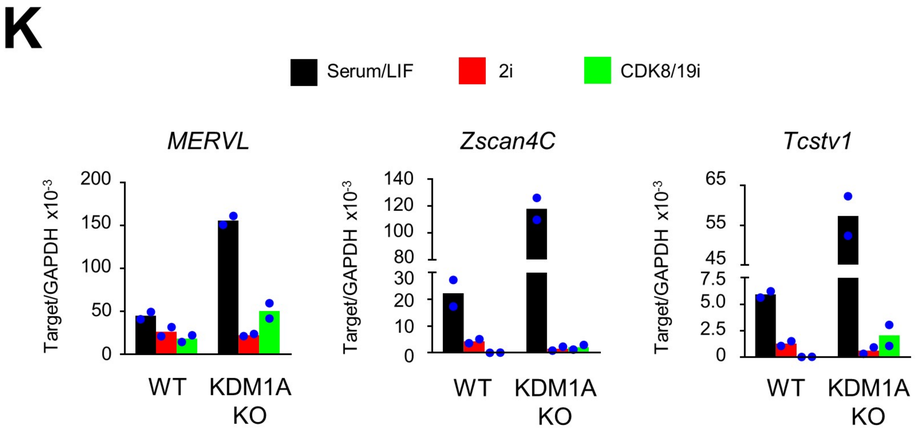


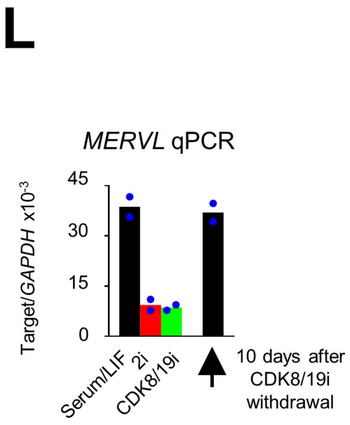


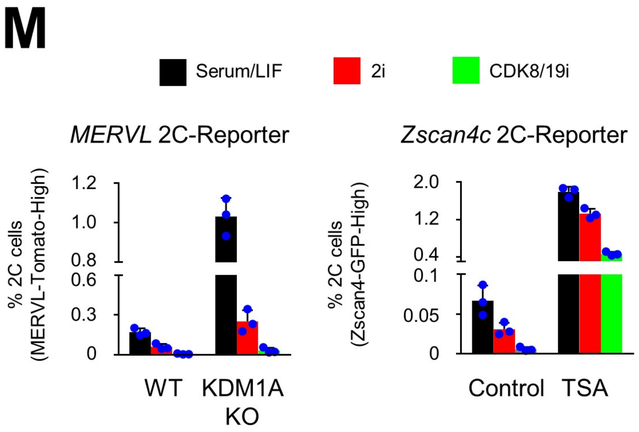




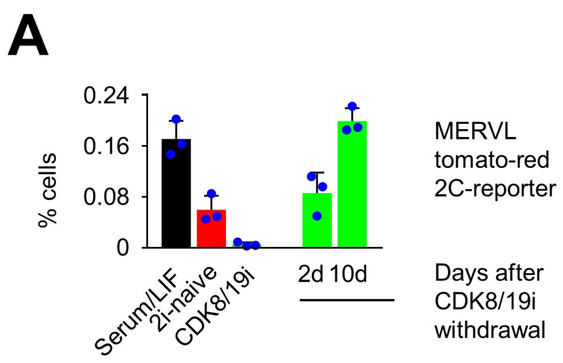


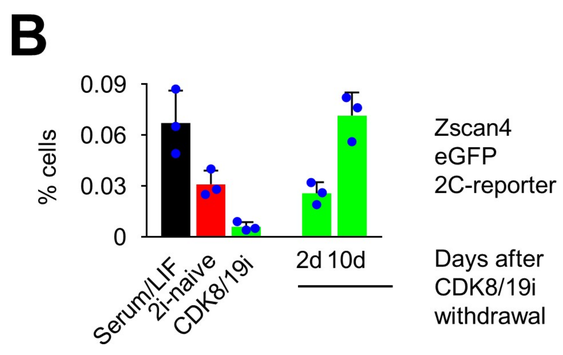


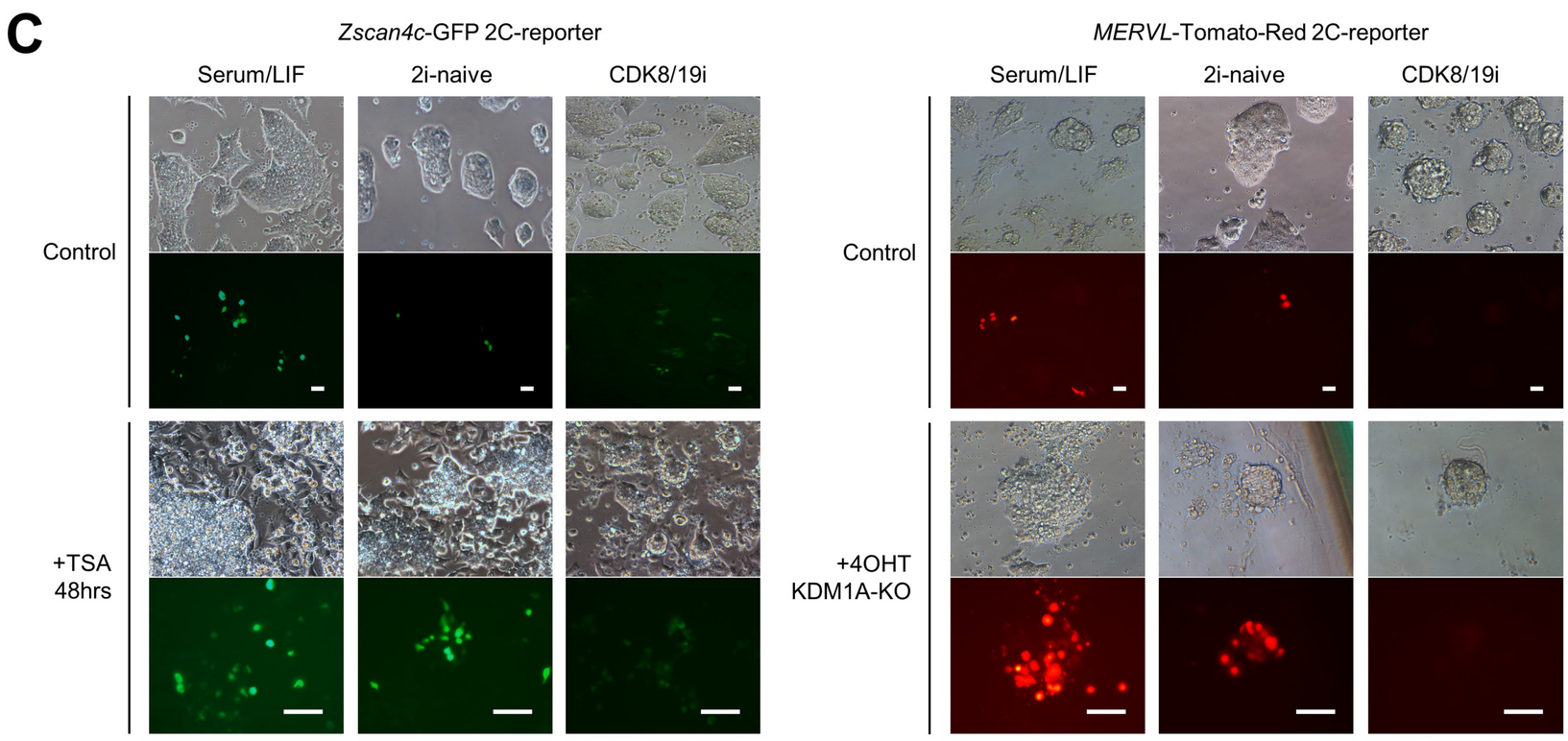




Extended Data Fig. 4







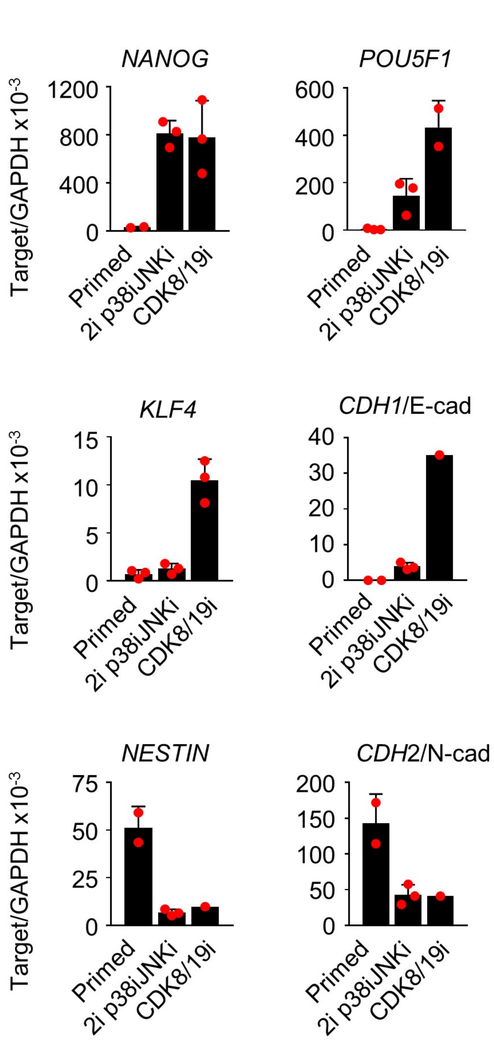
D

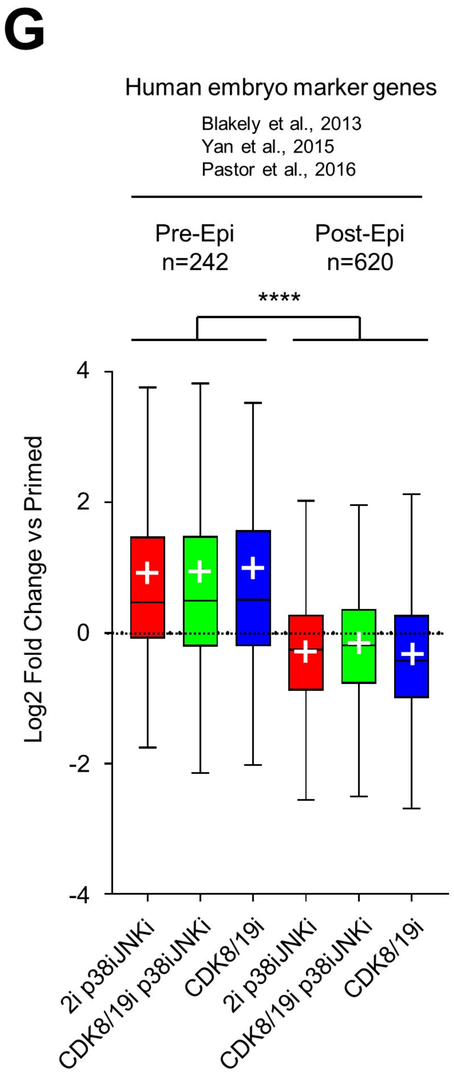
	Up 2i (645)	Up CDK8i (875)	Down 2i (1157)	Down CDK8i (822)
Marks et al., 2012	275 1.747e-123	164 2.180e-20	247 8.257e-265	250 8.230e-55
Kolodziejczyk et al., 2015	189 3.697e-79	103 1.937e-10	172 3.038e-83	94 2.0e-32
Fidalgo et al., 2015	268 1.565e-152	162 2.860e-34	567 5.278e-127	110 5.20e-31

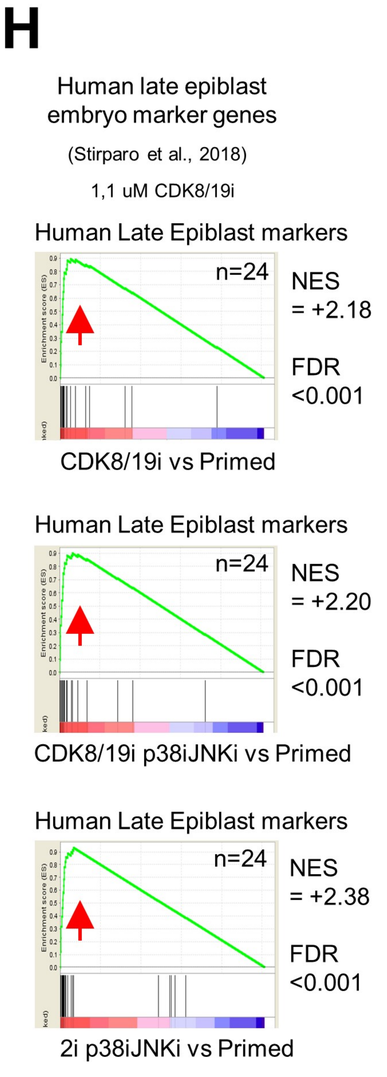
E

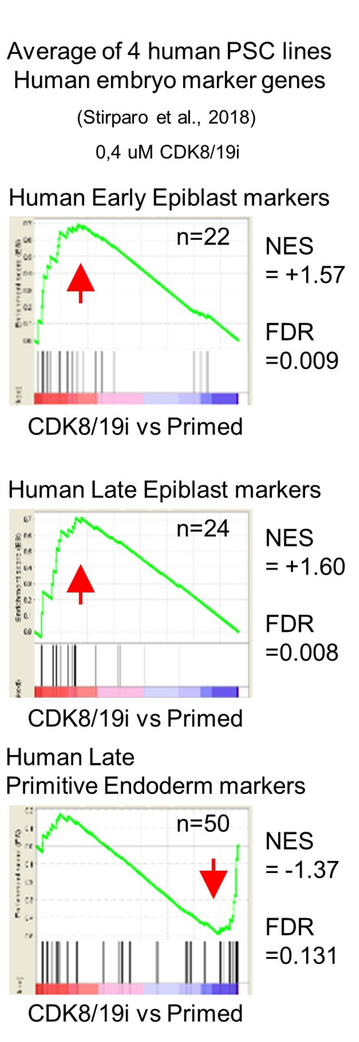
		Boroviak et al., 2015 Morula (446)	ICM (335)	Pre EPI (128)	Post EPI (207)
Naïve state	2i-CNIO (504)	27 0.001	**40** **4.385e-13**	**21** **5.560e-10**	8 0.332
	8i-CNIO (613)	27 0.014	**43** **3.337e-62**	**27** **3.465e-13**	7 -
	2i Marks et al. (1051)	44 0.005	**60** **9.540e-13**	25 8.069e-07	7 -
	2i Kolodziejczyk et al. (802)	44 1.546e-05	**60** **4.451e-18**	**30** **9.219e-13**	5 -
	Fidalgo et al. (759)	47 2.645e-07	**59** **1.438e-18**	21 6.442e-07	4 -

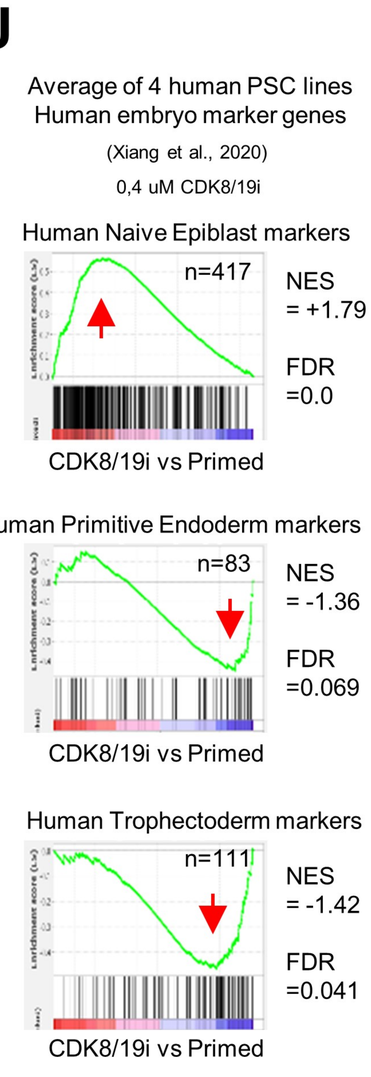
		Boroviak et al., 2015 Morula (446)	ICM (335)	Pre EPI (128)	Post EPI (207)
Primed state	2i-CNIO (926)	38 0.012	24 0.169	9 0.327	**72** **2.789e-37**
	8i-CNIO (663)	30 0.007	24 0.007	4 -	**47** **4.125e-22**
	2i Marks et al. (1450)	43 0.366	29 **0.435**	10 -	**92** **1.154e-41**
	2i Kolodziejczyk et al. (379)	15 0.113	7 -	5 0.189	**56** **1.484e-43**
	2i Fidalgo et al. (538)	23 0.029	4 -	4 -	**44** **3.773e-23**





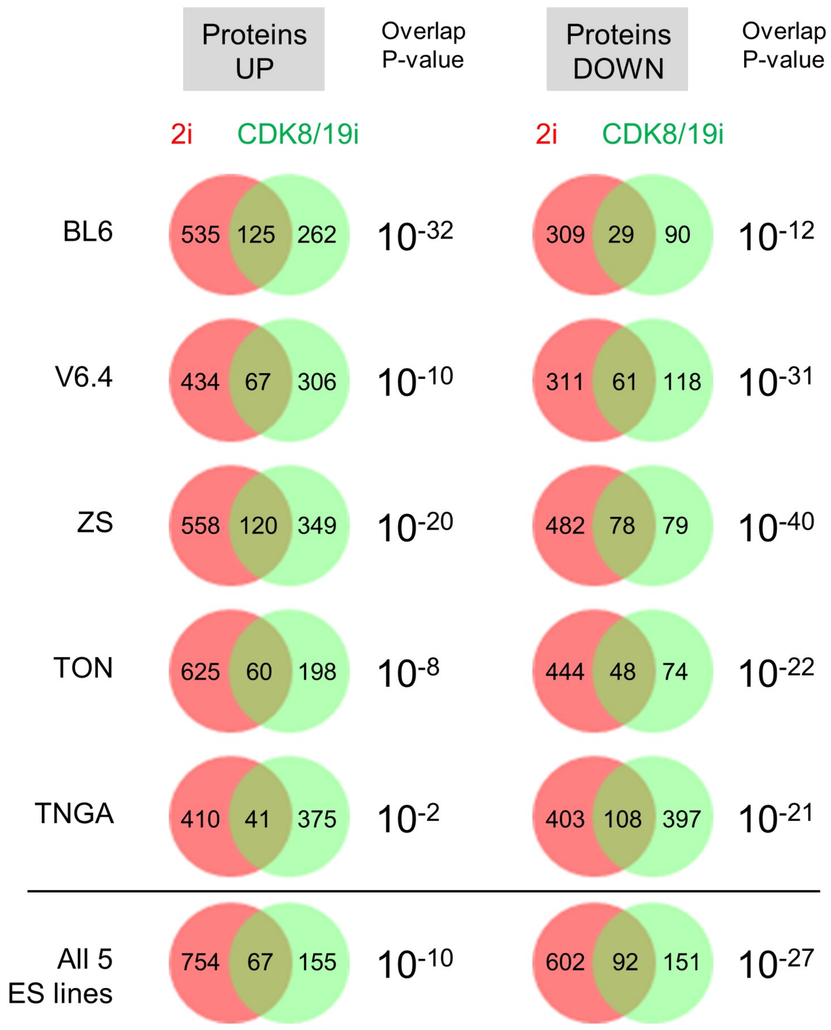






Extended Data Fig. 5

A



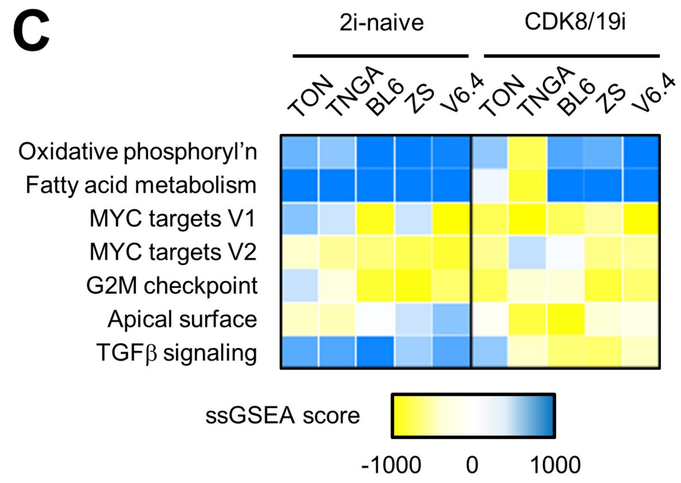
		UP Cdk8i	DOWN Cdk8i
All 5 cell lines	UP 2i	67	30
	DOWN 2i	25	92
		UP Cdk8i	DOWN Cdk8i
BL6	UP 2i	125	11
	DOWN 2i	12	29
		UP Cdk8i	DOWN Cdk8i
V6.4	UP 2i	67	3
	DOWN 2i	13	61
		UP Cdk8i	DOWN Cdk8i
ZS	UP 2i	120	5
	DOWN 2i	18	78
		UP Cdk8i	DOWN Cdk8i
TON	UP 2i	60	8
	DOWN 2i	16	48
		UP Cdk8i	DOWN Cdk8i
TNGA	UP 2i	41	44
	DOWN 2i	35	107

B

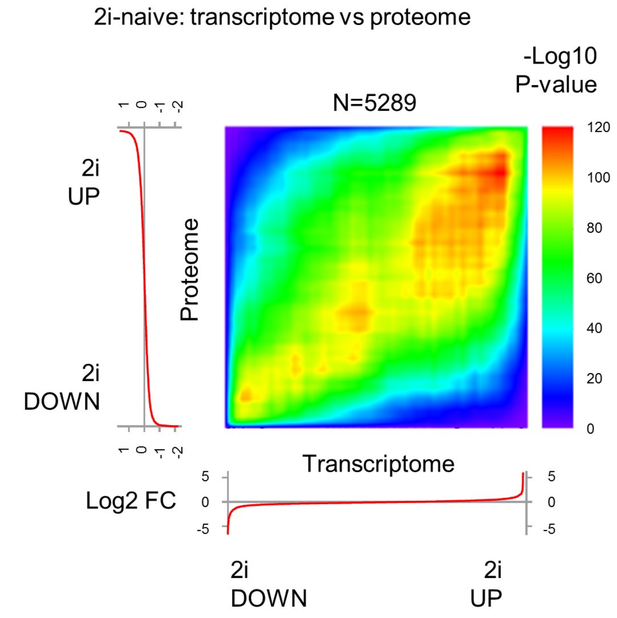
Differentially expressed proteins

	UP		DOWN	
	2i	CDK8i	2i	CDK8i
5mESCs	821	222	694	243
TNGA	451	416	510	504
ZS	678	469	560	157
V64	501	373	372	179
BL6	660	387	338	119
TON	685	258	492	122

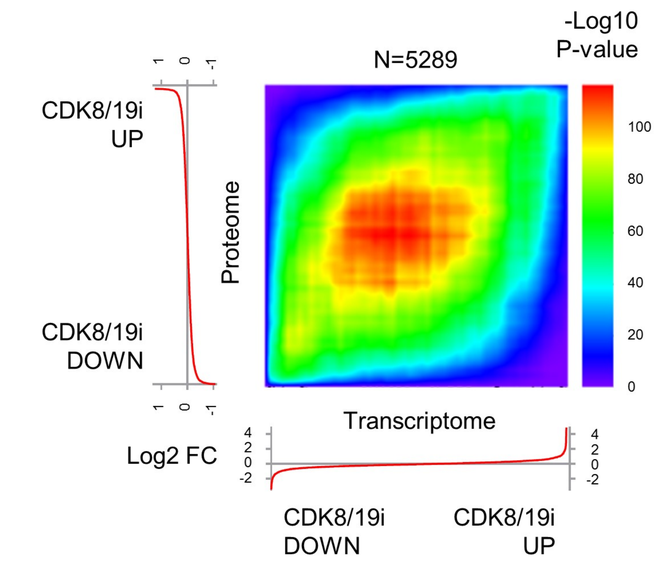
C



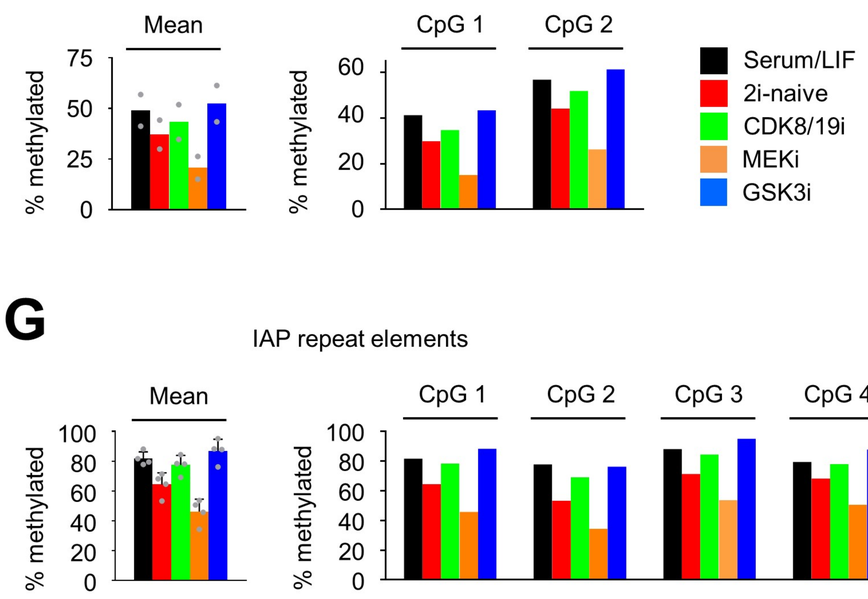
D



E

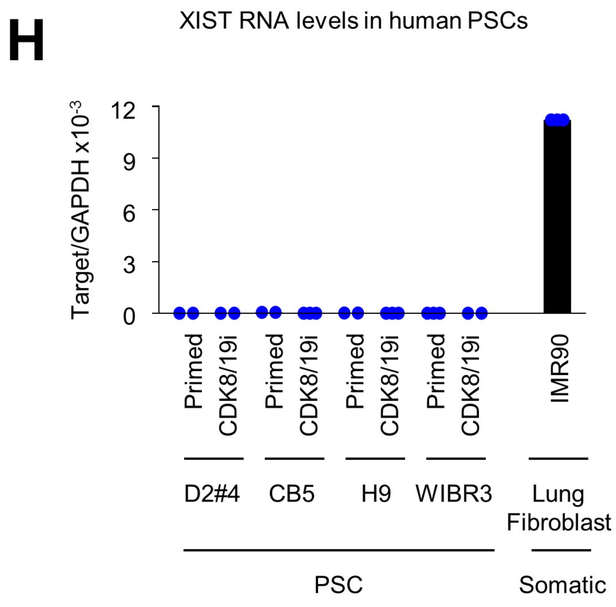


F

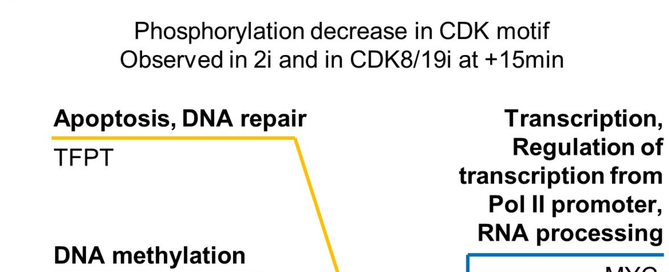


G

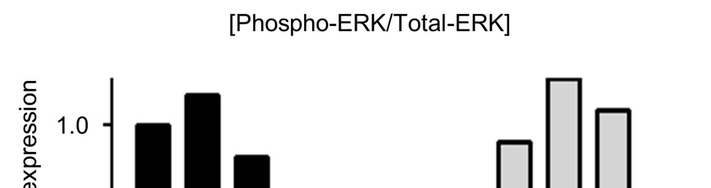
H



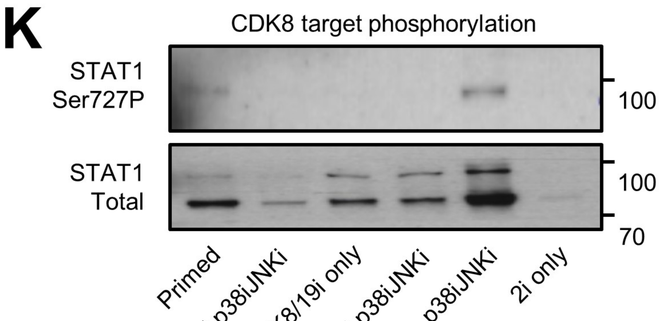
I



J

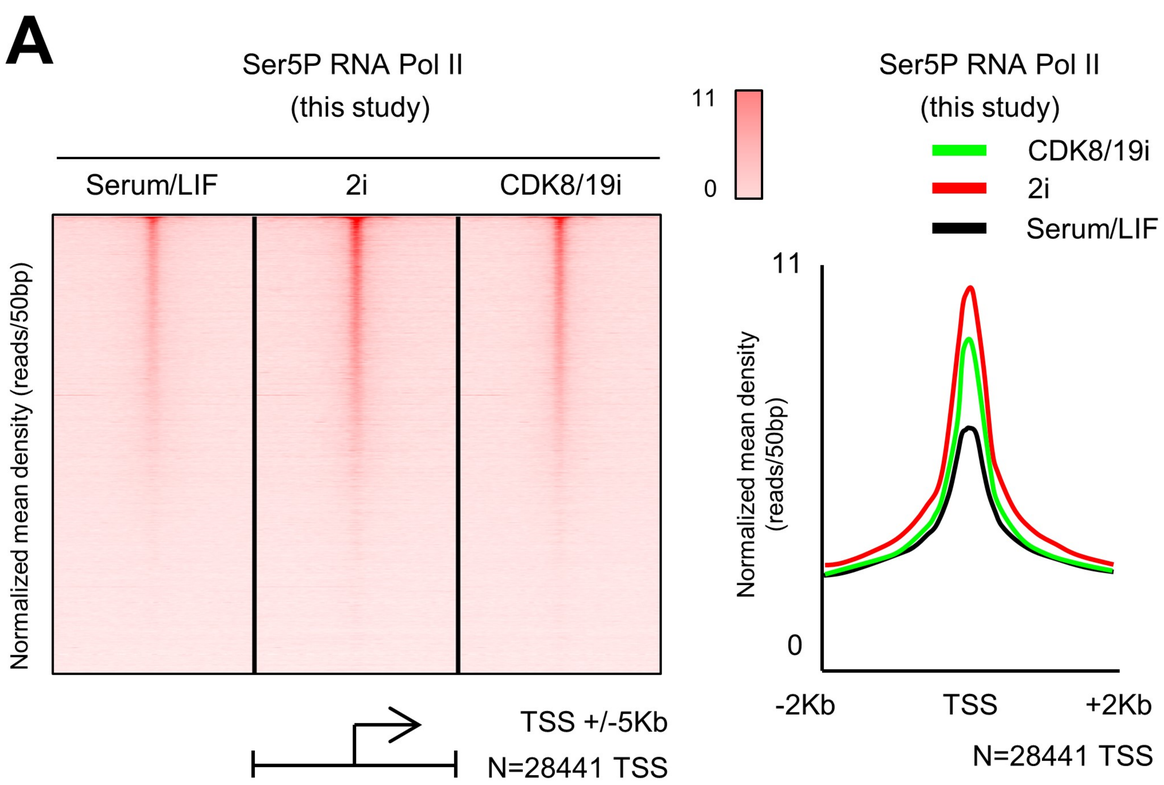


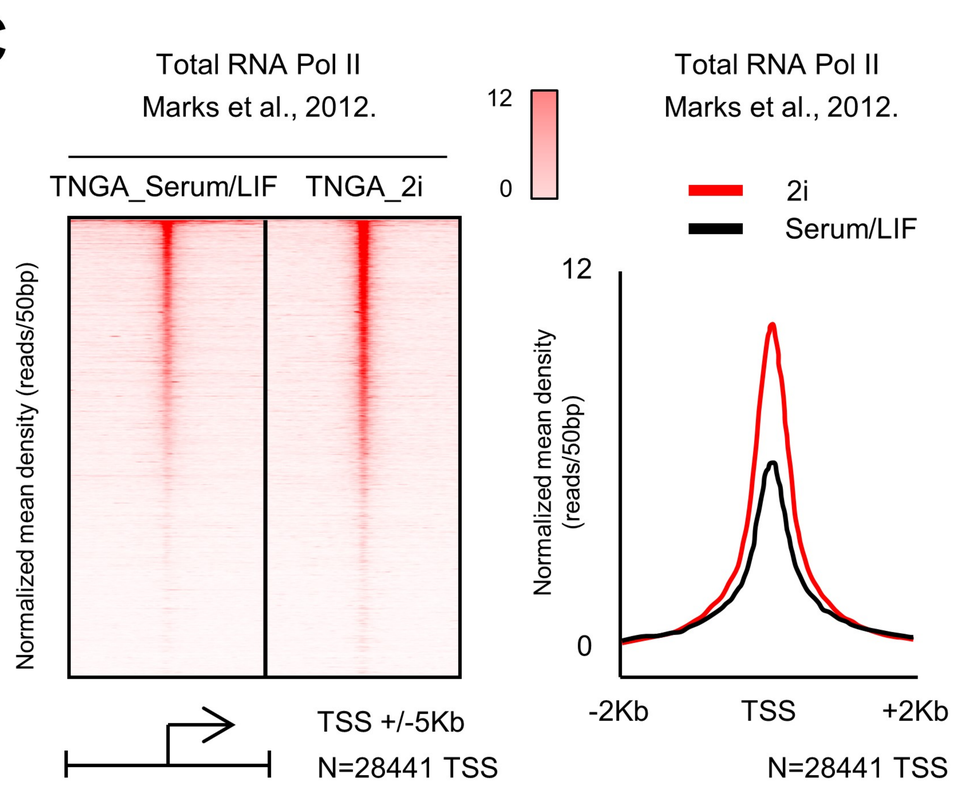
K

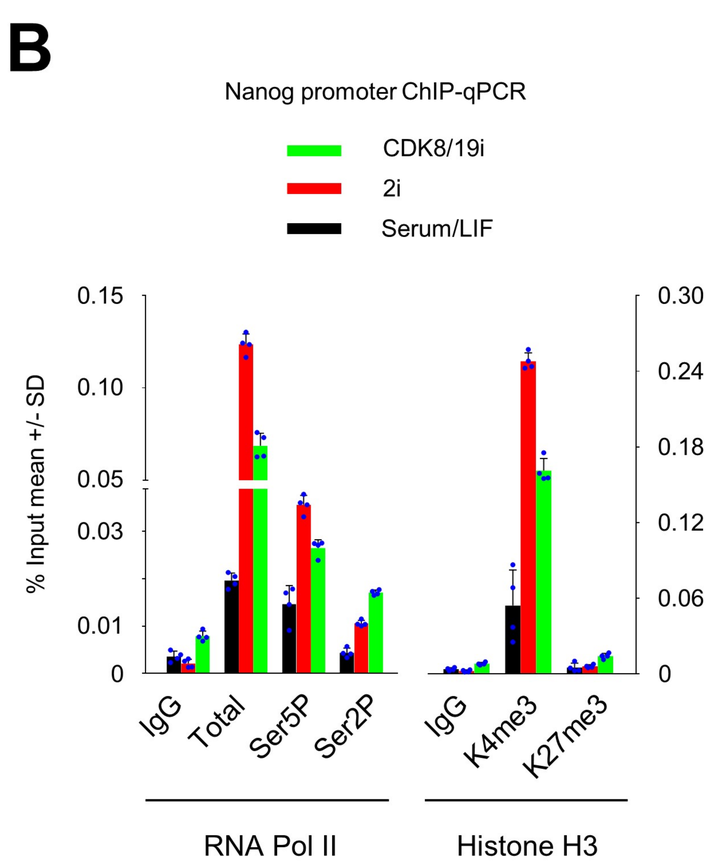


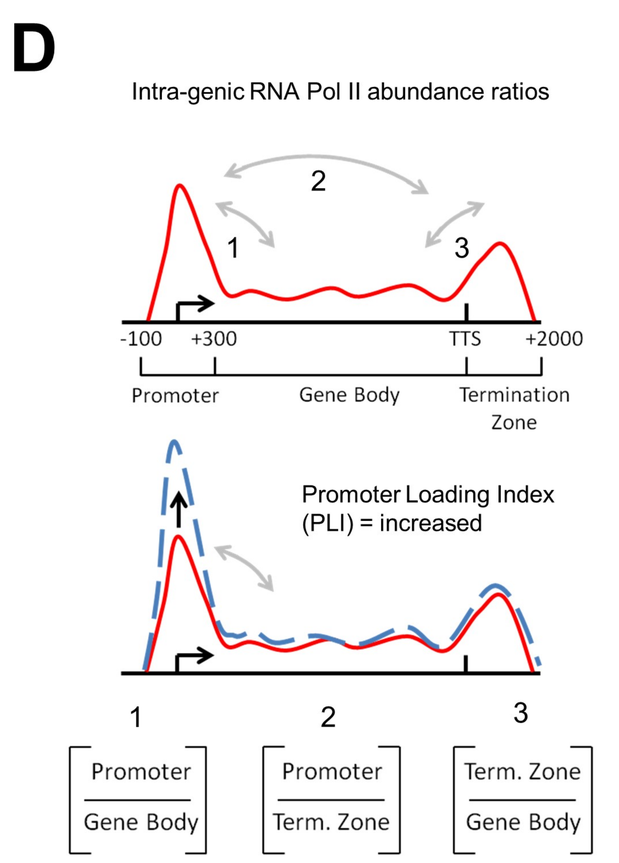
L

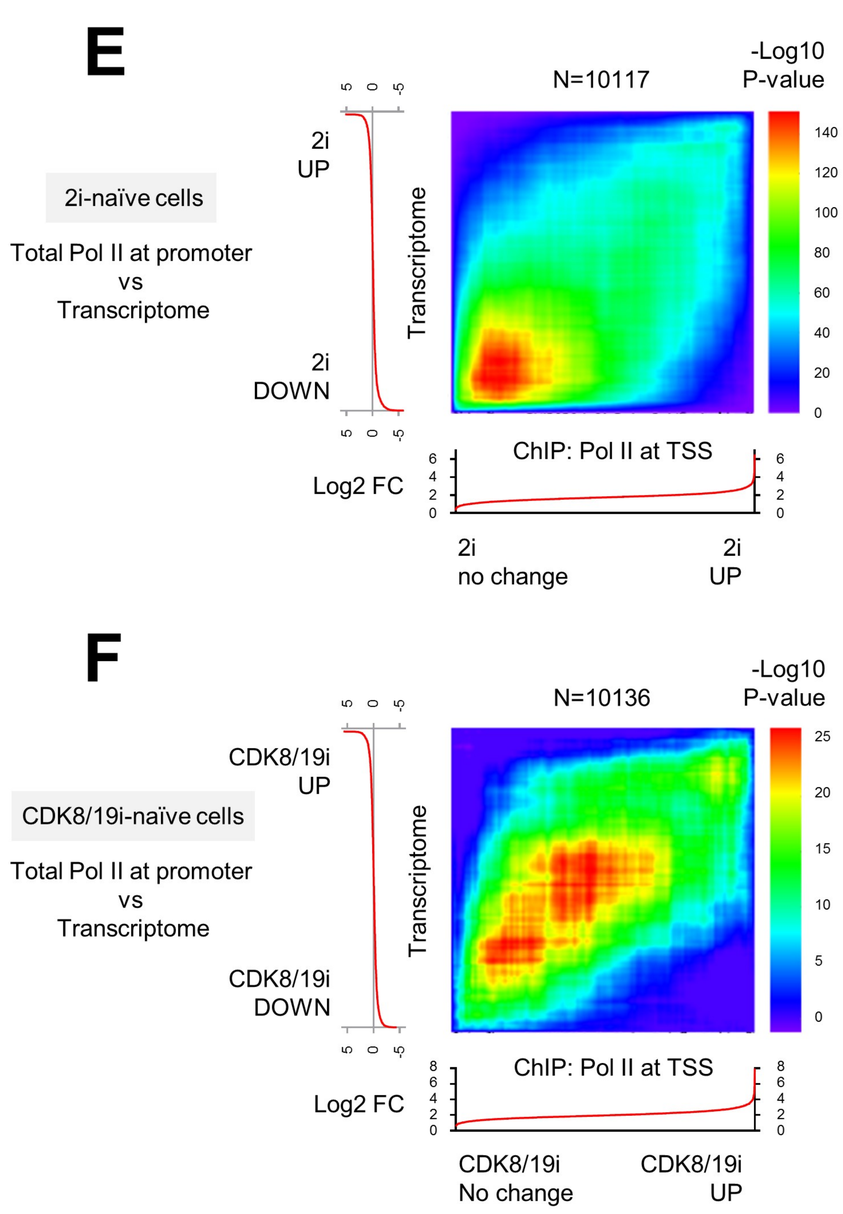
Extended Data Fig. 6

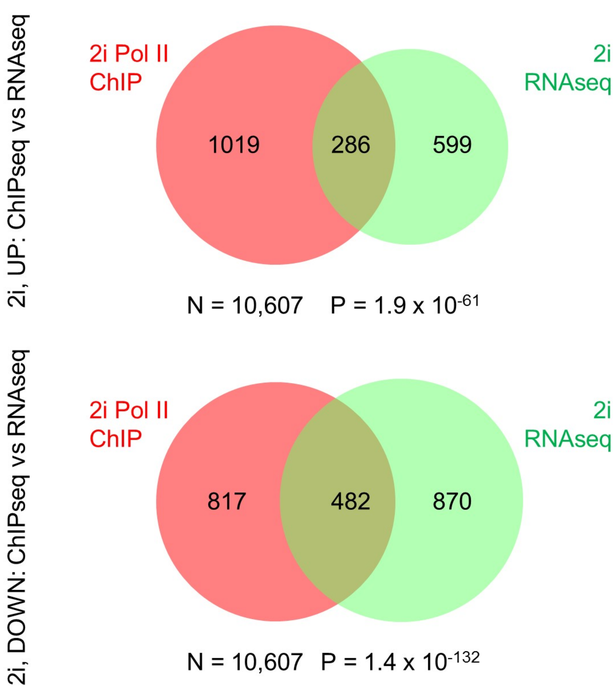


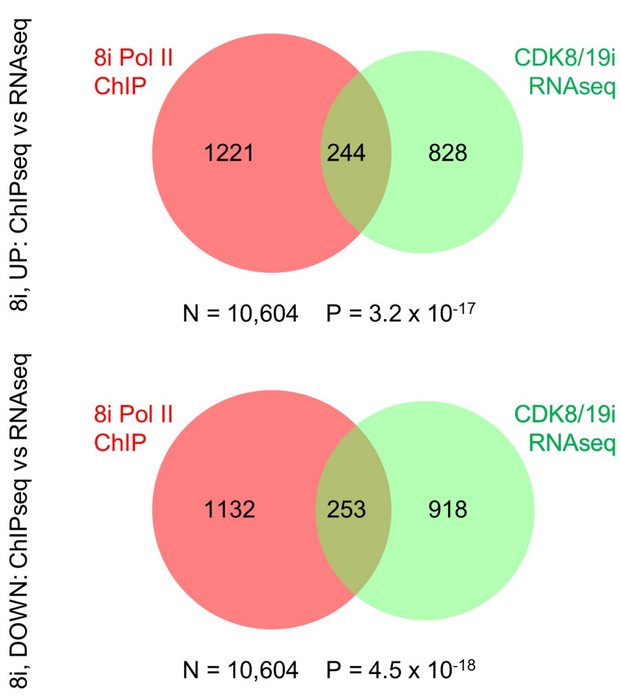


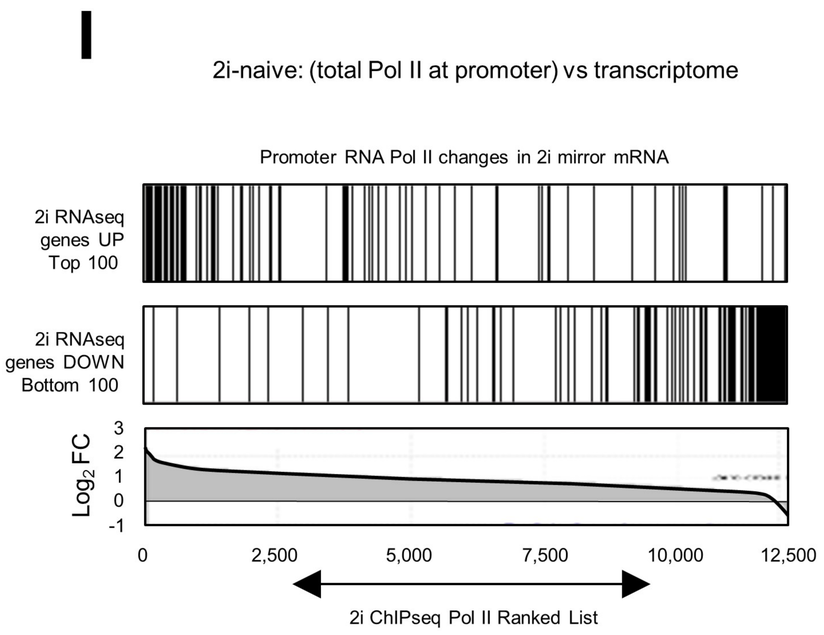




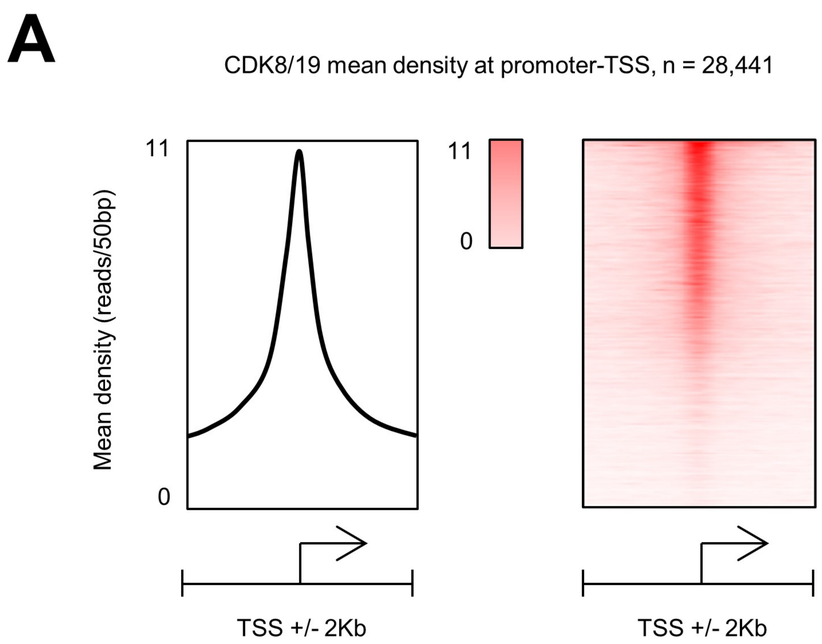


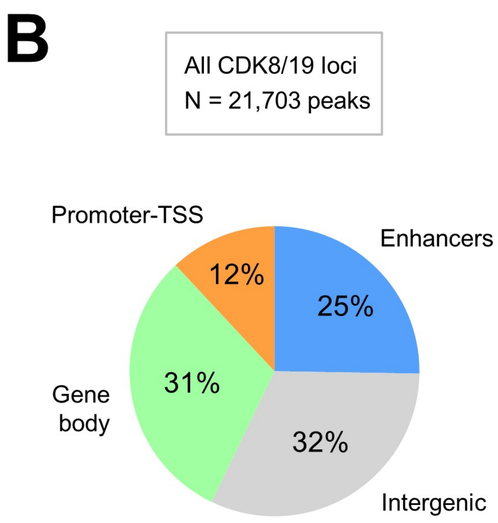


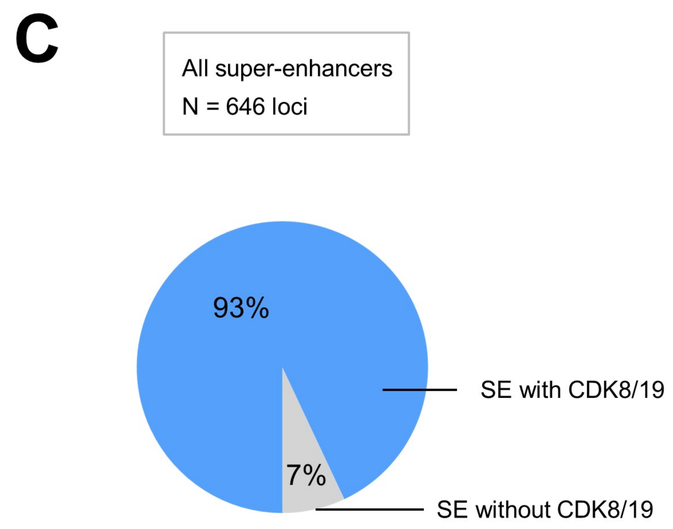


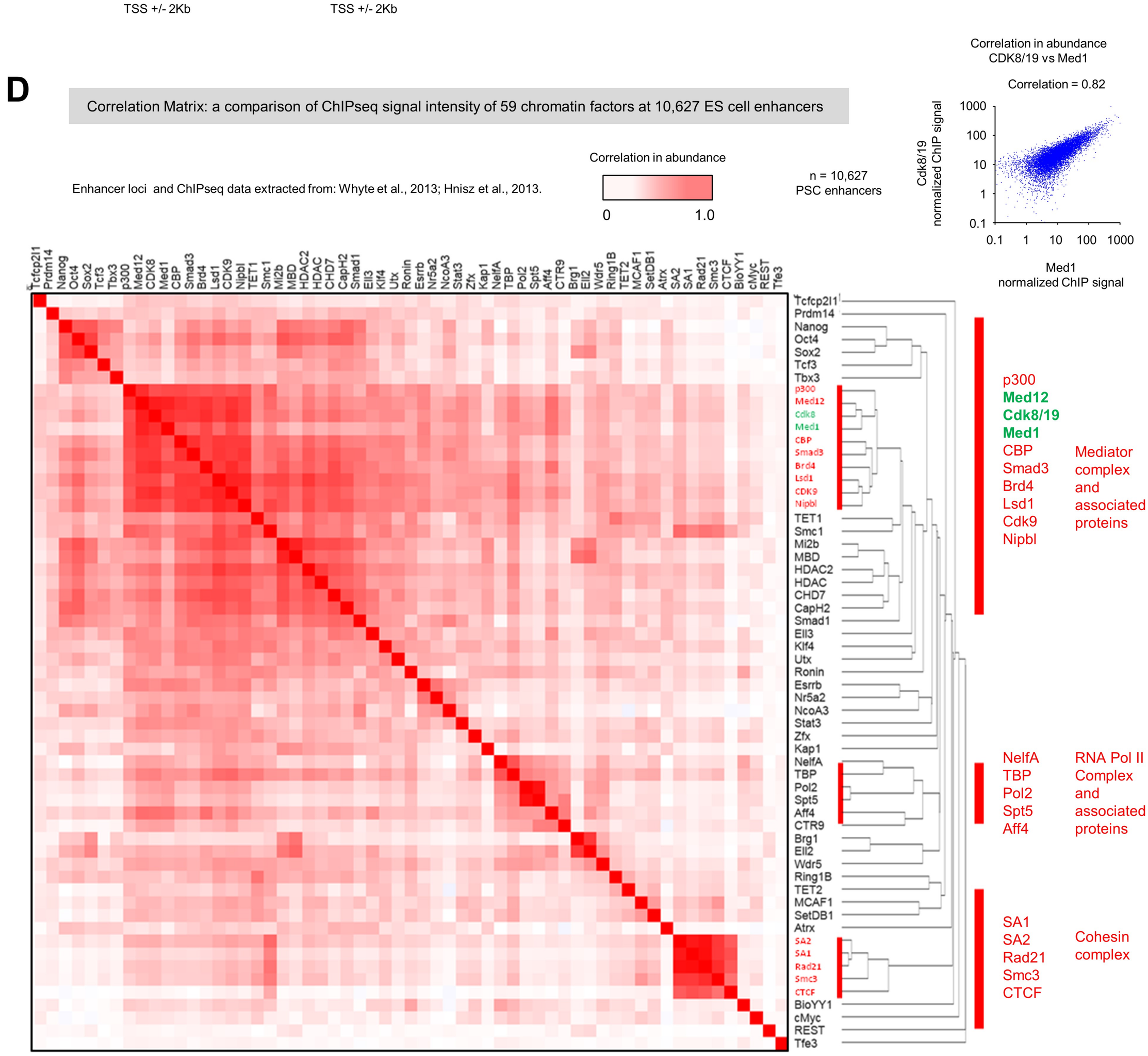


Extended Data Fig. 7

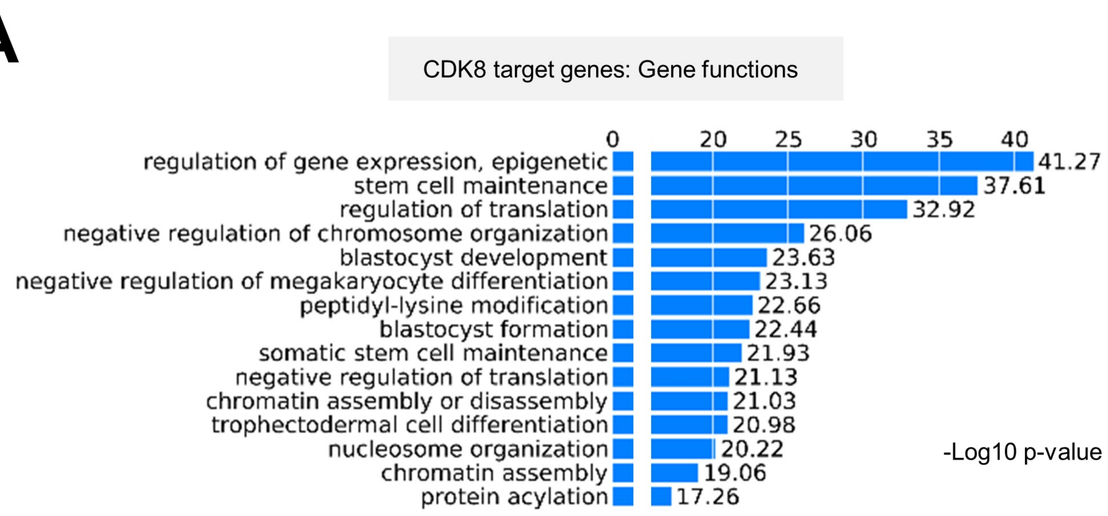


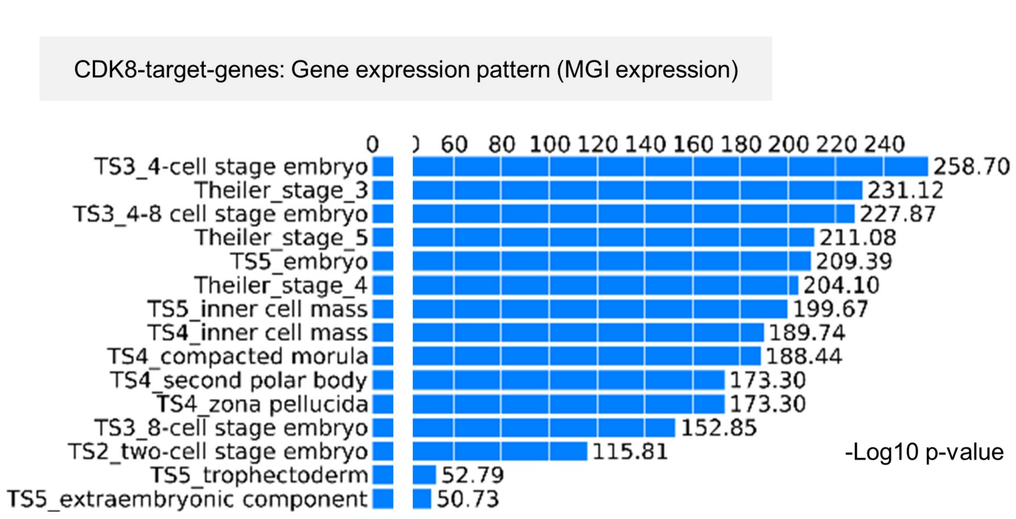


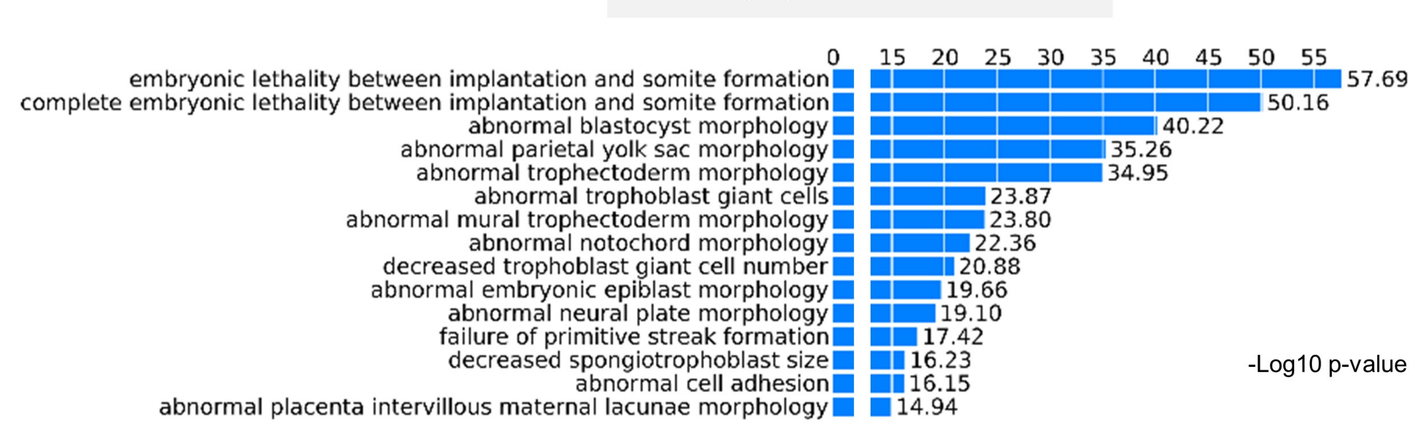


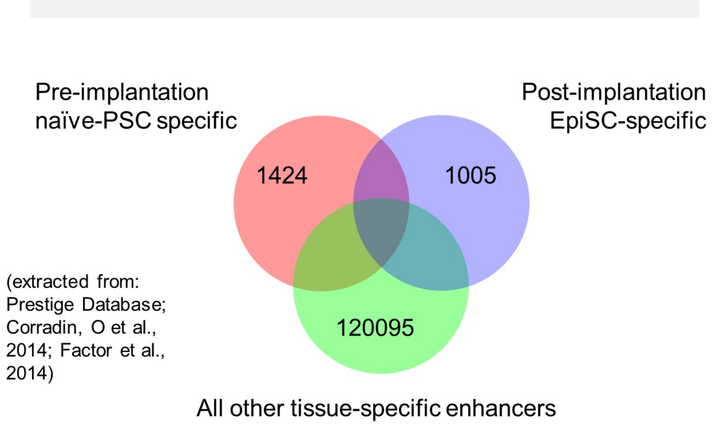


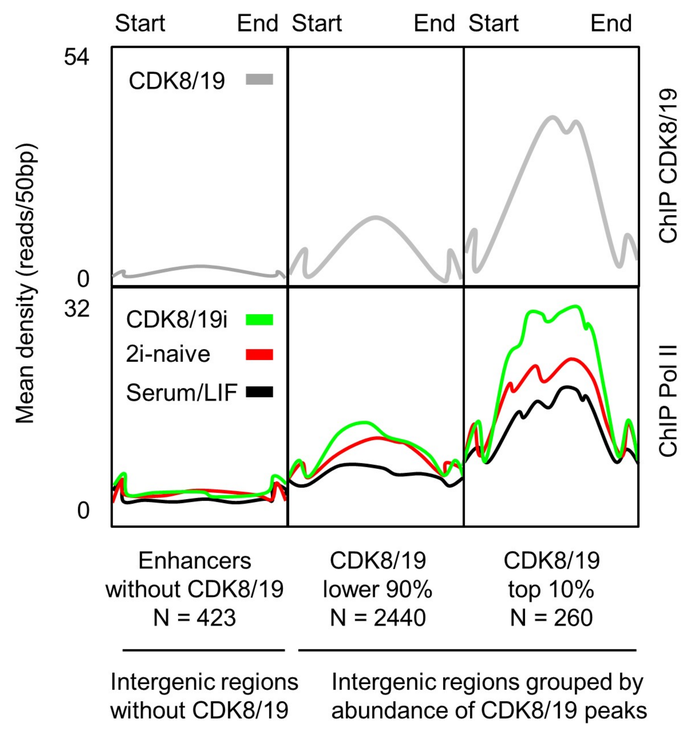
Extended Data Fig. 8

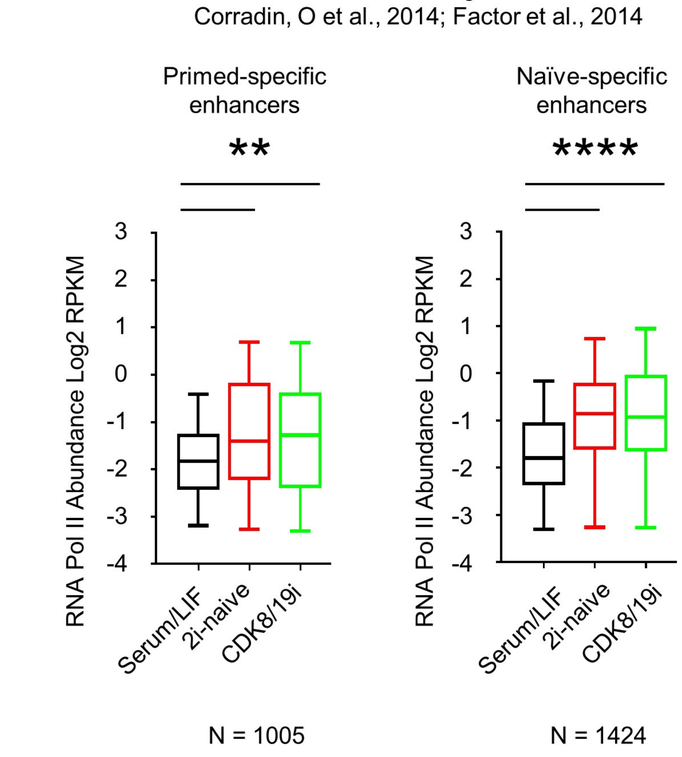


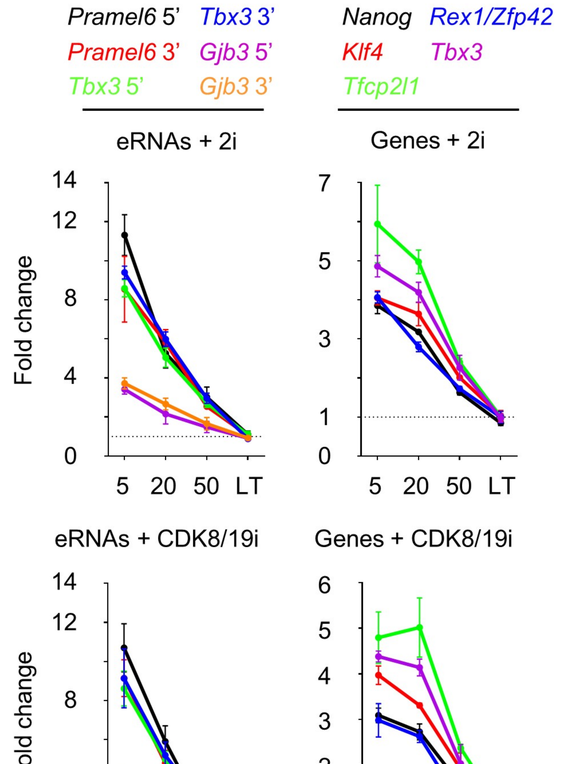


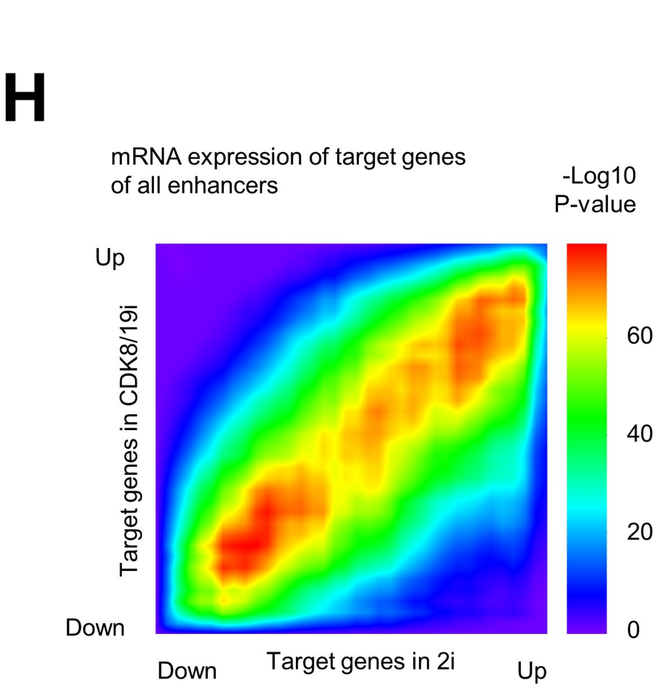






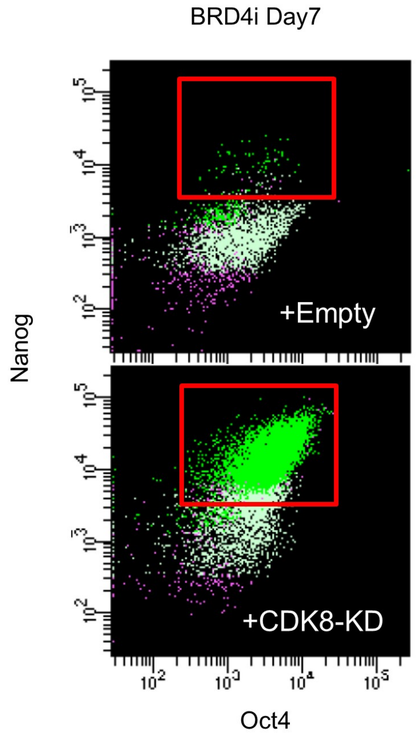


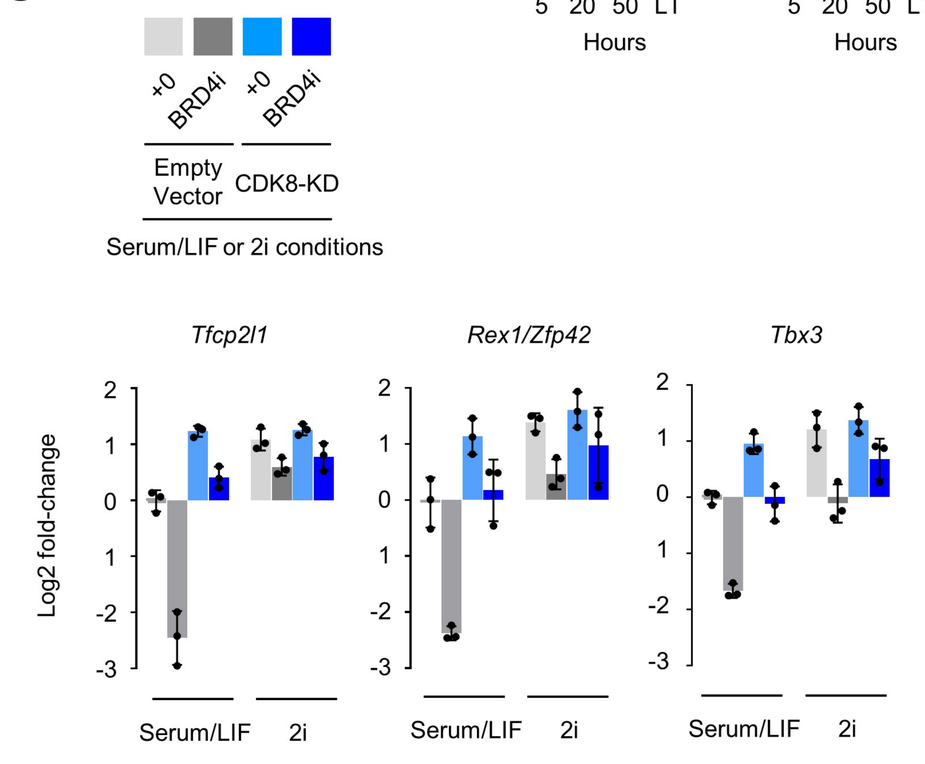




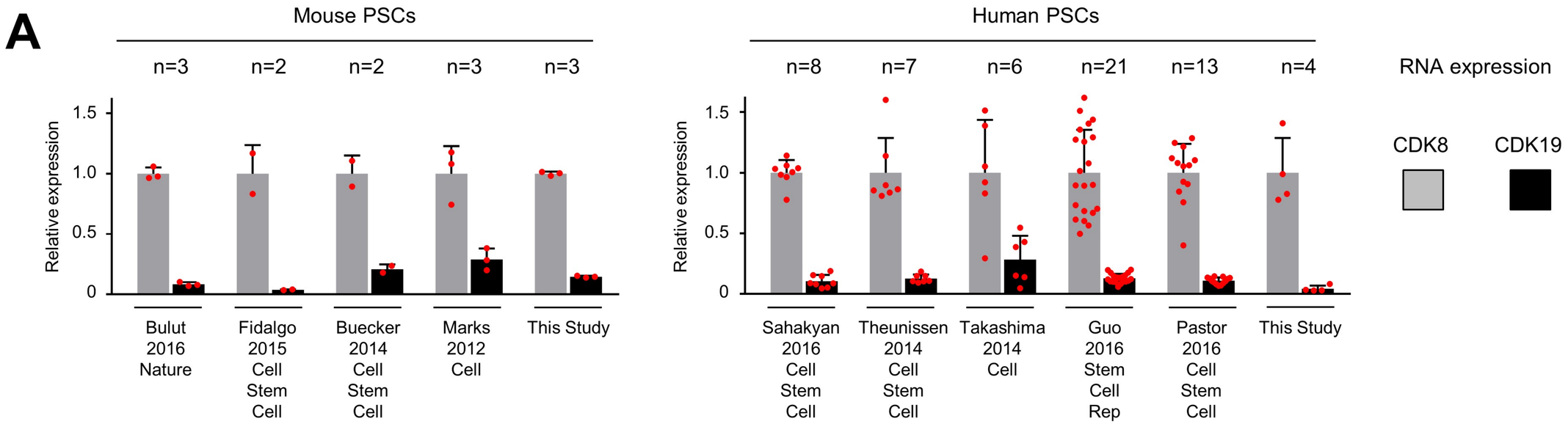
All PSC enhancers
N = 10,627
Hnisz et al., 2013.

GREAT analysis
Single-nearest target gene
N = 3553 target mRNAs





Extended Data Fig. 9



Extended Data Fig. 10