SMAD2/3-SMYD2 and developmental transcription factors form positive feedback loops with cell cycle inhibitors to guide tissue formation

Stefania Militi
University of Oxford
Reshma Nibhani
University of Oxford
Martin Pook
University of Oxford
Siim Pauklin
siim.pauklin@ndorms.ox.ac.uk

University of Oxford https://orcid.org/0000-0001-8367-3670

Research Article

Keywords: Cancer stem cells, Pancreatic cancer, Human pluripotent stem cells, Cell Cycle, Epigenetics, TGFβ/ACTIVIN-SMAD2/3, Differentiation

Posted Date: February 28th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3994324/v1

Additional Declarations: The authors declare no competing interests.

SMAD2/3-SMYD2 and developmental transcription factors form positive feedback loops with cell cycle inhibitors to guide tissue formation

Stefania Militi[1], Reshma Nibhani[1], Martin Pook[1], Siim Pauklin[1,*]

1. Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Old Road, Headington, Oxford, OX3 7LD, UK.

* Correspondence: siim.pauklin@ndorms.ox.ac.uk

Abstract.

Tissue formation and organ homeostasis is achieved by precise coordination of proliferation and differentiation of stem cells and progenitors. While deregulation of these processes can result in degenerative disease or cancer, their molecular interplays remain unclear. Here we show that the switch of human pluripotent stem cell (hPSC) self-renewal to differentiation is associated with the induction of distinct Cyclin Dependent Kinase Inhibitors (CDKIs). In hPSCs, Activin/Nodal/TGFβ signalling maintains CDKIs in a poised state via SMAD2/3-NANOG-OCT4-EZH2-SNON transcriptional complex. Upon gradual differentiation, CDKIs are induced by successive transcriptional complexes between SMAD2/3-SMYD2 and developmental regulators such as EOMES, thereby lengthening the G1 phase. This, in turn, induces SMAD2/3 transcriptional activity by blocking its linker phosphorylation. Such SMAD2/3-CDKI positive feedback loops drive the exit from pluripotency and stepwise cell fate specification that could be harnessed for producing cells for therapeutic applications. Our study uncovers fundamental mechanisms how cell fate specification is interconnected to cell cycle dynamics and provides insight to autonomous circuitries governing tissue self-formation.

Running Title: Positive feedback loops direct stem cell differentiation.

Keywords: Cancer stem cells; Pancreatic cancer; Human pluripotent stem cells; Cell Cycle; Epigenetics; TGFβ/ACTIVIN-SMAD2/3; Differentiation

Highlights

- TGFβ/ACTIVIN-SMAD2/3-EZH2 cooperates with OCT4 and NANOG in self-renewing hPSCs to keep CDKIs in an epigenetically poised state for rapid activation.
- Upon the initiation of endoderm differentiation SMAD2/3 switches its binding to SMYD2 and EOMES to induce CDKI expression.
- CDKIs lengthen G1 phase and thereby increase the time of hyperactivated ACTIVIN/NODAL-SMAD2/3 that drives cell specification.
- SMAD2/3-CDKI and TF circuitries can be harnessed for guiding tissue self-formation for biomedical applications.

Introduction.

Coordination of stem cell proliferation and differentiation is essential for normal development, organ homeostasis and tissue repair. A direct interplay between cell cycle progression and differentiation has been suggested in somatic stem cells in the skin, brain, gut and hematopoietic system [1-3]. Of particular interest, differentiation is often associated with a change in cell cycle length [4,5], suggesting that mechanisms controlling cell cycle progression could be involved in cell fate decisions.

The progression of cell cycle in mammalian cells is primarily controlled by Cyclins and Cyclin Dependent Kinases (CDKs), which affect key transcriptional regulators such as Retinoblastoma protein (pRb). The activity of the Cyclin D/CDK complexes and thus cell proliferation is constrained by Cyclin Dependent Kinase Inhibitors (CDKIs), which are subdivided into two families. The INK4 proteins (p16INK4a, p15INK4b, p18INK4c, and p19INK4d) bind to CDK4 and CDK6, and inhibit their kinase activities by interfering with their association with Cyclin D proteins [6] while the Kip/Cip proteins (p21Cip1, p27Kip1 and p57Kip2) inhibit Cyclin E-CDK2 [6,7]. Importantly, the Cyclin D-CDK4/6 complex can also bind to Kip/Cip CDKIs. However, this interaction enhances Cyclin D-CDK4/6 activity since proteins such as p27 appear to limit INK4 CDKIs' capacity to bind this complex [8-10]. Hence, a complex combination of INK4 and KIP/CIP protein regulations determine cell cycle progression.

Besides their function as tumor suppressors, INK4 and KIP/CIP proteins are known to impact cellular differentiation and organogenesis. Genetic studies in the mouse have demonstrated that INK4 and KIP/CIP genes are required for normal development [11-14]. For example, absence of p18 causes widespread organomegaly in mice [15] while mice lacking p57 die at birth and have multiple developmental abnormalities including hyperplasia and delayed differentiation [16,17]. On the other hand, p27 is necessary for normal differentiation of neural stem cells (NSCs) and mESCs [18,19]. CDKIs

are also essential for stem cell differentiation in adult organs. p21 can limit hair follicle quiescence while p27 has a key function in neuronal differentiation in the cortex [20,21]. Considered collectively, these reports suggest that CDKIs have a function during stem cell and progenitor cell fate specification beyond their tumour suppressive functions. However, the precise interplay between, signals of differentiation and transcriptional networks regulating CDKI expression remain to be fully elucidated.

Human pluripotent stem cells (hPSCs) represent a unique model system to study such mechanisms. These pluripotent cells can self-renew indefinitely while maintaining the capacity to differentiate into the three primary germ layers neuroectoderm, mesoderm and endoderm. Accordingly, hPSCs have been used by multiple groups to study the mechanisms by which cell cycle control differentiation [22-25]. Here we took advantage of this model system to investigate the mechanisms by which CKIs can coordinate cell cycle and differentiation. We found that INK4 and KIP/CIP display a tissue specific expression during hPSC differentiation and that their induction is controlled by a distinct set of developmental signalling pathways including ACTIVIN/NODAL-SMAD2/3. We uncovered that ACTIVIN/NODAL establishes a characteristic CDKI expression pattern in hPSCs via an activating NANOG-SMAD2/3 complex on p27 locus and instead a transcriptionally repressive NANOG-SMAD2/3-SNON complex on other CDKI loci. These transcriptional complexes are substituted during the exit from the stem cell self-renewal state and entry into cell fate specification. The induction of CDKI proteins upon differentiation via developmental master regulators such as the EOMES-SMAD2/3 complex results in an increase of the G1 phase length especially during endoderm specification. This extended G1 phase enables SMAD2/3 to accumulate into the nucleus of cells differentiating toward endoderm and allowing the transcriptional circuitry associated with ACTIVIN/NODAL signalling to be strongly enhanced. These results reveal a positive feedback loops between developmental transcription factors and

CKIs (p15/p18/p27/p57), which dynamically guides the specification of stem cells along different lineages. Collectively, our results unravel the molecular mechanisms by which cell fate decisions are controlled by the cell cycle machinery and describe a dynamical coordination of lineage specification with cell cycle progression.

Results.

CDKIs are induced during hPSC differentiation and display distinct expression patters in germ layers.

To explore the interplays between cell cycle progression and cell fate decisions, we decided to firstly, characterise the regulation of CDKIs in pluripotent self-renewing hPSCs and during their differentiation. We determined the cell cycle profile of hPSCs differentiating toward endoderm, mesoderm and neuroectoderm by EdU incorporation analyses, which revealed that differentiation systematically results in an increase of the G1 phase length. However, this increase was markedly different between each differentiated cell type (Fig. 1a-b). Lengthening of G1 phase remained limited in neuroectoderm while being considerably increased in the endoderm lineage. To find out if any cells are entering the G0 phase in endoderm differentiation, we generated H9 hPSCs with a three-colour FUCCI system, which allows for the detection of cells in early G1, late G1, G1/S transition, S/G2/M and G0 phases. Differentiation of these H9-FUCCI cells to the three germ layers indicated that cells only increase in G0 phase in endoderm differentiation whereas neuroectoderm and mesoderm cells do not increase cell fraction in G0 phase (Supplementary Fig. 1a). Collectively, these data support previous reports [4,5,24], showing that differentiation is associated with a change

in cell cycle. It also suggests that factors controlling the length of G1 phase could be induced upon differentiation of hPSCs.

To further validate this hypothesis, we investigated the expression pattern of CDKIs during hPSC differentiation since these proteins represent the main inhibitors of cell cycle progression. Immunostaining, Q-PCR and Western blot analyses revealed that most of the CDKIs (p14, p15, p16, p18 and p57) are not expressed in hPSCs with the exception of p27 in some cells, and to a lesser extent p21 (Fig. 1c-e). Similar analyses revealed that CDKIs display distinct expression patterns during germ layer specification (Fig. 1d-e; Supplementary Fig. 1b-d). p18, p15, p14, p27 and p57 showed a specific induction during endoderm specification while mesoderm differentiation was accompanied by an induction of p27 and p16. Finally, only p21 was significantly increased during neuroectoderm specification (Fig. 1d-e; Supplementary Fig. 1b). Further coexpression analyses of OCT4 or NANOG with CDKIs indicate the general non-overlapping expression of these pluripotency factors with CDKIs in early endoderm differentiation at day 1, except p27 (Supplementary Fig. 1c-d). We also analysed the expression of CDKIs by IF during germ layer differentiation. As expected, the positive control of differentiation indicated the loss of pluripotency factor SOX2, OCT4 and NANOG expression during germ layer differentiation at 72h, except for SOX2, which has a high expression in neuroectoderm (Supplementary Fig. 1e). Full IF analyses during endoderm at 72h (Supplementary Fig. 2a), mesoderm at 72h (Supplementary Fig. 2b) and neuroectoderm differentiation at 72h (Supplementary Fig. 2c), provided detailed insight to CDKI protein induction at the single-cell level and showed most extensive expression of CDKIs in endoderm germ layer. Endoderm cells at day 3 were also used for co-immunostaining for CDKIs and the marker of proliferation ki67, which indicated that ki67 signal is present in cells that do not express CDKIs (Supplementary Fig. 3a). Together these data suggest that self-renewing pluripotent stem cells

keep most of the CDKIs in a repressed state and avoid their expression, while CDKIs could have a tissue specific function during hPSC differentiation.

Next, we analysed the precise timing of CDKI induction upon endoderm differentiation of hPSCs since the specification of this germ layer is associated with the upregulation of several CDKIs. We took advantage of FUCCI-hPSCs since these cells can be synchronised in different phases of the cell cycle by cell sorting [22]. In this case, FUCCI-hPSCs were synchronised to early G1 (Fig. 1g) and then differentiated to endoderm. This approach results in a population of cells, which more homogenously differentiate to endoderm while progressing with the same timing through the cell cycle [26]. hPSCs differentiating to endoderm using this approach were examined for the expression of mesendoderm/endoderm and CDKI markers every 6 hours (Fig. 1h and Supplementary Information). These analyses revealed that CDKIs are induced concomitantly with early mesendoderm markers such as Mixl1 prior the induction of the definitive endoderm marker SOX17. This very rapid induction could enable CDKIs to inhibit Cyclin/CDK complexes during cell cycle progression upon differentiation thereby explaining the lengthening of the G1 phase characterising the cell cycle profile of endoderm cells.

Altogether, these results suggest that CDKI expression in tightly controlled in pluripotent stem cells and tissue specific induction of CDKIs determine the cell cycle profile of hPSC derivatives after differentiation into the three germ layers.

A compound library screen of epigenetic inhibitors identifies EZH2 and SMYD2 as factors governing hPSC self-renewal and differentiation.

To uncover novel epigenetic regulators of hPSCs, we used three pluripotency markers (OCT4, CD133, SSEA4) for identifying pluripotent and differentiated cells in the experiments. The H9 hPSC

used for this screening were genetically engineered, with the sequence coding for green fluorescent protein (eGFP) integrated in the endogenous locus, resulting in controlled expression of a OCT4-eGFP fusion protein driven by the endogenous *OCT4* promoter. The CD133 and SSEA4 are widely used pluripotency markers expressed in hPSCs but lost in differentiated hPSCs.

We hypothesised that the self-renewal of hPSCs and the early cell fate decisions of hPSCs leading to CDKI expression during germ layer differentiation are controlled by epigenetic mechanisms. To address this, we performed a focused library screen consisting of validated small molecule inhibitors (Supplementary Table 4) targeting epigenetic regulators such as 'readers, writers and erasers' of a histone code [32]. These experiments aimed to identify molecular targets of small molecule compounds which specifically inhibit epigenetic regulators and thereby affect pluripotency marker-expressing cells (Fig. 2a). We measured pluripotency marker (CD133, OCT4, SSEA4) expression by flow cytometry and hPSC cell growth after incubating the cells with the compounds for 5 days, which allowed us to identify effective compounds that impact pluripotency marker expressing cells while also detecting cells that do not express these markers which represent differentiated hPSCs. The compound library used in our experiments consisted of 142 compounds that have been verified to be active and targeting specific epigenetic modifying enzymes (Fig. 2b). Overall, this screening identified compounds that significantly and reliably reduced the relative percentage of marker positive hPSCs and negative (CD133-/OCT4-/SSEA4-) differentiated hPSCs (Fig. 2c-d).

The screening identified novel compounds that target distinct epigenetic regulatory components and altered the percentage of pluripotent cells versus differentiated cells. Among the top candidate compounds that increased differentiation evidenced by increased pluripotency marker negative cells were EZH2 inhibitors (GSK343, UNC1999, CPI-169) (Fig. 2c-d). In contrast, among the top candidate compounds that decreased background differentiation evidenced by

decreased marker negative cells were SMYD2 inhibitors (PFI-5, LLY-507, BAY-598) (Fig. 2c-d). These results suggested that EZH2 inhibition could induce differentiation of hPSCs whereas SMYD2 inhibition could block differentiation of hPSCs. EZH2 is a subunit of polycomb repressive complex 2 (PRC2), which catalyzes the trimethylation of H3K27 at developmental gene promoters [27]. This maintains the transcriptionally repressive state of genes associated with differentiation to retain stem cell pluripotency in hPSCs [28]. The SMYD (SET and MYND domain containing protein) family contains five members (SMYD1-5) that share a catalytic SET domain and an MYND motif, which are involved in lysine methylation and protein–protein interaction, respectively. SMYD2 has been reported to suppress cell proliferation by mediating H3K36me2 deposition [29], and in transcriptional regulation by mediating H3K4me1 [30]. In addition, SMYD2 is reported to methylate the K370 site of p53 repressing its antitumor effect [31], whereas SMYD2 induces RB1 methylation at lysine 810 in some cancers that promotes cell cycle progression of these malignant cells [32].

To further validate EZH2 and SMYD2 as a target we used two EZH2 chemical inhibitors (GSK343 and UNC1999) and two SMYD2 chemical inhibitors (PFI-5 and LLY-507) to investigate their effects on hPSCs (Fig. 2e-f). Treatment of hPSCs with EZH2 inhibitors indicated a loss of pluripotency and increased differentiation based on cell morphology (Fig. 2e), while qPCR analyses revealed a reduction of core pluripotency factors OCT4, NANOG and SOX2 with concomitant differentiation mostly to endoderm (T, EOMES, SOX17) and less to mesoderm (MESP1, MESP2) but not to neuroectoderm (SOX1, PAX6) (Fig. 2f, h). We also observed increased differentiation to endoderm and less to mesoderm by qPCR analyses of EZH2 knockdown hPSCs (Fig. 2e, g-h; Supplementary Fig. 3b). In contrast, SMYD2 inhibitors did not increase the differentiated morphology of hPSCs and did not increase spontaneous differentiation of hPSCs to germ layers but had a consistently lower background signal of endoderm markers (Fig. 2e-f; Supplementary Fig. 3b). Therefore, we investigated SMYD2 knockdown effects upon endoderm differentiation. We

found that cells differentiated to endoderm for 36h with SMYD2 inhibitors indicated higher expression of pluripotency markers OCT4, NANOG and SOX2 while there is a reduced expression of T, EOMES and SOX17 compared to control endoderm differentiation (Fig. 2i-j). SMYD2 knockdown also reduced differentiation to endoderm as shown by qPCR analyses and immunofluorescence microscopy of SMYD2 knockdown (Fig. 2i-j).

Next, we hypothesised that EZH2 and SMYD2 could regulate the expression of CDKIs in hPSCs and during endoderm differentiation. To investigate this, we analysed the expression of CDKIs upon the treatment with EZH2 and SMYD2 inhibitors. EZH2 inhibition resulted in the induction of p15, p18 and p57 in hPSCs upon both chemical inhibition and genetic depletion of EZH2 (Fig. 2k-l). In contrast, SMYD2 inhibition did not upregulate CDKI expression in hPSCs (Fig. 2k), but SMYD2 chemical inhibition and genetic depletion both attenuated the induction of CDKIs during endoderm differentiation (Fig. 2m).

Collectively, our compound screening identified EZH2 and SMYD2 as epigenetic factors regulating hPSC self-renewal and endoderm differentiation that also govern CDKI expression in hPSCs and early differentiation to endoderm (Fig. 2n).

Developmental signalling pathways ACTIVIN, BMP4, FGF and PI3K control the expression of different CKIs during cell fate specification.

The differential expression of CDKIs in germ layers indicated that their expression could be controlled by different sets of signalling pathways. To investigate this hypothesis, we blocked individually each pathway (ACTIVIN A/TGFβ, BMP4, FGF2, PI3K) known to be involved in endoderm specification and then measured the expression of each CDKI by qPCR. Inhibition of ACTIVIN/NODAL signalling resulted in the loss of p15, p18, p27 and p57 expression, while FGF2 inhibition decreased p18 expression and BMP4 inhibition reduced p15 and p14 induction (Fig. 3a)

while increased ACTIVIN and BMP4 induced these CDKIs respectively (Supplementary Fig. 3c-e). Finally, absence of PI3K inhibition caused a loss of p15 and p57 expression, while increasing p18 expression (Fig. 3a). Therefore, developmental pathways directing cell fate choice also appear to control expression of CDKIs, uncovering a crosstalk between differentiation signals and cell cycle regulation.

To validate the interconnection of ACTIVIN/TGFβ/NODAL-SMAD2/3 signalling and CDKIs, we analysed SMAD2/3 ChIP-seq data in hPSCs. Interestingly, we found SMAD2/3 binding onto the regulatory regions in the proximity of p14/p16, p15, p18, p21, p27 and p57 loci in pluripotent cells despite the absence of active CDKI expression in pluripotent condition (Fig. 3b). To gain insight to the timing and putative recruitment of SMAD2/3 onto CDKI loci, we performed SMAD2/3 ChIP-QPCR every 12 hours after induction of endoderm specification. SMAD2/3 proteins showed binding already in self-renewing pluripotent hPSCs but also during endoderm differentiation (Supplementary Fig. 3f), suggesting that ACTIVIN/NODAL-SMAD2/3 signalling could directly control the expression of CDKIs differently in pluripotent cells and differentiating cells. Promoter-luciferase assays further indicated the existence of an ACTIVIN/TGFβ/NODAL signalling dependent and SMAD2/3-mediated induction of CDKIs through their promoter regions in endoderm cells (Supplementary Fig. 3g).

Collectively, developmental signalling pathways, especially ACTIVIN/TGFβ-SMAD2/3, directly control the expression of CDKIs during stem cell differentiation. However, it raised the intriguing question on the nature of the molecular mechanisms that provide stem cells with a distinct repressed CDKI expression signature that is rapidly altered upon the initiation of stem cell differentiation.

TGFβ/ACTIVIN-SMAD2/3-EZH2 cooperates with OCT4, NANOG and SOX2 in self-renewing hPSCs to keep CDKIs in an epigenetically poised state for rapid activation.

Since SMAD2/3 cooperates with sequence-specific transcription factors in gene regulation [33], we decided to investigate this mechanism in CDKI regulation. By first performing transcription factor motif analyses within SMAD2/3 binding peaks we found the enrichment of OCT4, NANOG and SOX2 binding motifs at the SMAD2/3 binding peaks at CDKI regulatory regions, suggesting that the core pluripotency transcription factors OCT4, NANOG and SOX2 could cooperate with SMAD2/3 in controlling CDKI expression in self-renewing pluripotent stem cells. SMAD2/3 is integrated in the "stemness" network in hPSCs and is involved in maintaining pluripotency by co-occupying loci that support self-renewal and the pluripotent state together with the key pluripotency factor NANOG, while at the same time allowing developmental genes to be rapidly induced upon differentiation [34-36]. Hence, CDKIs could represent genes that are similarly regulated by these stemness pathways.

ChIP-sequencing analysis of OCT4, NANOG and SOX2 revealed their binding on most CDKI loci in hPSCs to overlapping regions with SMAD2/3 (Fig. 3b). Since SMAD2/3, NANOG and OCT4 all bind to CDKI regulatory regions, we decided to explore the possible cooperation between SMAD2/3 and NANOG on CDKI loci in hPSCs. Sequential ChIP of OCT4 followed by SMAD2/3 (Fig. 3c) and NANOG ChIP followed by SMAD2/3 (Fig. 3d) showed that these stemness factors form a transcriptional complex on CDKI loci p18, p15 and p27 and p57.

Next, to determine the effect of the core pluripotency factors on CDKI expression, we focused on OCT4 and NANOG, and performed its knockdown in hPSCs. SMAD2/3 and OCT4 CHIP experiments in Scramble and OCT4 iKD cells indicated reduced SMAD2/3 binding to CDKI loci in the absence of OCT4 (Fig. 3e). CDKI loci are also marked by both activating H3K4me3 and repressing EZH2-dependent H3K27me3 modifications (Fig. 3b), thus representing bivalent marks

that are usually associated with developmental loci that are induced during tissue specification [37]. OCT4 iKD reduced the binding of EZH2 on CDKI loci (Fig. 3e), whereas OCT4 iKD led to the upregulation of CDKIs both at the mRNA and protein levels (Fig. 3f-g). We also observed a decrease in the EZH2-regulated repressive H3K27me3 bivalency mark on p18, p15, p21 and p57 loci whereas the activating H3K4me3 bivalency mark showed no changes (Fig. 2i). Co-transfection of an OCT4 expression vector and p18, p15 and p57 promoter-luciferase constructs reduced luciferase signal upon transient overexpression of OCT4 (Fig. 2j).

SMAD2/3 and NANOG CHIP experiments also indicated reduced SMAD2/3 binding to CDKI loci in the absence of NANOG (Fig. 3d; Supplementary Fig. 3h). Furthermore, NANOG KD led to the upregulation of CDKIs both at the mRNA and protein levels (Fig. 3h), and decreased in the repressive H3K27me3 bivalency mark on p18, p15, p21 and p57 loci whereas the activating H3K4me3 bivalency mark showed no changes (Supplementary Fig. 3i). As for OCT4, co-transfection of a NANOG expression vector and p18, p15 and p57 promoter-luciferase constructs showed a reduced luciferase signal upon transient overexpression of NANOG (Fig. 3j).

Altogether, these data uncovered that the SMAD2/3-OCT4-NANOG-EZH2 complex seems to maintain the characteristic CDKI expression pattern in hPSCs. However, the precise epigenetic mechanisms mediating the changes in CDKI expression still remained unclear.

SMAD2/3-EZH2 maintains bivalent chromatin signatures on CDKI loci in self-renewing stem cells together with SNON.

The SMAD2/3 repressor protein SNON/SKIL has been shown to suppress primitive streak and definitive endoderm genes in hPSCs, which helps to maintain pluripotency by suppressing differentiation [38]. We therefore hypothesised that SNON could also mediate the transcriptionally suppressive function of SMAD2/3-NANOG-OCT4 complex on CDKI loci. SNON ChIP data analyses

revealed its potential binding to the same regions as SMAD2/3, OCT4 and NANOG in leukemic cells (Fig. 3b), although no such information was available for hPSCs. Hence, we performed SNON ChIP qPCR experiments that indicated its enrichment on p15, p18, p21 and p57 loci, which was lost upon OCT4 KD (Fig. 3e) and NANOG KD (Supplementary Fig. 3h), and ACTIVIN/TGFβ-SMAD2/3 pathway inhibition with SB431542 (Supplementary Fig. 3j), indicating that NANOG and SMAD2/3 are involved in the recruitment of SNON onto CDKI loci. Lastly, co-transfection of CDKI promoter-luciferase constructs with a SNON expressing plasmid reduced the luciferase signal (Fig. 3j), indicating its repressive function on CDKI loci in hPSCs. These results show that SNON supports the pluripotent state by restricting the induction of CDKIs in hPSCs via a transcriptionally repressive complex together with SMAD2/3 and the core pluripotency factors.

Collectively, our results uncover an intricate control of CDKI expression in hPSCs: the expression of other CDKIs is maintained in a poised state in pluripotent cells by an OCT4-NANOG-SMAD2/3-SNON complex.

SMAD2/3 switches its binding to SMYD2 and EOMES to induce CDKI expression during the initiation of endoderm differentiation.

ACTIVIN/TGFβ/NODAL has seemingly opposing functions in maintaining hPSCs in their pluripotent state but also inducing their differentiation to endoderm. TGFβ/ACTIVIN-SMAD2/3 signalling is crucial for regulating the self-renewal and pluripotency of hPSCs by SMAD2/3-mediated transcriptional regulation of stem cell loci [39], whereas it also induces definitive endoderm formation and EMT thus leading to the formation of endoderm cells with mesenchymal characteristics [26]. This raises the question of how CDKI expression is regulated in pluripotent versus differentiating cells.

SMAD2/3 ChIP-seq analyses revealed the binding of SMAD2/3 to the CDKI regulatory regions in endoderm cells, some of which were overlapping with the binding regions found in pluripotent cells whereas there were also some new binding sites (Fig. 3b). The motif analyses of SMAD2/3 peaks in the proximity of CDKI loci in endoderm indicated the presence of transcription factor EOMES (Fig. 4a). This supported the notion that SMAD2/3 can switch its transcriptional partners during differentiation and this mechanism is essential for endoderm specification [26,35,40,41]. Indeed, we have shown previously that SMAD2/3 interacts with NANOG in hPSCs to maintain pluripotency but with EOMES to drive endoderm formation [26,40]. Thus, we hypothesised that SMAD2/3 could cooperate with EOMES to induce CDKI expression upon differentiation. EOMES ChIP-Seq analyses in endoderm showed that this master regulator of differentiation can indeed be found on p18, p15 and p57 loci on regions overlapping with SMAD2/3 binding, and ChIP-QPCR validated these observations (Fig. 3b). Importantly, sequential ChIP of EOMES and SMAD2/3 demonstrated that these transcription factors are part of the same protein complexes co-binding these genomic regions (Fig. 4b). Furthermore, knockdown of EOMES expression during endoderm differentiation resulted in a decrease of CDKI expression (Fig. 4c-d). Promoter-luciferase assays provided further evidence that EOMES is able to directly induce p15, p18 and p57 expression in endoderm cells (Fig. 4e).

To further determine the molecular function of EOMES in the induction of CDKI expression, we performed SMAD2/3 ChIP experiments in hPSCs knocked down for EOMES and differentiating toward the endoderm lineage. These experiments revealed that EOMES is necessary for the recruitment of SMAD2/3 to these loci (Fig. 4f). Reversely, inhibition of SMAD2/3 did not affect the binding of EOMES (Fig. 4h). Thus, EOMES seems to be necessary for SMAD2/3 binding on CDKI loci but not vice versa. To characterise the timing of NANOG and EOMES on CDKI loci, we performed ChIP-QPCR of these factors every 12h during endoderm differentiation (Fig. 4h-j). NANOG binding

to CDKI loci was lost by 24h of endoderm differentiation while EOMES binding started to increase at this time point on CDKI loci, indicating an exchange of factors from the key pluripotency factor NANOG to the inducer of definitive endoderm EOMES. At the same time, SMAD2/3 binding was detected in pluripotent conditions as well as during endoderm differentiation (Fig. 4h-j).

Next, we analysed the deposition of histone modification on CDKI loci during endoderm differentiation and the effect of EOMES. The knockdown of EOMES led to the reduction of transcriptionally activating modification H3K4me3 whereas there was a concomitant increase in repressive H3K27me3 (Fig. 4k). This led us to investigate the binding of SMYD2 on CDKI loci since we had identified it in compound screening and showed its function as a regulator of CDKI loci. ChIP-qPCR indicated that SMYD2 bound to the same regions near CDKI loci as SMAD2/3 and EOMES, and the inhibition of ACTIVIN/TGFβ signalling with SB431542 reduced its binding. This suggested that SMAD2/3 help to recruit SMYD2 to CDKI regulatory regions (Fig. 4l). In addition, sequential ChIP of EOMES and SMYD2 demonstrated that these factors are part of the same protein complexes co-binding CDKI genomic regions (Fig. 4m). Lastly, SMYD2 binding is decreased to CDKI loci upon EOMES KD, indicating that EOMES contributes to SMYD2 recruitment to CDKI loci during endoderm differentiation.

Collectively, these data suggest that pluripotency factors such as NANOG with EZH2 and SNON are replaced on CDKI loci during the switching from self-renewal to differentiation, whereas EOMES together with SMAD2/3 and SMYD2 cooperate to directly induce the expression of p15, p18 and p57 CDKI during endoderm differentiation (Fig. 4o).

CDKIs direct lineage specification of human pluripotent stem cells.

The germ layer specific induction of CDKIs during hPSC differentiation suggested a potential function for them not only in cell cycle regulation but also in cell fate decisions. Therefore, we

decided to determine the role of CDKIs during endoderm differentiation using gain- and loss-of-function studies. We first knocked down each individual CDKI (p14/p16, p15, p18, p21, p27 and p57; Fig. 5a) using stable overexpression of shRNA as described previously [22]. A decrease in expression (>80%) in day 2 differentiated cells was confirmed for CDKIs at both protein and mRNA levels (Supplementary Fig. 4a-b). As expected due to the absence of CDKI expression in hPSCs, the morphology and number of pluripotent colonies derived after transfections was not affected when compared to control, except for a decrease in p27 (Fig. 5a, Supplementary Fig. 4c), suggesting that p27 could impact hPSCs self-renewal. Nonetheless, Q-PCR analyses for assessing the background differentiation revealed that hPSCs with a knockdown for the expression of p15, p18 and p57 show a decrease in endoderm marker expression (SOX17, EOMES and GSC; Supplementary Fig. 4d) while neuroectoderm markers PAX6 and SOX1 were increased (Supplementary Fig. 4d). On the other hand, hPSCs with a knockdown for p21 expression showed the opposite effect and p16 knockdown was associated with a decrease in mesoderm markers such as Brachyury/T and Mesp1 (Supplementary Fig. 4d). Finally, knockdown of p27 only caused limited change in differentiation marker expression (Supplementary Fig. 4d). Flow cytometry analyses confirmed these observations (Supplementary Fig. 4e-h) thereby demonstrating that decrease in CDKI family members increased specific background of differentiation.

To determine the function of CDKIs on germ layer specification, hPSCs with a knockdown for p14/p16, p15, p18, p21, p27 and p57 expression were differentiated into endoderm, mesoderm and neuroectoderm. Q-PCR analyses revealed that the absence of p18, p15 and p57 reduced the capacity of hPSCs to differentiate into endoderm and into mesoderm as shown by a decrease in mesendoderm markers and an increase in pluripotency markers (Fig. 5b and Supplementary Fig. 4i-j). On the other hand, the same hPSC lines displayed an enhanced capacity to differentiate toward the neuroectoderm lineage (Fig. 5b and Supplementary Fig. 4k). In contrast, loss of p21

improved endoderm differentiation of hPSCs while neuroectoderm specification was less efficient and mesendoderm was not affected (Fig. 5b and Supplementary Fig. 4i-k). Finally, p16 KD improved mesoderm specification while other germ layers were not significantly affected (Fig. 5b and Supplementary Fig. 4i). We confirmed these results at the protein level by flow cytometry (Fig. 5c-d and Supplementary Fig. 4j-k) and immunostaining (Fig. 5e). Finally, we performed teratoma assays in mice on hPSCs knocked down for p18, p21 and p57 expression (Fig. 5f-g). Histological analyses revealed that the tumours generated from hPSC with a knockdown for p18 and p57 contain an increased quantity of neuroectodermal tissues to the detriment of endodermal tissues (Fig. 5g). In contrast, hPSCs with a knockdown for p21 produced more mesoderm and endoderm derivatives and less neuroectoderm when compared to teratomas from Scramble shRNA (Fig. 5g). Considered together, these data demonstrate that CDKIs impact germ layer differentiation *in vitro* and that a specific combination of CDKIs is involved in hPSC differentiation to achieve particular cell fate choice.

To further confirm this hypothesis, we investigated the effects of CDKI overexpression on differentiation of hPSCs. For that, we expressed individual CDKIs constitutively in hPSCs (Supplementary Fig. 5a) as previously described [22]. CDKI overexpression strongly reduced colony formation after transfection (p21, p27, p57, p18 and p15 in Fig. 6a) suggesting a negative effect on cell survival and/or self-renewal. The remaining colonies displayed a level of CDKIs comparable or lower than the natural level of expression observed in endoderm cells with the exception of p14 that had higher expression (Fig. 6b, Supplementary Fig. 5b). As expected, CDKI overexpressing hPSCs grew very slowly and overexpression of p15, p18 and p57 resulted in the lengthening of the G1 phase while other CDKIs had weaker effects (Fig. 6c-d). CDKI overexpressing hPSCs were also strikingly more prone to spontaneous differentiation (Fig. 6b). These observations were confirmed by immunostaining and Q-PCR analyses showing that overexpression of p15, p18 and p57

increased the background expression of endoderm markers while reducing pluripotency and neuroectoderm markers (Fig. 6e-f; Supplementary Fig. 6a). p16 overexpression modestly induced mesoderm markers while blocking neuroectoderm. Only p21 overexpression induced the neuroectoderm markers Pax6 (Fig. 6f). These results were further confirmed by flow cytometry analyses (Supplementary Fig. 5d-g).

Next, we studied the effect of CDKI overexpression during hPSC specification toward endoderm, mesoderm and neuroectoderm. Flow cytometry analyses and immunostaining showed that overexpression of p15, p18 and p57 during endoderm differentiation with a less than optimal concentration of Activin A, which results in inefficient differentiation, increased endoderm marker expression while decreasing pluripotency markers (Fig. 6g and Supplementary Fig. 5h-i). Furthermore, the same overexpression decreased neuroectoderm differentiation efficiency while increasing pluripotency markers (Fig. 6h and Supplementary Fig. 5j).

Collectively, these functional studies showed that CDKIs promote or limit the specification of hPSCs toward specific germ layers while determining the length of their cell cycle.

Manipulation of positive feedback loops between SMAD2/3-CDKI and developmental transcription factors can be harnessed for guiding tissue self-formation for biomedical applications.

We decided to delineate further the molecular mechanisms by which CDKIs could influence cell fate choice of hPSCs. We have shown previously that CDK4/6-Cyclin D complex can regulate the transcriptional activity associated with ACTIVIN/NODAL signalling by inhibiting the cytoplasmic-nuclear shuttling of SMAD2/3 [22] during the late part of the G1 phase. Therefore, we hypothesised that CDKIs could influence hPSC fate choice by blocking Cyclin D-CDK4/6 inhibition of SMAD2/3. To validate this hypothesis, we transitorily overexpressed p18 or p15 in FUCCI-hPSCs

and then studied the capacity of the resulting cells to initiate endoderm differentiation during late G1 phase when Cyclin D-CDK4/6 block SMAD2/3 activity (Fig. 7a-b). Transient expression of CDKIs in late G1 phase enhanced endoderm marker expression as analysed by Q-PCR and flow cytometry (Fig. 7b-c). Furthermore, SMAD2/3 ChIP-QPCR in the early and late G1 phase revealed that p18 overexpression allows SMAD2/3 binding to the endoderm loci EOMES, Mixl1 and GSC (Supplementary Fig. 6b) while SMAD2/3 is usually blocked to access these genomic regions by Cyclin D-CDK4/6 activity [22]. We also performed SMAD2/3 ChIP-QPCR at different time points during endoderm differentiation of hPSCs with a knockdown for p18, p15 and p57, and observed that SMAD2/3 binding to endodermal loci Mixl1, EOMES and SOX17 was strongly reduced (Supplementary Fig. 6c-e). Next, we determined the subcellular localisation of SMAD2/3 during differentiation in the presence of absence of CDKIs. These experiments showed that knockdown of CDKI expression caused cytoplasmic accumulation of SMAD2/3 away from chromatin during endoderm differentiation, while overexpression of the same CDKIs had the reverse effect, resulting in SMAD2/3 accumulation onto chromatin (Fig. 7d; Supplementary Fig. 6f). Furthermore, overexpression of p15, p18 and p57 increased the transcriptional activity of SMAD3 (Fig. 7e; Supplementary Fig. 6g) while CDKIs did not affect the transcriptional activity of SMAD3-EPSM which contains mutations at CDK4/6 phosphorylation sites (T178V, S203A, S207A, S212A; [42]) (Fig. 7e), suggesting that these linker residues mediate CDKI effects on SMAD2/3. These results demonstrate that the CKIs promote the activity of Activin/Nodal signalling by limiting the inhibitory action of Cyclin D-CDK4/6 on SMAD2/3-mediated transcription.

Importantly, the induction of CDKIs upon differentiation also results in the lengthening of the G1 phase. Cell cycle analyses showed that overexpression of p18 lengthened the G1 phase of differentiating endoderm cells while the knockdown of the same CDKI decrease the size of G1 (Fig. 7f). This change in the cell cycle profile could enable SMAD2/3 to bind for a prolonged period of

time onto endodermal genes and thus to increase the expression of master regulators necessary for cell fate commitment. Following this hypothesis, we explored if endoderm-promoting CKIs (p15, p18 and p57) could affect SMAD2/3 binding and activity during germ layer specification. Since endoderm differentiation is accompanied by a significant extension of the G1 phase length, we wanted to gain insight to the dynamics of SMAD2/3 activity according to the cell cycle phases. For this we transfected FUCCI-hPSCs with the SBE4-luciferase construct and analysed SMAD2/3 transcriptional activity in pluripotent cells and day 1 endoderm in the presence or absence of p18 overexpression (Supplementary Fig. 7a). Luciferase assay revealed that the transcriptional activity of SMAD2/3 is elevated in endoderm cells overexpressing p18 and thus displaying an increased G1 enhanced SMAD2/3 activity. Next, we analysed the dynamics of SMAD2/3 binding on endoderm loci during differentiation in the presence or absence of p18 overexpression (Supplementary Fig. 7b). SMAD2/3 ChIP-QPCR revealed a transient binding of SMAD2/3 to its developmental target loci in early G1 phase of pluripotent cells, while its binding was extended across the whole G1 phase in endoderm cells (Supplementary Fig. 7b) especially in the presence of p18. Similar experiments performed on H3K36me3 revealed that an increase in this histone modification marking active transcription was not only concomitant with elevated SMAD2/3 binding in the extended G1 phase of endoderm cells but was also further enhanced by p18 overexpression (Supplementary Fig. 7b). Importantly, similar observations were made with the genomic regions including the cell cycle inhibitor p15, p18 and p57 loci (Supplementary Fig. 7c), showing that SMAD2/3 activity is increased by CDKIs and in turn, SMAD2/3 induces CDKI expression thereby forming a regulatory circuitry with these cell cycle inhibitors (Supplementary Fig. 7c). Altogether, these results show that CDKIs limit the activity of Cyclin D/CDK4-6 complexes, which concomitantly increase of the length of the G1 phase both of which increase of transcriptional activity of SMAD2/3 (Supplementary Fig. 7d).

Cells have a remarkable capacity to self-organise and form complex tissues without apparent external guidance, suggesting inherent developmental programs that are activated self-sufficiently in a step-wise manner. The positive feedback loops between SMAD2/3-CDKI and developmental transcription factors could provide insight to the mechanisms of such developmental circuitries that can gradually drive cell fate specification and complex tissue formation since each step initiates the next molecular process (Supplementary Fig. 7e). This provides insight to developmental processes during organogenesis in early development. Hence, we hypothesised that it might be possible to make use of the discovered positive feedback loops between SMAD2/3-CDKI and developmental transcription factors for guiding tissue self-formation for biomedical applications. To begin to address this issue, we used an inducible p18 CRISPR/Cas-KRAB knockdown system that relies on the expression of p18 gRNAs under DOX-inducible conditions that can be turned on and off upon the addition of Doxycyclin. We established a p18 iKD in hIPSC line and also utilised the small molecule compound PD0332991 for inhibiting CDK4/6 by adding it to the media. We cultured the hIPSCs in organoid condition and adapted a previously published differentiation protocol [22,26] for pancreatic insulin-producing beta-cells for this experimental condition (Fig. 7g). Each hiPSC step-wise differentiation stage, definitive endoderm, dorsal foregut, pancreatic progenitors/pancreatic endoderm, and pancreatic islet cells, were supplemented with either DOX for p18 knockdown, or with PD0332991 for CDK4/6 inhibition. QPCR analysis of these treatments indicated that p18 iKD and PD0332991 shift cell fate specification by reducing specification toward pancreatic beta-cell fate, as in definitive endoderm (SOX17 and FOXA2) and pancreatic islet cell (NGN, INSULIN) specification stage. At the same time, p18 iKD and PD0332991 further shift cell fate specification toward other possible cell specification routes (stage 1: mesoderm markers MESP1, MESP2; stage 4: alpha cell marker GSC and delta cell marker SST). In contrast, p18 iKD and PD0332991 treatment during dorsal foregut (HNF4A, HLXB9) and pancreatic

progenitor (PDX1, SOX9) stage, improve cell fate specification toward these cell identities while reducing the other possible cell specification routes (stage 2: ventral foregut markers SOX17 and CER; stage 3: hepatic endoderm markers AFP and HEX). These results indicated that the dynamical activation and inactivation of the SMAD2/3-CDKI-developmental TF circuitry contributes to guiding cell fate specification upon multiple routes available for the differentiating cells.

We also examined the expression of some key markers by immunostaining, and found that temporal CDK4/6 inhibition with PD0332991 at pancreatic progenitor stage further improves the specification toward PDX1 positive pancreatic progenitor cells (Fig. 7i), NGN3 (Fig. 7j) and INSULIN expressing beta-cells at the islet cell specification stage (Fig. 7k). On the other hand, the constant inhibition of CDK4/6 with PD0332991 throughout the differentiation process reduces the formation of PDX1 positive cells, and interestingly, alters the spatial expression of NGN3 and INSULIN (Fig. 7k). Collectively, these results suggest that the dynamic fluctuation of CDK4/6 inhibition during pancreatic islet fate specification contributes to efficient differentiation. Furthermore, manipulation of the positive feedback loops between SMAD2/3-CDKI and developmental transcription factors can be intersected for guiding tissue self-formation for producing patient-specific cell types such as pancreatic islet cells that have biomedical applicability.

Discussion

Our study has uncovered EZH2 and SMYD2 as novel epigenetic regulators of the switching from self-renewal to differentiation, and a function for cell cycle inhibitors in directing lineage specification of human stem cells. Our results show that self-renewal and differentiation signals such as ACTIVIN maintain the stem cell-specific expression profile of CDKIs. Accordingly, hPSCs keep most CDKIs in a poised state via NANOG-OCT4-SMAD2/3-EZH2-SNON complex and ready for

rapid induction. In response to developmental signals that trigger stem cell differentiation, the bivalently marked (H3K4me3 and H3K27me3) poised CDKIs are induced only few hours after the induction of differentiation via cooperation between SMAD2/3 and developmental master regulators such as EOMES and epigenetic regulator SMYD2. CDKIs can then inhibit in part the activity of the Cyclin D-CDK4/6 complex in the next G1 phase during which endoderm cells are specified. This mechanism allows CDKIs not only to extend the G1 phase but concomitantly enhance the transcriptional activity of ACTIVIN/NODAL/TGFß signalling which can then gradually activate the expression of master regulators of cell fate choice such as SOX17. This positive feedback loop appears to be important for hPSCs to commit toward the endodermal fate and to enable ACTIVIN/NODAL/TGFß signalling to convert from signalling that protects pluripotency to signalling that promotes differentiation (Fig. 7k).

These mechanisms could explain how cell fate commitment occurs upon differentiation of a diversity of stem cells. Indeed, TGFß signalling has been shown to directly control the expression of p15, p21, p27 and p57 in tissue specific progenitors located in the central nervous system, bones, hematopoietic system and liver [43-46]. Therefore, the mechanisms uncovered by our current study could be relevant for a diversity of cell types in which the induction of CDKIs by TGFß will be not only important for blocking their proliferation but also for enhancing the activity of differentiation signals necessary for their final maturation/commitment.

Similarly, a decrease in CDKI expression is a common process in cancer and our results suggest that this mechanism could not only lead to increased capacity to proliferation but also a decreased capacity to sense differentiation signals. Indeed, TGFß signalling has been shown to control the expression of several CDKIs in a diversity of cancer lines [46]. As uncovered by our results, the maintenance of p27 expression in hPSCs via SMAD2/3-NANOG-OCT4 is particularly interesting considering the dual role of p27 as a tumor suppressor as well as an oncogene [9]. It suggests that

this molecular machinery could also play a role in promoting the proliferation of cancer cells or cancer stem cells. OCT4 and NANOG expression has indeed been linked to various tumors [47-50]. Furthermore, the suppression of other CDKIs by SMAD2/3-NANOG-OCT4-EZH2-SNON in hPSCs could also support the proliferation capacity of cancer stem cells, considering the tumor-promoting effects of SNON [51-54]. Hence, the results uncovered by our study shed light to the interconnection between cell cycle regulation and cell fate decisions, which are likely to be relevant for a broad range of cellular settings including stem cells, non-malignant progenitors and cancer stem cells.

The role of CDKIs in embryogenesis is less clear, since absence of these proteins in mice do not result in severe phenotypes during early development. p18 is among the more abundantly expressed INK4-family CDKIs and is detected throughout mouse embryogenesis, p15 is less abundant and 19(ARF)/p16 show relatively little expression in prenatal tissues while they become more abundantly expressed in adult tissue [55] where they act as tumor suppressors [56,57]. Hence, among these CDKIs, p18 seems to play a more important role during early development, which is also underlined by the phenotype of the mouse knockout for p18, which show an increased body size and widespread organomegaly [15]. Interestingly, p18 mRNA expression during development in the mouse matches our *in vitro* differentiation results for human stem cells in that it shows the strongest induction during differentiation among the CDKIs and is particularly expressed in endoderm. However, it is known that CDKIs can have at least partially redundant activity, which could mask their function during gastrulation. For instance mice with a knockout for p27 have an increased body size like p18 KO mice [57], while p18/p27 double-null mice have an even more increased body size compared to both single knockouts [15]. Such an additive effect on body size is not seen in p18/p21 double KO mice [58], suggesting CDKI-specific redundancy. Furthermore, p18 can cooperate with p27 during liver regeneration *in vivo* by regulating DNA synthesis and G1/S

phase progression [59], and it is likely to involve also p57 [60]. Due to the emergence of CRISPR/Cas9 mediated genome-editing system [61], the questions of CDKI cooperation in various tissues could be revisited by simultaneous multi-gene knockouts. These studies would also benefit from genome-wide analysis of epigenetic changes, which has become possible only in recent years [62].

Importantly, our results also suggest that CDKI could function in a tissue specific manner. Indeed, we observed that p21 was functional important for neuronal differentiation while p16 promoted mesoderm and other CDKIs endoderm differentiation. Therefore, the mechanisms involving CDKI in differentiation are likely to be much more complex. Indeed, several studies have suggested that KIP/CIP family CDKIs could be located on the chromatin and interfere directly with transcriptional activity of specific genes. p21 has been shown to bind E2F1 [63] while p27 interacts with p130-E2F4 complex to recruit transcriptional co-repressors such as Sin3A and histone deacetylases [64]. p27 can also interact with Neurogenin-2 to promote the differentiation of neuronal progenitors in the cortex [20] while p57 can form a complex with MyoD to promote expression of muscle-specific genes [65]. Therefore, CDKIs could have a function during cell fate specification as transcriptional regulators besides their role as cell cycle inhibitors.

Since CDKIs form a regulatory loop with key signaling pathways to promote specification toward the endoderm lineage such regulation could be used in the future for devising universal method of differentiation which could work with any hIPSCs lines for a diversity of therapeutic applications, disease modeling or *in vivo* reprogramming [66-68]. The positive feedback regulation of CDKIs could also be exploited for guiding lineage specification of complex tissues during *in vitro* organogenesis [67,69], by controlled expression of distinct CDKIs in a dynamic, stepwise and lineage specific manner for deriving organoid systems. Furthermore, the function of CDKIs in aging cells as inducers of senescence and as tumor suppressors [70,71], could be particularly relevant for

overcoming the current roadblock in regenerative medicine in deriving cell types with adult cell characteristics instead of their immature fetal-like properties [68,72].

One of the main objectives in biomedicine is using hPSCs for differentiating them to functional cell types such as pancreatic beta-cells for biomedical applications, but the low efficiency of cell differentiation has remained a challenge. The identified positive feedback circuitries between SMAD2/3, CDKIs and developmental transcription factors such as EOMES could be beneficial for spontaneous self-organisation of cells and tissue formation. To examine the practical utility of our findings, we differentiated hPSCs to endoderm and further to pancreatic lineage by manipulating the activity of p18 and CDK4/6 with a small compound that impacted the different stages and efficiency of pancreatic insulin-producing beta-cell differentiation. Hence, our findings have practical significance for more efficient generation of hIPSC-derived cell types for biomedical applications. This could be particularly useful in regenerative medicine and cell replacement therapies by differentiating patient-derived iPSCs to functional cell types such as insulin-producing beta-cells for Type I diabetes patients.

In summary, our results unravel the molecular mechanisms by which cell fate decisions are controlled by the cell cycle machinery and describe a dynamical coordination of lineage specification with cell cycle progression. This knowledge could be utilised for novel strategies in directed differentiation of stem cells toward clinically relevant cell types as well as for more robust production of functional cells for drug screening and disease modelling. Furthermore, given the function of CDKIs as tumor suppressors, the identification of the regulatory machinery controlling CDKI expression suggests candidate pathways to the deregulation of the cell cycle in tumorigenic processes and cancer stem cells.

Declarations

Ethical Approval and Consent to participate

The use of H9 hPSCs was consented by WiCell.

Consent for publication

The use of H9 hPSCs was consented by WiCell.

Availability of supporting data

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Siim Pauklin (siim.pauklin@ndorms.ox.ac.uk).

Newly generated materials associated with the paper should be requested by contacting the lead contact. Data and code availability:

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

Competing interests

The authors declare no competing interests.

Funding

SP was supported by a Cancer Research UK Career Development Fellowship, Grant ID C59392/A25064, and Federation of European Biochemical Societies long-term fellowship. SWC

was supported by the High-Level Talent Recruitment Programme for Academic and Research Platform Construction (Reference Number: 5000105) from Wenzhou-Kean University.

Authors' contributions

SM, RN, MP performed experiments, analysed the data, and contributed to the writing of the manuscript. SP conceived the research, performed experiments, analysed the data, supervised the study and wrote the manuscript.

Acknowledgements.

We are grateful to Prof Udo Oppermann, Dr Martin Philpott, Dr James Dunford, Prof Ludovic Vallier, Dr Inmaculada Moreno and Prof Deborah Burks for help and support. This work was supported by a Cancer Research UK Career Development Fellowship, Grant ID C59392/A25064, Federation of European Biochemical Societies long-term fellowship (SP). We thank the flow cytometry and cell sorting staff at the NIHR BRC-Cambridge Phenotyping Hub at the University of Cambridge, the staff at the Botnar Research Centre at the University of Oxford, and the Flow Cytometry Facility at the MRC Weatherall Institute of Molecular Medicine at the University of Oxford.

Authors' information

There are no relevant additional information for authors.

Materials and Methods.

Detailed description of experimental procedures is included in Supplemental Information.

References

1. Fuchs, E. (2009). The tortoise and the hair: slow-cycling cells in the stem cell race. Cell *137*, 811-819. 10.1016/j.cell.2009.05.002.
2. Lange, C., and Calegari, F. (2010). Cdks and cyclins link G1 length and differentiation of embryonic, neural and hematopoietic stem cells. Cell Cycle *9*, 1893-1900.
3. Li, L., and Clevers, H. (2010). Coexistence of quiescent and active adult stem cells in mammals. Science *327*, 542-545. 10.1126/science.1180794.
4. Coronado, D., Godet, M., Bourillot, P.Y., Tapponnier, Y., Bernat, A., Petit, M., Afanassieff, M., Markossian, S., Malashicheva, A., Iacone, R., et al. (2013). A short G1 phase is an intrinsic determinant of naive embryonic stem cell pluripotency. Stem cell research *10*, 118-131. 10.1016/j.scr.2012.10.004.
5. Savatier, P., Lapillonne, H., van Grunsven, L.A., Rudkin, B.B., and Samarut, J. (1996). Withdrawal of differentiation inhibitory activity/leukemia inhibitory factor up-regulates D-type cyclins and cyclin-dependent kinase inhibitors in mouse embryonic stem cells. Oncogene *12*, 309-322.
6. Sherr, C.J., and Roberts, J.M. (1999). CDK inhibitors: positive and negative regulators of G1-phase progression. Genes Dev *13*, 1501-1512.
7. Lim, S., and Kaldis, P. (2013). Cdks, cyclins and CKIs: roles beyond cell cycle regulation. Development *140*, 3079-3093. 10.1242/dev.091744.
8. Polyak, K., Kato, J.Y., Solomon, M.J., Sherr, C.J., Massague, J., Roberts, J.M., and Koff, A. (1994). p27Kip1, a cyclin-Cdk inhibitor, links transforming growth factor-beta and contact inhibition to cell cycle arrest. Genes Dev *8*, 9-22.
9. Besson, A., Hwang, H.C., Cicero, S., Donovan, S.L., Gurian-West, M., Johnson, D., Clurman, B.E., Dyer, M.A., and Roberts, J.M. (2007). Discovery of an oncogenic activity in p27Kip1 that causes stem cell expansion and a multiple tumor phenotype. Genes Dev *21*, 1731-1746. 10.1101/gad.1556607.
10. Toyoshima, H., and Hunter, T. (1994). p27, a novel inhibitor of G1 cyclin-Cdk protein kinase activity, is related to p21. Cell *78*, 67-74.
11. Buchold, G.M., Magyar, P.L., Arumugam, R., Lee, M.M., and O'Brien, D.A. (2007). p19Ink4d and p18Ink4c cyclin-dependent kinase inhibitors in the male reproductive axis. Mol Reprod Dev *74*, 997-1007. 10.1002/mrd.20716.
12. Cunningham, J.J., Levine, E.M., Zindy, F., Goloubeva, O., Roussel, M.F., and Smeyne, R.J. (2002). The cyclin-dependent kinase inhibitors p19(Ink4d) and p27(Kip1) are coexpressed in select retinal cells and act cooperatively to control cell cycle exit. Molecular and cellular neurosciences *19*, 359-374. 10.1006/mcne.2001.1090.
13. Dyer, M.A., and Cepko, C.L. (2001). p27Kip1 and p57Kip2 regulate proliferation in distinct retinal progenitor cell populations. J Neurosci *21*, 4259-4271.
14. Zindy, F., den Besten, W., Chen, B., Rehg, J.E., Latres, E., Barbacid, M., Pollard, J.W., Sherr, C.J., Cohen, P.E., and Roussel, M.F. (2001). Control of spermatogenesis in mice by the cyclin D-dependent kinase inhibitors p18(Ink4c) and p19(Ink4d). Molecular and cellular biology *21*, 3244-3255. 10.1128/MCB.21.9.3244-3255.2001.
15. Franklin, D.S., Godfrey, V.L., Lee, H., Kovalev, G.I., Schoonhoven, R., Chen-Kiang, S., Su, L., and Xiong, Y. (1998). CDK inhibitors p18(INK4c) and p27(Kip1) mediate two separate pathways to collaboratively suppress pituitary tumorigenesis. Genes Dev *12*, 2899-2911.
16. Yan, Y., Frisen, J., Lee, M.H., Massague, J., and Barbacid, M. (1997). Ablation of the CDK inhibitor p57Kip2 results in increased apoptosis and delayed differentiation during mouse development. Genes Dev *11*, 973-983.

17. Zhang, P., Liegeois, N.J., Wong, C., Finegold, M., Hou, H., Thompson, J.C., Silverman, A., Harper, J.W., DePinho, R.A., and Elledge, S.J. (1997). Altered cell differentiation and proliferation in mice lacking p57KIP2 indicates a role in Beckwith-Wiedemann syndrome. Nature *387*, 151-158. 10.1038/387151a0.
18. Li, H., Collado, M., Villasante, A., Matheu, A., Lynch, C.J., Canamero, M., Rizzoti, K., Carneiro, C., Martinez, G., Vidal, A., et al. (2012). p27(Kip1) directly represses Sox2 during embryonic stem cell differentiation. Cell stem cell *11*, 845-852. 10.1016/j.stem.2012.09.014.
19. Marques-Torrejon, M.A., Porlan, E., Banito, A., Gomez-Ibarlucea, E., Lopez-Contreras, A.J., Fernandez-Capetillo, O., Vidal, A., Gil, J., Torres, J., and Farinas, I. (2013). Cyclin-dependent kinase inhibitor p21 controls adult neural stem cell expansion by regulating Sox2 gene expression. Cell stem cell *12*, 88-100. 10.1016/j.stem.2012.12.001.
20. Nguyen, L., Besson, A., Heng, J.I., Schuurmans, C., Teboul, L., Parras, C., Philpott, A., Roberts, J.M., and Guillemot, F. (2006). p27kip1 independently promotes neuronal differentiation and migration in the cerebral cortex. Genes Dev *20*, 1511-1524. 10.1101/gad.377106.
21. Godin, J.D., Thomas, N., Laguesse, S., Malinouskaya, L., Close, P., Malaise, O., Purnelle, A., Raineteau, O., Campbell, K., Fero, M., et al. (2012). p27(Kip1) is a microtubule-associated protein that promotes microtubule polymerization during neuron migration. Developmental cell *23*, 729-744. 10.1016/j.devcel.2012.08.006.
22. Pauklin, S., and Vallier, L. (2013). The cell-cycle state of stem cells determines cell fate propensity. Cell *155*, 135-147. 10.1016/j.cell.2013.08.031.
23. Singh, A.M., Chappell, J., Trost, R., Lin, L., Wang, T., Tang, J., Wu, H., Zhao, S., Jin, P., and Dalton, S. (2013). Cell-cycle control of developmentally regulated transcription factors accounts for heterogeneity in human pluripotent cells. Stem cell reports *1*, 532-544. 10.1016/j.stemcr.2013.10.009.
24. Calder, A., Roth-Albin, I., Bhatia, S., Pilquil, C., Lee, J.H., Bhatia, M., Levadoux-Martin, M., McNicol, J., Russell, J., Collins, T., and Draper, J.S. (2012). Lengthened G1 Phase Indicates Differentiation Status in Human Embryonic Stem Cells. Stem Cells Dev. 10.1089/scd.2012.0168.
25. Gonzales, K.A., Liang, H., Lim, Y.S., Chan, Y.S., Yeo, J.C., Tan, C.P., Gao, B., Le, B., Tan, Z.Y., Low, K.Y., et al. (2015). Deterministic Restriction on Pluripotent State Dissolution by Cell-Cycle Pathways. Cell *162*, 564-579. 10.1016/j.cell.2015.07.001.
26. Madrigal, P., Deng, S., Feng, Y., Militi, S., Goh, K.J., Nibhani, R., Grandy, R., Osnato, A., Ortmann, D., Brown, S., and Pauklin, S. (2023). Epigenetic and transcriptional regulations prime cell fate before division during human pluripotent stem cell differentiation. Nat Commun *14*, 405. 10.1038/s41467-023-36116-9.
27. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043. 10.1126/science.1076997.
28. Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353. nature04733 [pii]
10.1038/nature04733.
29. Brown, M.A., Sims, R.J., 3rd, Gottlieb, P.D., and Tucker, P.W. (2006). Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex. Mol Cancer *5*, 26. 10.1186/1476-4598-5-26.

30. Abu-Farha, M., Lambert, J.P., Al-Madhoun, A.S., Elisma, F., Skerjanc, I.S., and Figeys, D. (2008). The tale of two domains: proteomics and genomics analysis of SMYD2, a new histone methyltransferase. Mol Cell Proteomics *7*, 560-572. 10.1074/mcp.M700271-MCP200.
31. Huang, J., Perez-Burgos, L., Placek, B.J., Sengupta, R., Richter, M., Dorsey, J.A., Kubicek, S., Opravil, S., Jenuwein, T., and Berger, S.L. (2006). Repression of p53 activity by Smyd2-mediated methylation. Nature *444*, 629-632. 10.1038/nature05287.
32. Cho, H.S., Hayami, S., Toyokawa, G., Maejima, K., Yamane, Y., Suzuki, T., Dohmae, N., Kogure, M., Kang, D., Neal, D.E., et al. (2012). RB1 methylation by SMYD2 enhances cell cycle progression through an increase of RB1 phosphorylation. Neoplasia *14*, 476-486. 10.1593/neo.12656.
33. Zwijsen, R.M., Wientjens, E., Klompmaker, R., van der Sman, J., Bernards, R., and Michalides, R.J. (1997). CDK-independent activation of estrogen receptor by cyclin D1. Cell *88*, 405-415.
34. Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L.E., Trotter, M.W., Cho, C.H., Martinez, A., Rugg-Gunn, P., et al. (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. Development *136*, 1339-1349. 10.1242/dev.033951.
35. Brown, S., Teo, A., Pauklin, S., Hannan, N., Cho, C.H., Lim, B., Vardy, L., Dunn, N.R., Trotter, M., Pedersen, R., and Vallier, L. (2011). Activin/Nodal signaling controls divergent transcriptional networks in human embryonic stem cells and in endoderm progenitors. Stem Cells *29*, 1176-1185. 10.1002/stem.666.
36. Pauklin, S., and Vallier, L. (2015). Activin/Nodal signalling in stem cells. Development *142*, 607-619. 10.1242/dev.091769.
37. Blanco, E., Gonzalez-Ramirez, M., Alcaine-Colet, A., Aranda, S., and Di Croce, L. (2020). The Bivalent Genome: Characterization, Structure, and Regulation. Trends in genetics : TIG *36*, 118-131. 10.1016/j.tig.2019.11.004.
38. Tsuneyoshi, N., Tan, E.K., Sadasivam, A., Poobalan, Y., Sumi, T., Nakatsuji, N., Suemori, H., and Dunn, N.R. (2012). The SMAD2/3 corepressor SNON maintains pluripotency through selective repression of mesendodermal genes in human ES cells. Genes Dev *26*, 2471-2476. 10.1101/gad.201772.112.
39. Bertero, A., Madrigal, P., Galli, A., Hubner, N.C., Moreno, I., Burks, D., Brown, S., Pedersen, R.A., Gaffney, D., Mendjan, S., et al. (2015). Activin/nodal signaling and NANOG orchestrate human embryonic stem cell fate decisions by controlling the H3K4me3 chromatin mark. Genes Dev *29*, 702-717. 10.1101/gad.255984.114.
40. Teo, A.K., Arnold, S.J., Trotter, M.W., Brown, S., Ang, L.T., Chng, Z., Robertson, E.J., Dunn, N.R., and Vallier, L. (2011). Pluripotency factors regulate definitive endoderm specification through eomesodermin. Genes Dev *25*, 238-250. gad.607311 [pii]
10.1101/gad.607311.
41. Mullen, A.C., Orlando, D.A., Newman, J.J., Loven, J., Kumar, R.M., Bilodeau, S., Reddy, J., Guenther, M.G., DeKoter, R.P., and Young, R.A. (2011). Master transcription factors determine cell-type-specific responses to TGF-beta signaling. Cell *147*, 565-576. 10.1016/j.cell.2011.08.050.
42. Kretzschmar, M., Doody, J., Timokhina, I., and Massague, J. (1999). A mechanism of repression of TGFbeta/ Smad signaling by oncogenic Ras. Genes Dev *13*, 804-816.
43. Palazuelos, J., Klingener, M., and Aguirre, A. (2014). TGFbeta signaling regulates the timing of CNS myelination by modulating oligodendrocyte progenitor cell cycle exit through SMAD3/4/FoxO1/Sp1. J Neurosci *34*, 7917-7930. 10.1523/JNEUROSCI.0363-14.2014.

44. Podkowa, M., Christova, T., Zhao, X., Jian, Y., and Attisano, L. (2013). p21-Activated kinase (PAK) is required for Bone Morphogenetic Protein (BMP)-induced dendritogenesis in cortical neurons. Molecular and cellular neurosciences *57*, 83-92. 10.1016/j.mcn.2013.10.005.
45. Chang, S.F., Chang, T.K., Peng, H.H., Yeh, Y.T., Lee, D.Y., Yeh, C.R., Zhou, J., Cheng, C.K., Chang, C.A., and Chiu, J.J. (2009). BMP-4 induction of arrest and differentiation of osteoblast-like cells via p21 CIP1 and p27 KIP1 regulation. Mol Endocrinol *23*, 1827-1838. 10.1210/me.2009-0143.
46. Scandura, J.M., Boccuni, P., Massague, J., and Nimer, S.D. (2004). Transforming growth factor beta-induced cell cycle arrest of human hematopoietic cells requires p57KIP2 up-regulation. Proc Natl Acad Sci U S A *101*, 15231-15236. 10.1073/pnas.0406771101.
47. Zhang, J., Espinoza, L.A., Kinders, R.J., Lawrence, S.M., Pfister, T.D., Zhou, M., Veenstra, T.D., Thorgeirsson, S.S., and Jessup, J.M. (2013). NANOG modulates stemness in human colorectal cancer. Oncogene *32*, 4397-4405. 10.1038/onc.2012.461.
48. Jeter, C.R., Liu, B., Liu, X., Chen, X., Liu, C., Calhoun-Davis, T., Repass, J., Zaehres, H., Shen, J.J., and Tang, D.G. (2011). NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. Oncogene *30*, 3833-3845. 10.1038/onc.2011.114.
49. Jeter, C.R., Badeaux, M., Choy, G., Chandra, D., Patrawala, L., Liu, C., Calhoun-Davis, T., Zaehres, H., Daley, G.Q., and Tang, D.G. (2009). Functional evidence that the self-renewal gene NANOG regulates human tumor development. Stem Cells *27*, 993-1005. 10.1002/stem.29.
50. Ibrahim, E.E., Babaei-Jadidi, R., Saadeddin, A., Spencer-Dene, B., Hossaini, S., Abuzinadah, M., Li, N., Fadhil, W., Ilyas, M., Bonnet, D., and Nateri, A.S. (2012). Embryonic NANOG activity defines colorectal cancer stem cells and modulates through AP1- and TCF-dependent mechanisms. Stem Cells *30*, 2076-2087. 10.1002/stem.1182.
51. Zhu, Q., Pearson-White, S., and Luo, K. (2005). Requirement for the SnoN oncoprotein in transforming growth factor beta-induced oncogenic transformation of fibroblast cells. Molecular and cellular biology *25*, 10731-10744. 10.1128/MCB.25.24.10731-10744.2005.
52. Krakowski, A.R., Laboureau, J., Mauviel, A., Bissell, M.J., and Luo, K. (2005). Cytoplasmic SnoN in normal tissues and nonmalignant cells antagonizes TGF-beta signaling by sequestration of the Smad proteins. Proc Natl Acad Sci U S A *102*, 12437-12442. 10.1073/pnas.0504107102.
53. He, J., Tegen, S.B., Krawitz, A.R., Martin, G.S., and Luo, K. (2003). The transforming activity of Ski and SnoN is dependent on their ability to repress the activity of Smad proteins. J Biol Chem *278*, 30540-30547. 10.1074/jbc.M304016200.
54. Edmiston, J.S., Yeudall, W.A., Chung, T.D., and Lebman, D.A. (2005). Inability of transforming growth factor-beta to cause SnoN degradation leads to resistance to transforming growth factor-beta-induced growth arrest in esophageal cancer cells. Cancer Res *65*, 4782-4788. 10.1158/0008-5472.CAN-04-4354.
55. Zindy, F., Quelle, D.E., Roussel, M.F., and Sherr, C.J. (1997). Expression of the p16INK4a tumor suppressor versus other INK4 family members during mouse development and aging. Oncogene *15*, 203-211. 10.1038/sj.onc.1201178.
56. Serrano, M., Lee, H., Chin, L., Cordon-Cardo, C., Beach, D., and DePinho, R.A. (1996). Role of the INK4a locus in tumor suppression and cell mortality. Cell *85*, 27-37.
57. Kamijo, T., Zindy, F., Roussel, M.F., Quelle, D.E., Downing, J.R., Ashmun, R.A., Grosveld, G., and Sherr, C.J. (1997). Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. Cell *91*, 649-659.

58. Deng, C., Zhang, P., Harper, J.W., Elledge, S.J., and Leder, P. (1995). Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. Cell *82*, 675-684.
59. Luedde, T., Rodriguez, M.E., Tacke, F., Xiong, Y., Brenner, D.A., and Trautwein, C. (2003). p18(INK4c) collaborates with other CDK-inhibitory proteins in the regenerating liver. Hepatology *37*, 833-841. 10.1053/jhep.2003.50136.
60. Awad, M.M., Sanders, J.A., and Gruppuso, P.A. (2000). A potential role for p15(Ink4b) and p57(Kip2) in liver development. FEBS Lett *483*, 160-164.
61. Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262-1278. 10.1016/j.cell.2014.05.010.
62. Rivera, C.M., and Ren, B. (2013). Mapping human epigenomes. Cell *155*, 39-55. 10.1016/j.cell.2013.09.011.
63. Delavaine, L., and La Thangue, N.B. (1999). Control of E2F activity by p21Waf1/Cip1. Oncogene *18*, 5381-5392. 10.1038/sj.onc.1202923.
64. Pippa, R., Espinosa, L., Gundem, G., Garcia-Escudero, R., Dominguez, A., Orlando, S., Gallastegui, E., Saiz, C., Besson, A., Pujol, M.J., et al. (2012). p27Kip1 represses transcription by direct interaction with p130/E2F4 at the promoters of target genes. Oncogene *31*, 4207-4220. 10.1038/onc.2011.582.
65. Reynaud, E.G., Leibovitch, M.P., Tintignac, L.A., Pelpel, K., Guillier, M., and Leibovitch, S.A. (2000). Stabilization of MyoD by direct binding to p57(Kip2). The Journal of biological chemistry *275*, 18767-18776. 10.1074/jbc.M907412199.
66. Ladewig, J., Koch, P., and Brustle, O. (2013). Leveling Waddington: the emergence of direct programming and the loss of cell fate hierarchies. Nat Rev Mol Cell Biol *14*, 225-236. 10.1038/nrm3543.
67. Sterneckert, J.L., Reinhardt, P., and Scholer, H.R. (2014). Investigating human disease using stem cell models. Nat Rev Genet *15*, 625-639. 10.1038/nrg3764.
68. Tabar, V., and Studer, L. (2014). Pluripotent stem cells in regenerative medicine: challenges and recent progress. Nat Rev Genet *15*, 82-92. 10.1038/nrg3563.
69. Sasai, Y. (2013). Next-generation regenerative medicine: organogenesis from stem cells in 3D culture. Cell stem cell *12*, 520-530. 10.1016/j.stem.2013.04.009.
70. Collado, M., Blasco, M.A., and Serrano, M. (2007). Cellular senescence in cancer and aging. Cell *130*, 223-233. 10.1016/j.cell.2007.07.003.
71. Kim, W.Y., and Sharpless, N.E. (2006). The regulation of INK4/ARF in cancer and aging. Cell *127*, 265-275. 10.1016/j.cell.2006.10.003.
72. Oh, J., Lee, Y.D., and Wagers, A.J. (2014). Stem cell aging: mechanisms, regulators and therapeutic opportunities. Nat Med *20*, 870-880. 10.1038/nm.3651.
73. Pauklin, S., Madrigal, P., Bertero, A., and Vallier, L. (2016). Initiation of stem cell differentiation involves cell cycle-dependent regulation of developmental genes by Cyclin D. Genes Dev *30*, 421-433. 10.1101/gad.271452.115.
74. Vallier, L., Touboul, T., Brown, S., Cho, C., Bilican, B., Alexander, M., Cedervall, J., Chandran, S., Ahrlund-Richter, L., Weber, A., and Pedersen, R.A. (2009). Signaling pathways controlling pluripotency and early cell fate decisions of human induced pluripotent stem cells. Stem Cells *27*, 2655-2666. 10.1002/stem.199.

Figure legends.

Figure 1: CDKIs are not expressed in self-renewing hPSCs while rapidly induced during germ layer differentiation with simultaneous changes in cell cycle length. (a) Representative dot blot graph of EdU incorporation and DNA content analysis for determining the cell cycle profile. (b) Endoderm cells have a strongly lengthened G1 phase compared to other germ layers. Cell cycle analysis by EdU incorporation during germ layer differentiation by collecting cells 72 hours after initiating endoderm, mesoderm or neuroectoderm differentiation. (c) CDKIs p14, p15, p16, p18, p21 and p57 are not expressed in hPSCs except for p27 protein. Immunostaining of CDKIs is shown in red colour co-stained with pluripotency markers Tra-1-60 or OCT4 in green colour. CDKI expression is not detectable except for p27 protein in undifferentiated hPSCs. Scale bar, 100μm. (d-e). CDKIs are specifically induced during the germ layer formation. Relative changes in CDKI (d) mRNA by euclidian hierarchical clustering or (e) protein compared to undifferentiated hPSCs. Z-scores in the heat map indicate the differential expression measured in number of standard deviations from the average level across all conditions. (f) CDKIs accumulate in the nucleus upon their induction in differentiating cells. Immunostaining of p57 and NANOG in a mix of pluripotent and differentiating cells at 36h of endoderm differentiation. Scale bar, 10μm. (g) Dot blot graph of unsynchronised FUCCI-hPSCs visualised by flow cytometry. The different cell cycle phases are marked grey lines, early G1 phase cells sorted are marked with red line. (h) CDKI induction is an early event during the initiation of differentiation. Dynamical changes of endoderm gene expression and CDKIs marks on synchronized cells differentiated toward definitive endoderm and analysed by Q-PCR. Statistical analysis was performed by 2-way ANOVA with multiple comparisons with Tukey correction and **** marks adjusted P-value <0.0001, *** is adjusted P-value <0.001, ** is adjusted P-value <0.01, * is adjusted P-value <0.05.

Figure 2: Compound screening identifies SMYD2 and EZH2 as regulators of hPSC self-renewal and differentiation. (a) Schematic representation of the small molecule screening process on hPSCs. (a) Compounds in the epigenetic library. (c-d) EZH2 inhibitors GSK343, UNC1999 and CPI-169 increase the fraction of pluripotent marker negative cells whereas SMYD2 inhibitors PFI-5, LLY-507 and BAY-598 decrease pluripotent marker negative cells compared to DMSO control treatment. (e) Representative colonies of hPSCs treated with EZH2 inhibitors, SMYD2 inhibitors or EZH2 and SMYD2 KD. (f) Treatment of hPSCs with EZH2 inhibitors increases endoderm and mesoderm marker expression. (g) EZH2 KD in hPSCs increases endoderm marker expression. (h) Immunostaining of EOMES and NANOG in EZH2 inhibitor treated hPSCs or EZH2 KD cells. (i) SMYD2 inhibition and knockdown reduces endoderm marker expression upon endoderm differentiation at 36h timepoint. (j) Immunostaining of EOMES and T in hPSCs differentiated to endoderm for 36h and treated with SMYD2 inhibitors or SMYD2 KD. (k) Several CDKIs are upregulated upon EZH2 inhibition in hPSCs. (l) EZH2 KD upregulates p15, p18 and p57 expression in hPSCs. (m) SMYD2 chemical inhibition and genetic knockdown downregulates several CDKIs during endoderm differentiation at 36h. (n) Schematic depiction of the effects of SMYD2 and EZH2 on pluripotency and differentiation in hPSCs. Statistical analysis was performed by 2-way ANOVA with multiple comparisons with Tukey correction and **** marks adjusted P-value <0.0001, *** is adjusted P-value <0.001, ** is adjusted P-value <0.01, * is adjusted P-value <0.05.

Figure 3: ACTIVIN/NODAL signalling establishes a stem cell-specific expression pattern of CDKIs with poised chromatin through a cooperation between NANOG/OCT4, SMAD2/3, EZH2 and SNON. (a) CDKI expression is distinctly regulated by developmental signalling pathways. Cells at day 2 endoderm were incubated with differentiation media lacking either ACTIVIN A, FGF2, BMP4 or PI3K and with the respective pathway inhibitors. Significant differences compared to OE GFP

compared to AFLyB sample and calculated by t-test are marked. Summary of CDKI regulation by developmental signalling pathways is shown in the table. + induction, – repression, = no change. (b) SMAD2/3, OCT4, NANOG, SOX2 and EZH2 bind to CDKI loci in human pluripotent cells. Histone H3K4me3 and H3K27me3 mark CDKI loci in pluripotent cells. SNON data represents leukemia cells. (c-d) OCT4 and NANOG form a complex with SMAD2/3 on CDKI loci in hPSCs. Sequential ChIP of (c) OCT4 and SMAD2/3 or (d) NANOG and SMAD2/3 in hPSC was performed and analysed by Q-PCR. Significant differences compared to IgG/SMAD2/3 sequential ChIP sample calculated by t-test are marked. (e) OCT4 knockdown results in reduced SMAD2/3, EZH2 and SNON binding on CDKI loci. ChIP of OCT4, SMAD2/3 and SNON was performed in Scramble and OCT4 KD cells to test their presence on the regions uncovered by genome-wide SMAD2/3 and OCT4 ChIP-seq experiments. Significant differences calculated by t-test are marked. NS – not significant. (f) immunostaining of p18 protein expression. Scale bar, 100μm. (g) OCT4 knockdown causes an increase in CDKIs except for p27 in hPSCs. Scramble and OCT4 KD cells were analysed by Q-PCR for determining the expression of CDKIs. (h) OCT4 knockdown causes an increase in CDKIs except for p27 in hPSCs. Scramble and OCT4 KD cells were analysed by Q-PCR for determining the expression of CDKIs. Significant differences calculated by t-test are marked. NS – not significant. (i) OCT4 knockdown causes a decrease in repressive bivalent mark H3K27me3 on p15, p18, p21 and p57 loci. Scramble and NANOG KD cells were analysed by ChIP-QPCR of H3K4me3 and H3K27me3 marks on CDKI loci. Significant differences calculated by two-way ANOVA are marked. (j) OCT4 and NANOG represses p15, p18 and p57 promoters. hPSCs were cotransfected with CDKI promoter-luciferase constructs and NANOG and/or SNON overexpression constructs. Luciferase signals were analysed 48h after transfection. Experiments represent three replicates. Statistical analysis was performed by 2-way ANOVA with multiple comparisons with Tukey correction and **** marks adjusted P-value <0.0001, *** is adjusted P-value <0.001, ** is adjusted P-value <0.01, * is adjusted P-value <0.05.

Figure 4: CDKIs are induced during endoderm differentiation by SMYD2-SMAD2/3-EOMES transcriptional complex. (a) The location of EOMES binding sites in the proximity of CKI loci relative to their transcription start site. Putative EOMES binding sites were uncovered by genome-wide analysis of EOMES binding in day 2 endoderm cells. (b) EOMES and SMAD2/3 form a complex on CDKI loci in endoderm cells. Sequential ChIP of EOMES and SMAD2/3 in day 2 endoderm was followed by Q-PCR analysis. Significant differences compared to IgG/SMAD2/3 sequential ChIP sample and calculated by t-test are marked. (c) EOMES knockdown in endoderm causes the loss of p18 expression. Immunostaining of p18 and EOMES in Scramble and EOMES KD cells differentiated to endoderm for 2 days. Scale bar, 20μm. (d) EOMES knockdown causes the loss of CDKI expression in endoderm cells. EOMES KD and Scramble cells were differentiated to endoderm and analysed by Q-PCR for determining CDKI expression. Significant differences compared to Scramble shRNA sample and calculated by t-test are marked. (e) EOMES induces CDKI expression via binding to regulatory regions on CDKI loci. CDKI promoter-luciferase construct harbouring putative EOMES binding regions were co-transfected with EOMES-expressing vector or GFP into day 2 endoderm cells and analysed 24h after transfection. Significant differences compared to OE GFP sample and calculated by t-test are marked. (f) SMAD2/3 binding to p15, p18 and p57 loci in endoderm depends on EOMES. SMAD2/3 and EOMES CHIP was performed in Scramble and EOMES KD cells in day 2 endoderm cells and analysed by Q-PCR. Significant differences compared to Scramble shRNA sample and calculated by t-test are marked. (g) Blocking SMAD2/3 binding to CDKI loci by inhibiting ACTIVIN/NODAL signalling with SB431542 (SB) does not abolish EOMES binding. Day 2 endoderm cells were treated with SB431542 for 2h and analysed by SMAD2/3 ChIP or EOMES ChIP. Significant differences compared to not SB431542 treated endoderm cells and calculated by t-test are marked. (h-j) Timeline of SMAD2/3, NANOG and EOMES binding to p18,

p15 and p57 loci during endoderm differentiation. (k) EOMES knockdown in endoderm cells results in the loss of activating histone mark H3K4me3 on CDKI loci and elevation of the repressing histone modification H3K27me3. Scramble and EOMES KD cells were differentiated to day 2 endoderm and analysed by H3K4me3 ChIP and H3K27me3 ChIP followed by Q-PCR. Significant differences compared to Scramble shRNA sample and calculated by t-test are marked. (l) SMYD2 binding to p15, p18 and p57 loci is reduced by ACTIVIN/NODAL signalling inhibitor SB431542. (m) EOMES and SMYD2 form a complex on CDKI loci in endoderm cells. Sequential ChIP of EOMES and SMYD2 in day 2 endoderm was followed by Q-PCR analysis. Significant differences compared to IgG/SMYD2 sequential ChIP sample and calculated by t-test are marked. (m) SMYD2 binding to CDKI loci is reduced upon EOMES KD in endoderm. Significant differences compared to Scramble shRNA sample and calculated by t-test are marked. (o) Graphical model depicting the absence of CKI expression in pluripotent cells and induction in endoderm cells by SMAD2/3 and EOMES. During endoderm differentiation, EOMES-SMAD2/3 complex binds to p18, p15 and p57 loci and induces their expression along with the endoderm genes. All data are shown as mean±s.d. (n=3).

Figure 5: Loss of CDKI expression affects the differentiation efficiency toward distinct cell fates. (a) Overexpression of CDKI shRNA does not disturb pluripotency and self-renewal of hPSCs except for p27. Representative colonies for each CDKI KD cell line. (b) p15, p18 and p57 knockdown reduces the efficiency of endoderm differentiation while p21 knockdown reduces the efficiency of neuroectoderm differentiation. Q-PCR analysis of pluripotency genes and germ layer specific markers in CKI KD cells differentiated toward endoderm and neuroectoderm. Significant differences compared to differentiated Scramble shRNA sample calculated by t-test are marked. (c-d) CDKI knockdown results in germ layer-specific effects on differentiation. Flow cytometry analysis of differentiation markers during (C) endoderm and (D) neuroectoderm differentiation.

Significant differences compared to differentiated Scramble shRNA sample calculated by t-test are marked. (e) p15, p18 and p57 knockdown cells have decreased endoderm specification capacity. Immunostaining of pluripotency and differentiation markers in Scramble and CDKI KD cells. Scale bar, 100μm. (f-g) Teratoma assays of CDKI knockdown in hESCs indicate changed differentiation efficiencies upon CDKI knockdown. (f) Schematic overview of teratoma assays. (g) Representative section of teratomas from Scramble, p18 KD, p21 KD and p57 KD cells. Undifferentiated cells were injected into the testes and allowed to form teratomas for 3 months before histological analysis.

Figure 6: Overexpression of CDKIs causes specific changes in germ layer formation capacity and extension of G1 phase length. (a) The number of colonies derived by stable expression is severely reduced in cells overexpressing CDKIs compared to OE GFP cells. Significant differences compared to OE GFP calculated by t-test are marked. (b) Overexpression of CDKI protein causes a loss of pluripotency and elevated differentiation morphology of hESCs as shown in representative colonies of OE CDKIs. (c) Graphical overview of cell cycle analysis in CDKI OE cells. (d) p18, p15 and p57 overexpression in hPSCs results in the lengthening of G1 phase. Cell cycle was analysed by EdU incorporation and flow cytometry. (e-f) Overexpression of p15, p18 and p57 induce endoderm differentiation. Differentiation markers were analysed by (e) immunostaining or (f) mRNA expression of pluripotency and differentiation markers in OE GFP and OE CDKI cells. Scale bar, 100μm. Significant differences compared to OE GFP calculated by t-test are marked. (g-h) CDKI overexpression results in germ layer-specific effects during differentiation. Flow cytometry analysis of differentiation markers during (g) endoderm and (h) neuroectoderm differentiation. Significant differences compared to OE vector sample calculated by t-test are marked.

Figure 7: Positive feedback loops between SMAD2/3-CDKI and developmental transcription factors can be harnessed for guiding tissue self-formation for biomedical applications. (a) Analysis of cell cycle dependent differentiation potential in FUCCI-hPSCs. Cells were transfected with CDKI expression constructs and synchronised by cell sorting to early G1 phase 48 hours after transfection. Thereafter, cells were differentiated to endoderm and analysed for endoderm markers. Scale bar, 20μm. (b) CDKI expression in late G1 allows for the induction of endoderm differentiation. Q-PCR analysis of early endoderm marker induction in control cells, p18 OE and p15 OE cells. Significant differences compared to OE vector calculated by t-test are marked. (c) CDKI expression in late G1 phase induces endoderm differentiation. Cells were treated as in (a) and analysed for SOX17 expression by flow cytometry at 24h and 48h time points. Significant differences compared to OE pTP6 vector calculated by two-way ANOVA are marked. (d) Relative abundance of SMAD2/3 in the cytoplasm and on chromatin upon CDKI knockdown or overexpression. Histone H3 and Actin are used as loading controls. (e) CDKIs induce SMAD2/3 transcriptional activity via inhibiting Cyclin D-CDK4/6 mediated phosphorylation of SMAD2/3 in its linker region. Cells were co-transfected with CDKI expressing plasmids and a luciferase construct under the regulation of a consensus SMAD Binding Element. Significant differences calculated by t-test are marked. (f) p18 expression regulates G1 phase length during endoderm differentiation. Cell cycle profile of p18 KD and p18 OE cells was compared to control cells by EdU incorporation. (f) Schematics of hIPSC organoid differentiation toward pancreatic tissue. Each DOX or PD0332991 treatment during the 1-4 specification stages lasted for 72 hours. Markers for each cell identity are shown. (g) QPCR analyses of markers during the pancreatic differentiation and the alternative cell fate markers were analysed at 1-4 specification stages. The treatments for DOX-mediated p18 knockdown and CDK4/6 inhibition with PD0332991 are shown in (g). (i-k) Dynamic PD0332991 treatment improves pancreatic differentiation while constant PD0332991 treatment impairs

differentiation as evidenced by reduced expression and altered spatial distribution of key pancreatic markers PDX1, NGN3 and INSULIN. Experiments represent three replicates. Statistical analysis was performed by 2-way ANOVA with multiple comparisons with Tukey correction and **** marks adjusted P-value <0.0001, *** is adjusted P-value <0.001, ** is adjusted P-value <0.01, * is adjusted P-value <0.05.

MATERIALS AND METHODS

Cell culture of hPSCs and FUCCI-hPSCs lines

hPSCs (H9 from WiCell) were grown in defined culture conditions as described previously [72]. H9 cells were passaged weekly using collagenase IV and maintained in chemically defined medium (CDM) supplemented with ACTIVIN A (10 ng/ml) and FGF2 (12 ng/ml). Pluripotent cells were maintained in Chemically Defined Media with BSA (CDM-BSA) supplemented with 10ng/ml recombinant human ACTIVIN A and 12ng/ml recombinant human FGF2 (both from Dr. Marko Hyvonen, Dept. of Biochemistry, University of Cambridge). Cells were passaged every 4-6 days with collagenase IV as clumps of 50-100 cells and dispensed at a density of 100-150 clumps/cm^2. The culture media was replaced 48 hours after the split and then every 24 hours. Alternative culture conditions were used to maintain hPSCs by maintaining cells on Vitronectin (StemCell Technologies)-coated plates in Essential 8 (E8) medium (Life technologies). Cells were passaged every 5-7 days using 0.5 uM EDTA and plated onto fresh vitronectin-coated plates in E8 medium. Medium was refreshed every second day. This change corresponds to modification of protocols in our lab and has no influence on experimental outcomes. The generation of FUCCI-hPSC lines has been described in [18] and are based on the FUCCI system described in [19].

In vitro differentiation of hPSCs

FUCCI-hPSCs were differentiated into endoderm as described previously [9]. Differentiation into endoderm was performed for up to 72 hours with a combination of cytokines as described in [22,73]. For cells sorted by FACS, the cells were collected and immediately placed into the endoderm differentiation media. Endoderm specification was performed in CDM with polyvinyl Alcohol

(CDM-PVA) prepared without insulin and supplemented with 50ng/ml FGF2, 1µM Ly-294002 (Promega), 100ng/ml ACTIVIN A, and 10ng/ml BMP4 (R&D) for 3 days. Alternatively, and for cells grown in E8 medium, H9 cells were plated as single cells onto gelatin/MEF-coated plates in E8 medium supplemented with 10uM Y-27632. The medium was refreshed the next day. Chemically Defined Media with Polyvinyl Alcohol (CDM-PVA) containing 100 ng/ml recombinant ACTIVIN A (CSCR, University of Cambridge), 80 ng/ml FGF2 (R&D Systems), 10 ng/ml BMP4 (CSCR, University of Cambridge), 10 µM LY29004 (Promega), and 3 µM CHIR99021 (Selleck Chemicals) was applied to the cells for 24 hours. The media was then replaced with fresh CDM-PVA supplemented with 100 ng/ml recombinant ACTIVIN A (CSCR, University of Cambridge), 80 ng/ml FGF2 (R&D Systems), 10 ng/ml BMP4 (CSCR, University of Cambridge) and 10 µM LY29004 (Promega). The next day, the media was removed and RPMI media supplemented with 1X B27 (Lifetech), 100 ng/ml ACTIVIN A, 80 ng/ml FGF2 and 1X non-essential amino acids (Lifetech) was added to the cells.

Generating INK4 and KIP/CIP Knockdown Cells

Previously validated shRNA expression vectors from Sigma-Aldrich (Supplementary Table 1) were transfected into H9 hPSCs with lipofectamine 2000 [22] and grown for 3 days. Cells were then cultured in the presence of puromycin until antibiotic-resistant colonies appeared. These were picked and characterized for knockdown efficiency.

Generating INK4 and KIP/CIP overexpressing Cells

For CDKI overexpression, cDNA sequences of p14, p15, p16, p18, p21, p27 and p57 were transferred into a pTP6 vector containing a CAG promoter. GFP and empty vector were used as controls. All inserts were confirmed by sequencing. Vectors were transfected into H9 hPSCs by lipofection [22] and grown for 3 days. Thereafter, cells with a stable integration were selected by

continuous presence of puromycin. Individual clones were picked, propagated and used for subsequent analyses.

qPCR and Immunostaining

Methods for Q-PCR and immunostaining have been described previously [74]. Q-PCR data are presented as the mean of three independent experiments and error bars indicate standard deviations. Antibodies and primer sequences have been listed in Supplementary Table 2 and 3.

Knockin hPSCs for OCT4

H9 hPSCs with OCT4-eGFP knockin were obtained from WiCell.

The small molecule screening library

The screening library contained concentrated small molecule compounds with verified biochemical activity against their targets. Most of the compounds target epigenetic regulators with high specificity (Supplementary Table 4).

Screening of the chemical compounds

The cells were grown in 96-well plates in standard growth medium with puromycin (1 µg/ml stock). Three technical replicates and three biological replicates were used for the screening. Cells were plated at a concentration of 10,000 cells in 100 µl of media per well in a 96-well plate. One day after plating the cells, the medium was changed to 90 µl standard growth medium supplemented with puromycin (0.5 µg/ml) and ACTIVIN A (10 ng/ml). On the same day, the compounds were added: first, 100x compound library dilutions were made, and 10uL of 100x diluted chemical was added to each well to obtain 1000x final dilution of the compounds. Cells were then cultured with

chemical compounds for five days with media change at day 0, day 2 and day 4 supplemented by fresh compounds. Each replicate was analyzed using Celigo Image Cytometer (Nexcelom) and flow cytometry. Cells were lifted and dissociated into single cells with Trypsin. Details on the antibodies that were used for flow cytometry are listed in Supplementary Table 5. The cells were incubated with 0.5 ug /ml final concentration of conjugated antibodies in 1% BSA-PBS for 40 minutes on ice and washing was repeated as before. The cells were then suspended in 300 uL 1% BSA-PBS with DAPI (1:2000) for live/dead separation and kept on ice to be used for the flow cytometry analysis.

Gene knockdown and overexpression

The stable knockdown and overexpression of SNON was performed by using previously published constructs [38] kindly provided by Prof. Ray Dunn at the Institute of Medical Biology, A*STAR (Agency for Science, Technology, and Research), Singapore. shRNA plasmid DNA was transfected into cells with Lipofectamine 3000 (Thermo Fischer Scientific) according to manufacturer guidelines. Puromycin was added to the growth media at 0.1ug/ml concentration and individual colonies were picked, expanded, and screened for gene knockdown compared to Scramble control transfected cells.

Nucleic acid extraction from cell lines

RNA was extracted using Direct-zol (TM) RNA extraction kit according to manufacturer protocol (Cambridge Bioscience, R2052). The quality of the RNA samples was verified using an RNA screen tape on a Tape-Station (Agilent). The RIN values for all samples were >7.5.

RNA Isolation and cDNA synthesis

Total RNA was isolated by RNeasy RNA Extraction Kit (Qiagen) according to manufacturer's guidelines. RNA was then eluted in 30µl of water and the concentration was measured using Nanodrop. The master mix was prepared as follows: 8µl 5x First-Strand Buffer (Invitrogen), 0.5µl Random primers (0.5 ug/ml) (Promega Cat. C1181), 1ul dNTP mix (10 mM each) (Promega Cat.U1515), 2 ul 0.1 M DTT, 0.5 ul RNase Out, 0.25 ul Superscript III Reverse Transcriptase (Life Technologies). 500 ng of total RNA into a separate tube with 11.75 ul RNase-free water. RNA was heated to 65 ^{o}C for 5 min and allowed to chill on ice for 2 min. 8.25 ul of the master mix were added to RNA. The reaction was incubated at 25 ^{o}C for 10 min and then at 42^{o}C for 50 min. The reaction was then inactivated by heating at 70 for 15 min.

RT-qPCR

2ng of synthesized cDNA was added to 5µl Power SYBR Mix (Life Tecnologies, 4368708 (Master Mix)) and 1.5µl 2µM of forward and reverse primers. RT-qPCR was performed on ViiA 7 machine with the following intervals: denaturation (95 ^{o}C) for 15s and a total of 40 cycles, annealing/extension (60 ^{o}C) for 60s, final extension (60 ^{o}C) for 10 minutes.

Flow cytometry for cell cycle analysis

Cells were collected and analysed using Fortessa (BD Bioscience). Passaging was performed a day 5, after which cells were plated again in spheroid conditions, with the same initial density of 5,000 cells / 1 mL medium. Compounds were added in day 7, and the treatment lasted for 72 hours. Experiment was performed in 3 replicates. The data was analysed in FlowJo.

Western blot analysis

Protein was isolated by lysing cells with RIPA Buffer (Sigma-Aldrich) supplemented by cOmplete EDTA-free protease inhibitor (Roche) and PhosSTOP™ (Sigma-Aldrich) and extracting the supernatant after high-speed centrifugation at 4°C. Protein quantification was performed using the Pierce BCA Protein Assay kit following the manufacturer's protocol. Isolated proteins were prepared for SDS-PAGE separation by dilution with 4× NuPAGE Sample buffer (Invitrogen), addition of NuPAGE™ Sample Reducing Agent ((10X), Invitrogen), 95°C for 5 minutes, and cooling. Isolated proteins were then analysed by Western blotting. Protein separation via SDS-PAGE was performed on a NuPAGE 4%–12% or 12% Bis-Tris gel (Life Technologies) with NuPAGE™ MOPS SDS Running Buffer (Life Technologies). Proteins were transferred to a PVDF membrane, blocked with 5% milk in PBS and 0.05% tween 20, probed with protein-specific antibodies, incubated with horseradish peroxidase-conjugated secondary antibodies, and visualized via enhanced chemiluminescence using the SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific). All antibodies (Supplementary Table 6) were diluted in 5% milk in PBS and 0.05% tween 20. Quantification was performed using ImageJ gel analysis tool.

Immunostaining

The immunostaining method has been described previously [22,39,73]. Cells were fixed for 20 minutes at 4°C in PBS 4% PFA (electron microscopy grade), rinsed three times with PBS, and blocked and permeabilized at the same time for 30 minutes at room temperature using PBS with 10% Donkey Serum (Biorad) and 0.1% Triton X-100 (Sigma). Incubation with primary antibodies diluted in PBS 1% Donkey Serum 0.1% Triton X-100 was performed overnight at 4°C. Samples were washed three times with PBS, and then incubated with AlexaFluor secondary antibodies for 1 hour at room temperature protected from light. Cells were finally washed three times with PBS, and Hoechst

(Sigma) was added to the first wash to stain nuclei. Images were acquired using a LSM 700 confocal microscope (Leica).

Chromatin Immunoprecipitation (ChIP)

All steps were performed on ice or at 4°C and ice-cold buffers and PBS were supplemented with 1mg/ml Leupeptin, 0.2mM PMSF, and 10mM NaButyrate were used unless otherwise stated. Approximately $5x10^6$ cells were used per sample and cross-linked with 1% formaldehyde for 15 minutes. Cross-linking was stopped by incubating samples with glycine at a final concentration of 0.125M for 5 minutes at room temperature, and the cells were washed with PBS followed by pelleting at 250g for 5 minutes. The pellet was re-suspended in 2ml ChIP Cell Lysis Buffer (CLB: 10 mM Tris pH8, 10 mM NaCl, 0.2% NP-40) and incubated for 10 minutes to lyse the plasma membranes. Nuclei were pelleted at 600g for 5 min, lysed in 1.25ml of ChIP Nuclear Lysis Buffer (NLB: 50 mM Tris pH8, 10mM EDTA, 1% SDS) for 10 minutes, and then 0.75ml of ChIP Dilution Buffer (DB: 20 mM Tris pH8, 2mM EDTA, 150mM NaCl, 0.01% SDS, 1% Triton X-100) was added to the samples. Chromatin was sonicated in 15ml Diagenode Bioruptor Pico water bath sonicator with an automated water cooling system, by performing 30 cycles of 30 seconds ON, 45 seconds OFF. This protocol resulted in the homogeneous generation of fragments of 100-400bp. Samples were clarified by centrifugation at 16000g for 10 minutes, and diluted with 3.5ml of DB. After pre-clearing with 10µg of non-immune IgG for 1h and 50µl of Protein G-Agarose for 2h, ChIP was performed overnight in rotation using specific antibodies (Table S2) or non-immune IgG as a control. After incubation for 1 hour with 30µl of Protein G-Agarose, beads were washed twice with ChIP Washing Buffer 1 (WB1: 20mM Tris pH8, 2mM EDTA, 50mM NaCl, 0.1% SDS, 1% Triton X-100), once with ChIP Washing Buffer 2 (WB2: 10mM Tris pH8, 1mM EDTA, 0.25M LiCl, 1% NP-40, 1% Deoxycholic acid), and twice with Tris-EDTA (TE: 10mM Tris pH8, 1mM EDTA). Precipitated DNA

was eluted with 150µl of ChIP Elution Buffer (EB: 100mM $NaHCO_3$) twice for 15 minutes at room temperature in rotation, and processed as follows in parallel with 300 µl of sonicated chromatin non-used for ChIP (Input). Cross-linking was reverted by adding NaCl to a final concentration of 300mM for protein-DNA de-crosslinking and incubated at 65°C for 5 hours and 1 µg RNase A (Sigma) to digest contaminating RNA. Finally, 60 µg of Proteinase K (Sigma) were added overnight at 45°C. DNA was extracted by sequential phenol-chloroform and chloroform extractions, and precipitated overnight at -80°C in 100mM NaAcetate, 66% ethanol and 50µg of glycogen (Ambion) as a carrier. After centrifugation at 16,000g for 1 hour at 4°C, DNA pellets were washed once with ice-cold 70% ethanol, and finally air dried. ChIP samples were resuspended in 30µl and 1:10 of the samples were used in qPCR for verifying the ChIP samples. Primers used for CHIP experiments are listed in Supplementary Table 7.

Cell fractionations.

Cells were harvested with trypsin and washed twice with cold PBS. For cytoplasmic lysis, cells were suspended in 5 times packed cell volume (1 ul PCV = 10^6 cells) equivalent of Isotonic Lysis Buffer (10 mM Tris HCl, pH 7.5, 3 mM CaCl, 2 mM $MgCl_2$, 0.32 M Sucrose, Complete protease inhibitors and phosphatase inhibitors), and incubated for 12 min on ice. Triton X-100 was added to a final concentration of 0.3% and incubated for 3 min. The suspension was centrifuged for 5 min at 1,500 rpm at 4 °C and the supernatant (cytoplasmic fraction) transferred to a fresh chilled tube. For nuclear lysis, nuclear pellets were resuspended in 2 x PCV Nuclear Lysis Buffer+Triton X-100 (50 mM Tris HCl, pH 7.5, 100 mM NaCl, 50 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 10% Glycerol, 0.3% Triton X-100, Complete protease inhibitors and phosphatase inhibitors) and dounce homogenized. The samples were incubated with gentle agitation for 30 min at 4 °C and then centrifuged with a Ti 70.1 rotor at 22,000 rpm for 30 min at 4 °C or with a Ti 45 rotor for 30 min at 20,000 rpm at 4 °C.

The chromatin pellets were dounce homogenized in 2 x PCV Nuclear Lysis Buffer+Triton X-100 and Benzonase until the pellets gave much less resistance. The samples were incubated at RT for 30 min and centrifuged with either a Ti 70.1 rotor for 30 min at 22,000 rpm at 4 °C or with a Ti 45 rotor for 30 min at 20,000 rpm at 4 °C.

Protein co-immunoprecipitation.

Samples were incubated with 5 ug of cross-linked antibodies for 12h at 4 °C. Beads were washed five times with ten bead volumes of Nuclear Lysis Buffer and eluted in SDS western blotting buffer (30 mM Tris pH 6.8, 10% Glycerol, 2% SDS, 0.36 M beta-mercaptoethanol (Sigma), 0.02% bromophenol blue) by heating at 90 °C for 5 min. Samples were analysed by standard western blotting techniques.

Flow cytometry.

Flow cytometry was carried out with a BD MoFlo flow cytometer and analysed by FloJo software. Cell cycle distribution was analysed by Click-It EdU incorporation Kit (Invitrogen) according to manufacturer's guidelines. Marker expression was analysed at various timepoints during differentiation by first dissociating cells into single cells with Cell Dissociation Buffer (Gibco) and fixing in 4% PFA for 20 min at 4°C. This was followed by permeabilisation and blocking with 10% serum + 0.1% Triton X-100 in PBS for 30 min at RT and incubation with primary antibody in 1% serum + 0.1% Triton X-100 for 2h at 4°C. After washing the samples three times with PBS, they were incubated with a secondary antibody for 2h at 4°C, washed three times with PBS and analysed by flow cytometry.

Cell sorting by FACS

FACS on FUCCI-hPSCs was performed as described before [18,19]. hPSCs were washed with PBS and detached from the plate by incubating them for 10 min at 37 °C in Cell Dissociation Buffer (Gibco). Cells were then washed with cold filter sterilised 1% BSA in PBS, before incubating cells in PBS 1% BSA with Tra-1-60 primary antibody (1:100) and Alexa Fluor 647 donkey α-mouse secondary antibody (1:1000) on ice for 20 min in the dark with occasional gentle mixing. The cells were then washed once with at least 50x pellet volume PBS 1% BSA, resuspended gently in 3ml sterile maintenance media, and subjected to cell sorting by gating Tra-1-60+ cells according to the mAG/mKO2 FUCCI signals for hPSCs or for mAG/mRFP/mKate2 FUCCI signals for PDAC cells. The cell sorting was performed with a BeckmanCoulter MoFlo MLS high-speed cell sorter by using parameters described previously [18], and the cells were sorted directly into collection tubes with 2 ml maintenance media.

Teratoma assays.

One million hESC were injected in the lumen of the testicle of 6 to 8-weeks-old SCID mice, and three animals were injected in each group. After 12 weeks, mice were sacrificed, and the testicles and tumours were dissected and fixed for 48h in Bouins solution (Sigma-Aldrich). The fixed tissues were then paraffin-embedded and processed according to standard procedures. Sections (5µm) were stained with hematoxylin/eosin and subsequently examined under bright-field microscope for the presence of tissues deriving from the three germ layers. Animal procedures were performed in accordance with the local committee on Animal Experimentation at Centro de Investigación Príncipe Felipe.

Luciferase Assay

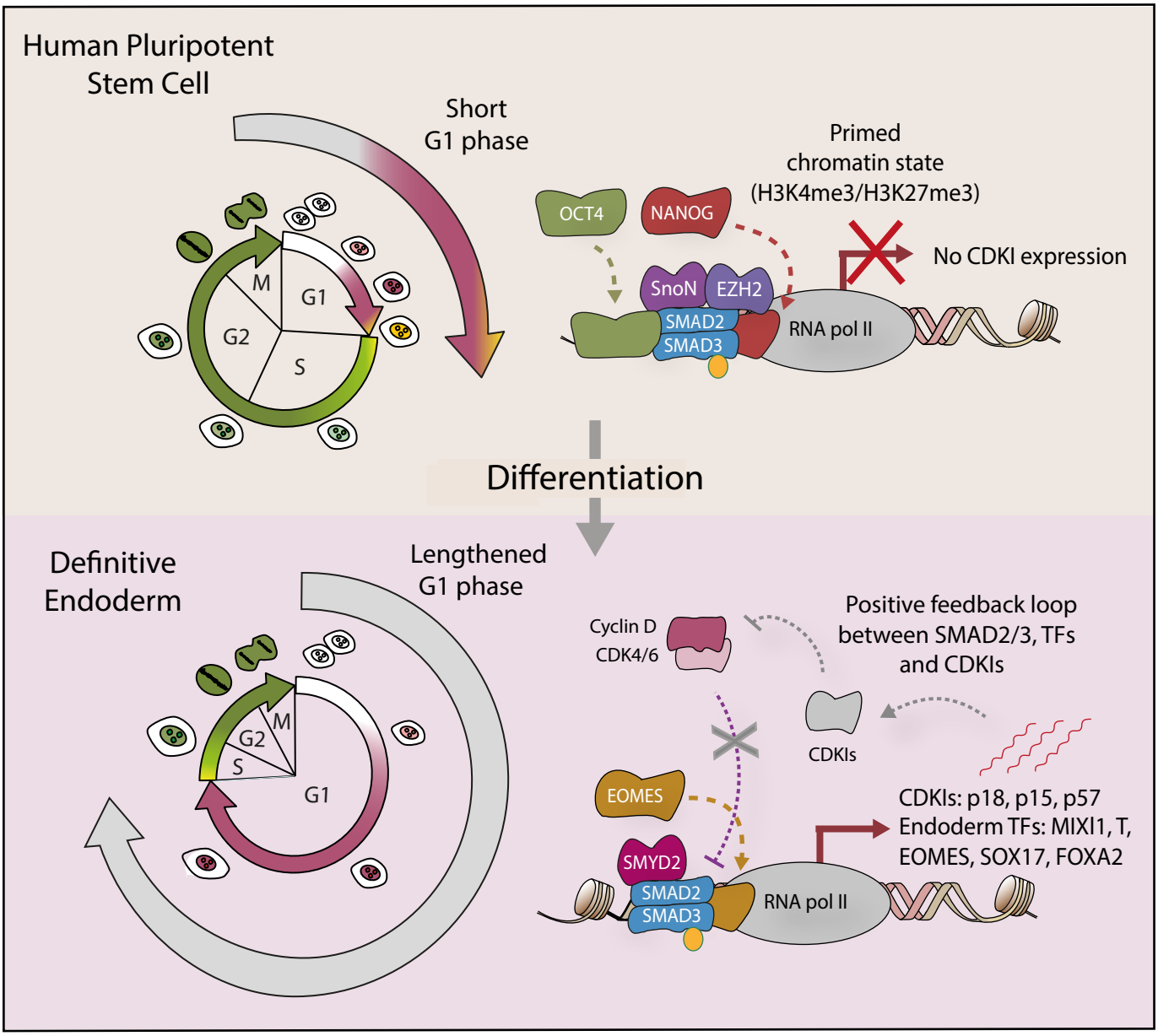
Cells were transfected with a SMAD2/3 reporter construct (SBE4-luciferase), SOX17 or GSC promoter constructs [35] and Renilla luciferase at a ratio of 10:1, using Lipofectamine 2000 (Invitrogen) [22]. Luciferase activity was measured with the dual luciferase assay kit following (Promega) manufacturer instructions. Firefly luciferase activity was normalized to Renilla luciferase activity for cell numbers and transfection efficiency. Samples were analyzed on a Glomax Luminometer and software.

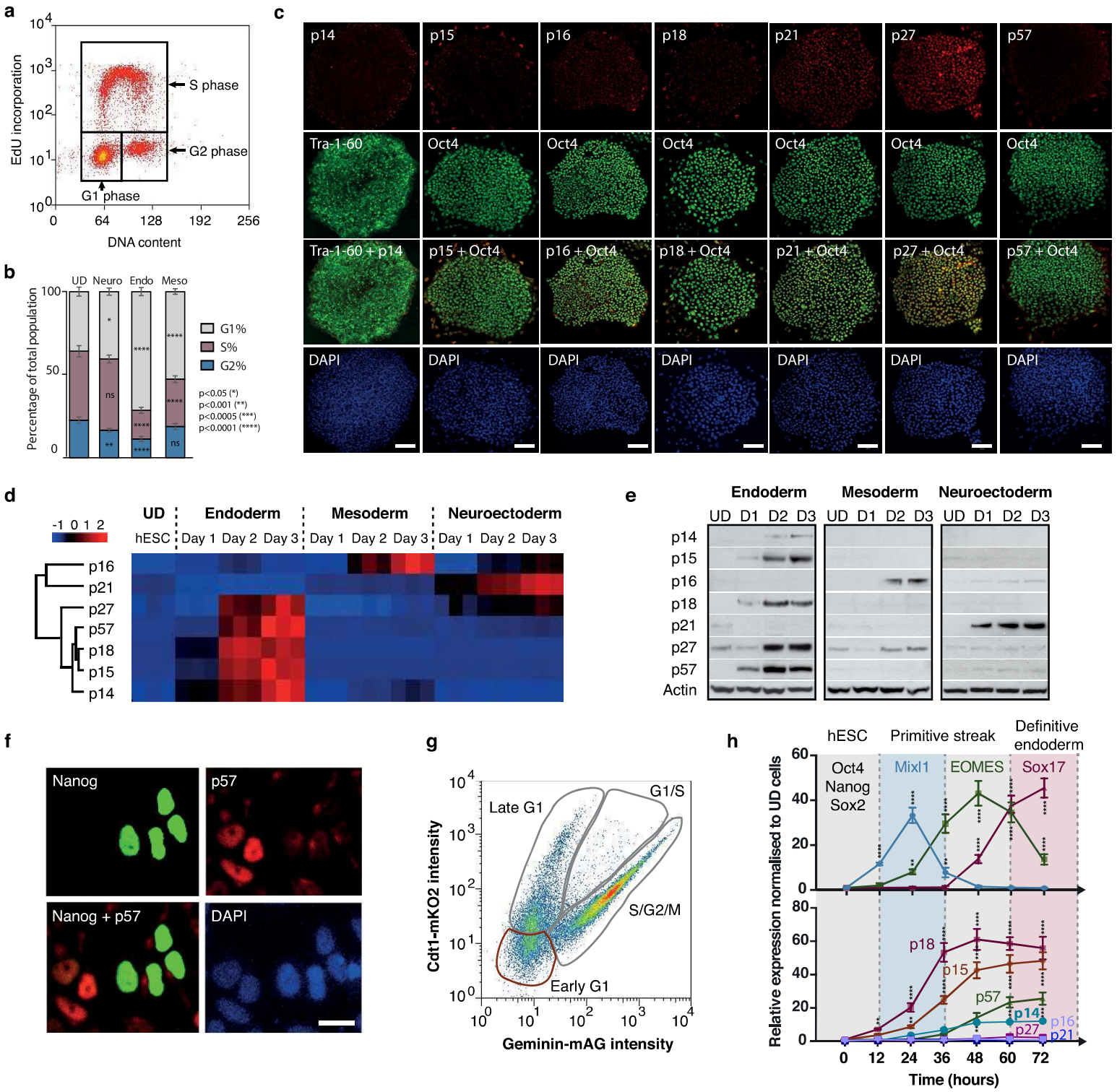
EdU incorporation assay

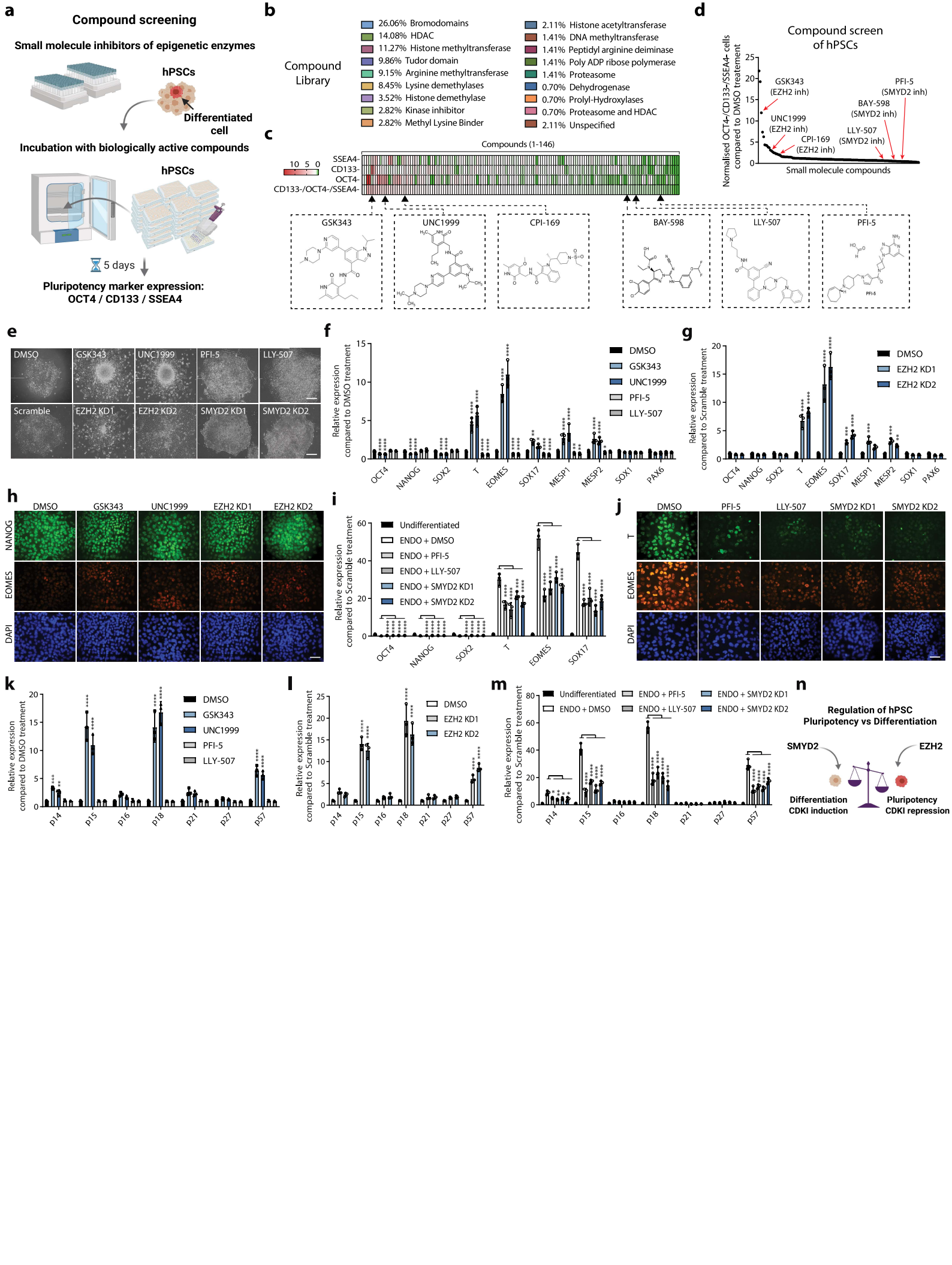
Cell cycle distribution was analysed by Click-It EdU incorporation Kit (Invitrogen) according to manufacturer's guidelines. Flow cytometry was carried out with a BD MoFlo flow cytometer and analysed by FloJo software. Cells were cultured in media collected from cells with different treatment conditions for 72h, replacing the media every 24 hours.

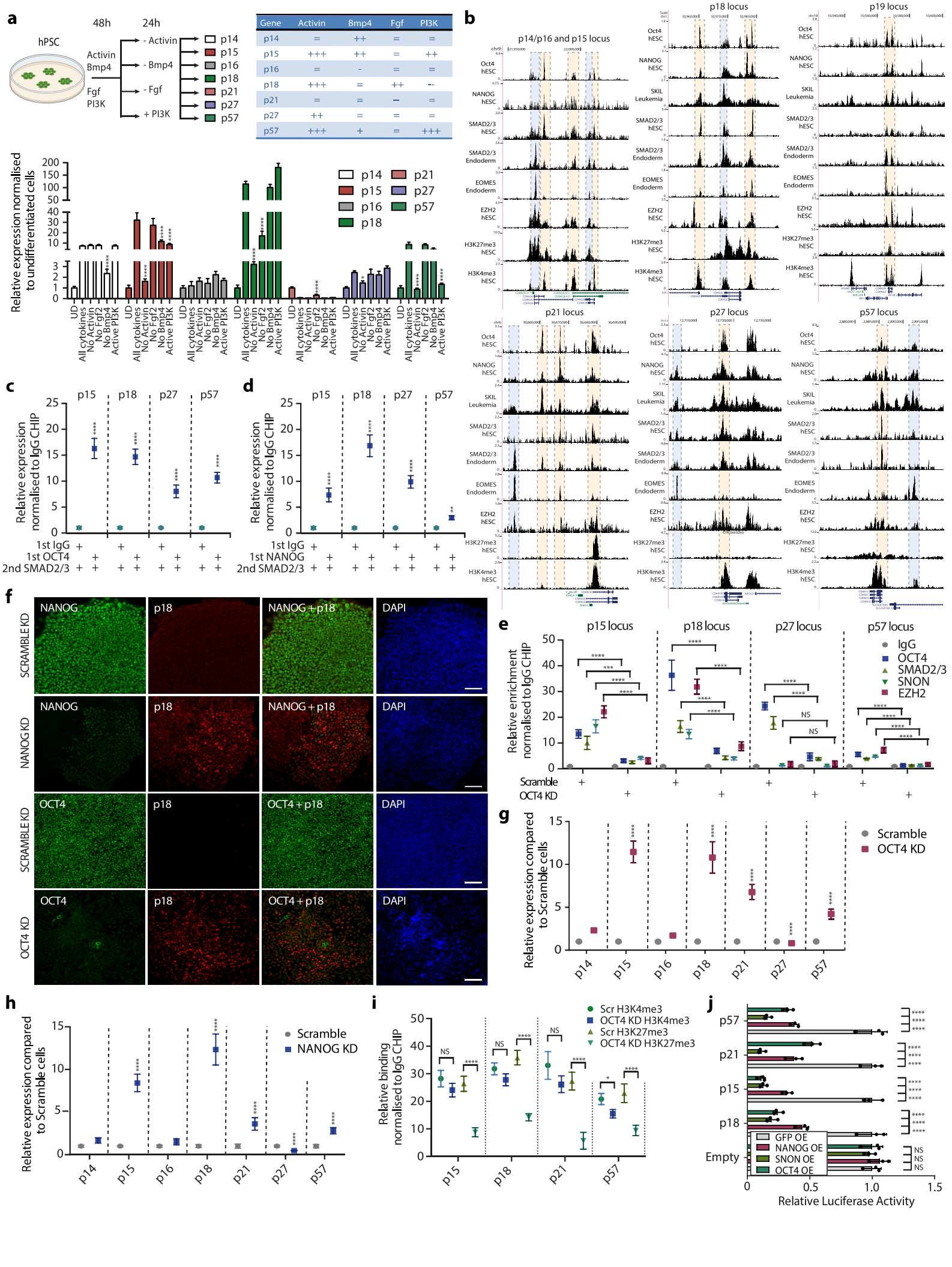
Statistical analysis.

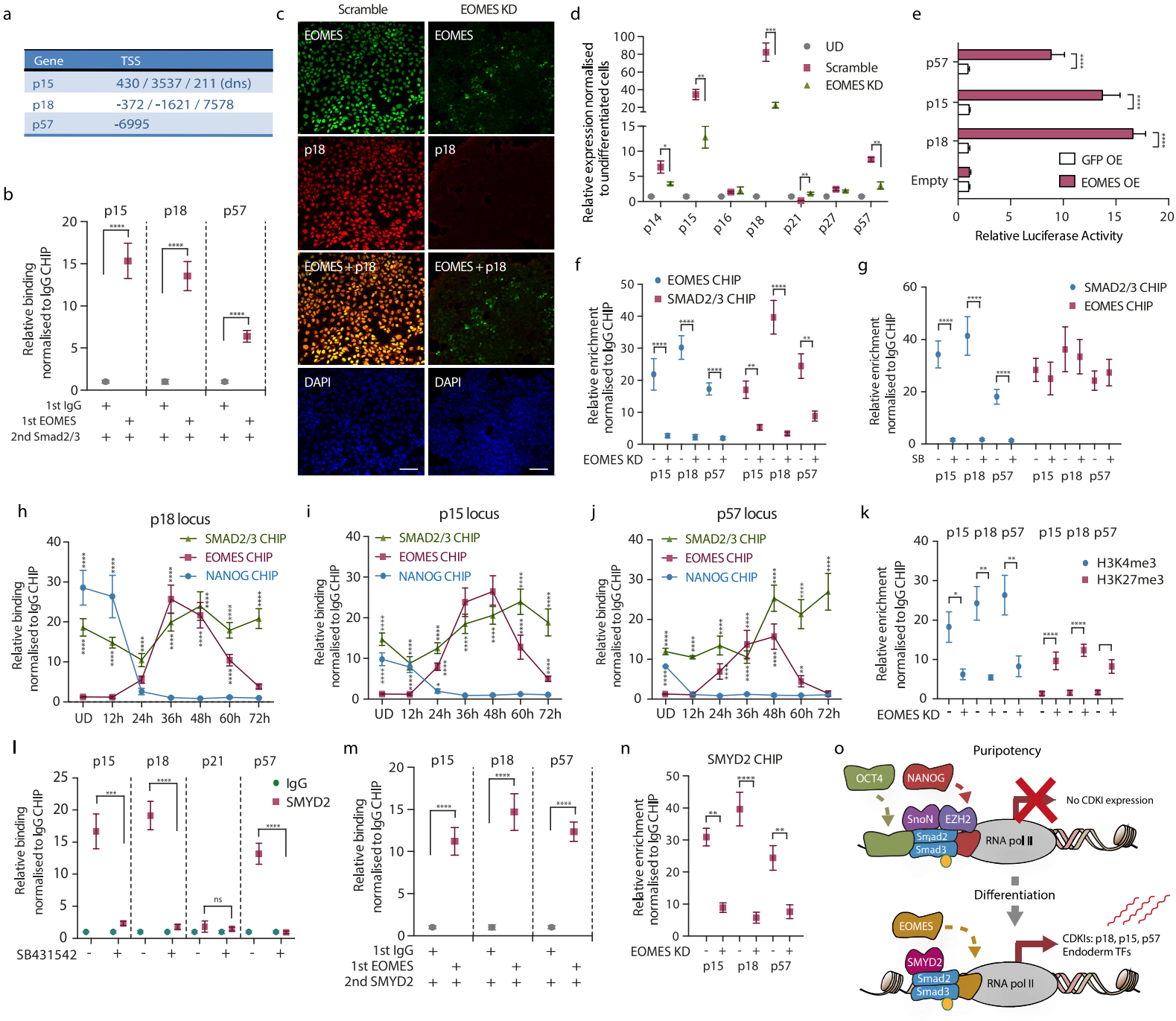
GraphPad Prism 6 was used for statistical analysis by performing t-test and two-way ANOVA tests followed by Bonferroni's corrected multiple comparisons between pairs of conditions. Unless otherwise indicated in the figure legends, we analysed three biological replicates for each data point in all graphs, and the level of significance was as follows: $P< 0.1$ (∗), $P< 0.05$ (∗∗), $P< 0.01$ (∗∗∗), and $P< 0.001$ (∗∗∗∗).

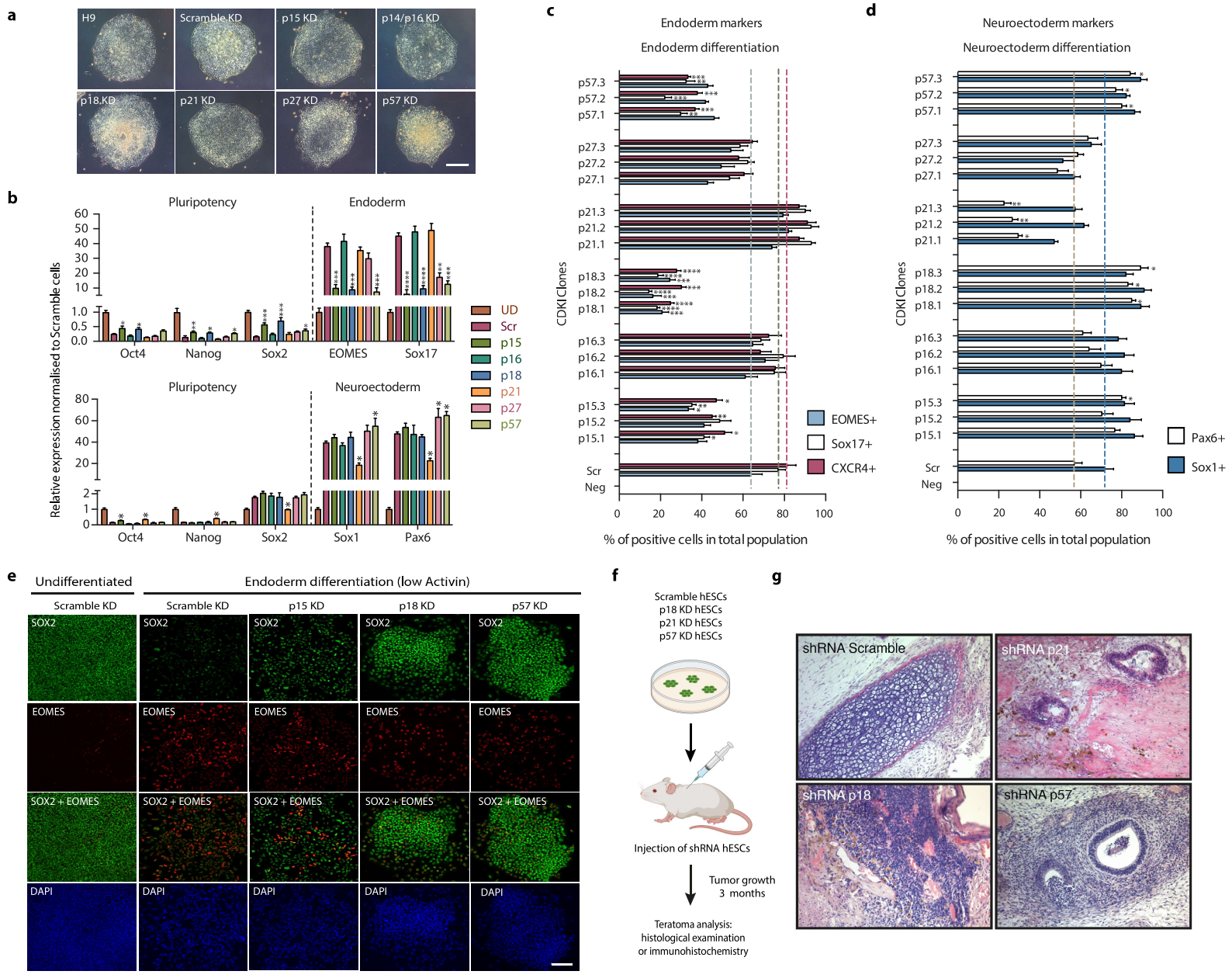

Human Pluripotent Stem Cell
Short G1 phase
M
G1
G2
S
Primed chromatin state (H3K4me3/H3K27me3)
OCT4
NANOG
SnoN
EZH2
SMAD2
SMAD3
RNA pol II
No CDKI expression
Differentiation
Definitive Endoderm
Lengthened G1 phase
M
G2
S
G1
Cyclin D
CDK4/6
Positive feedback loop between SMAD2/3, TFs and CDKIs
CDKIs
EOMES
SMYD2
SMAD2
SMAD3
RNA pol II
CDKIs: p18, p15, p57
Endoderm TFs: MIXl1, T, EOMES, SOX17, FOXA2

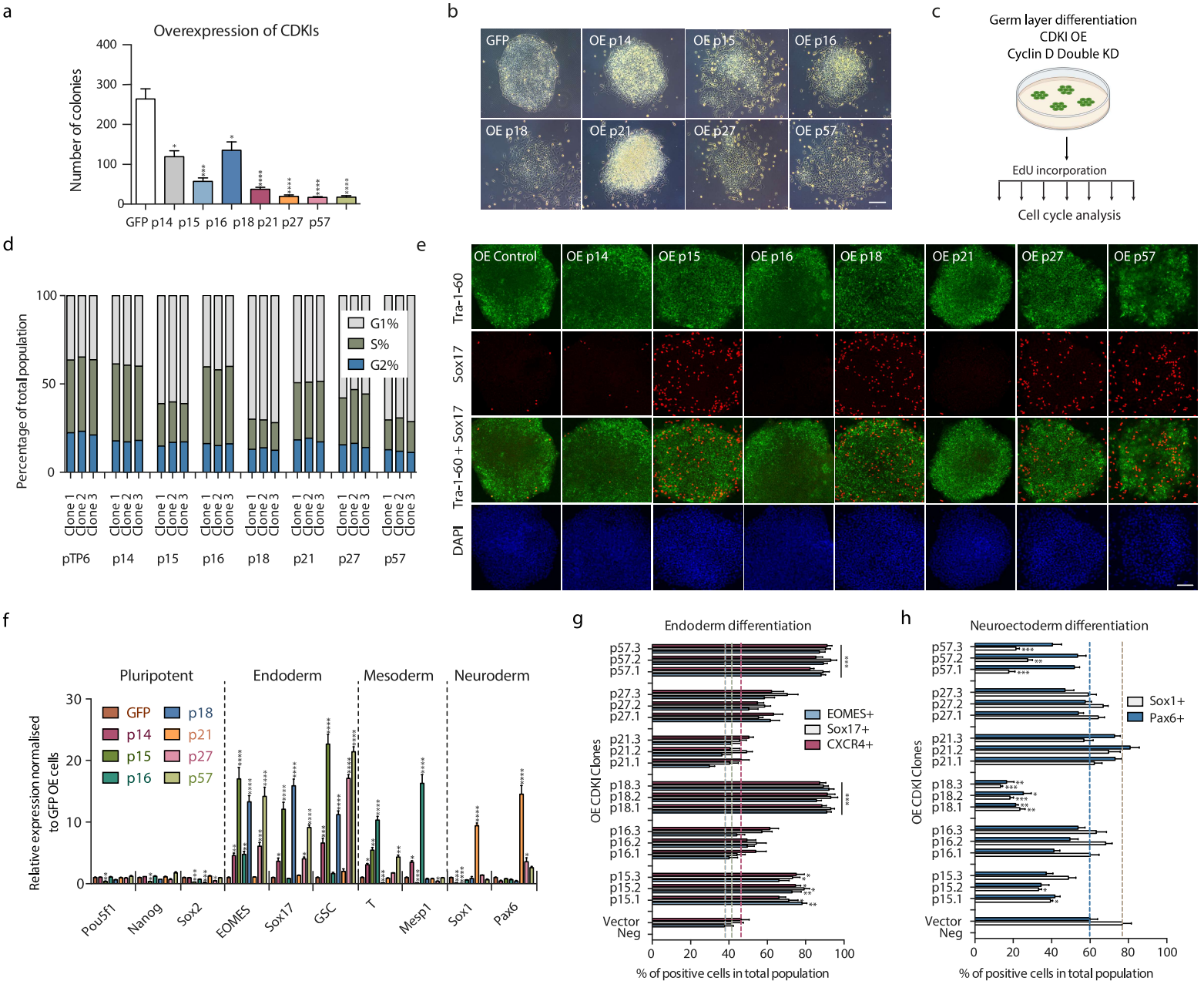

a
EdU incorporation
DNA content
S phase
G2 phase
G1 phase
b
UD Neuro Endo Meso
Percentage of total population
G1%
S%
G2%
p<0.05 (*)
p<0.001 (**)
p<0.0005 (***)
p<0.0001 (****)
c
p14 p15 p16 p18 p21 p27 p57
Tra-1-60 Oct4
Tra-1-60 + p14
p15 + Oct4 p16 + Oct4 p18 + Oct4 p21 + Oct4 p27 + Oct4 p57 + Oct4
DAPI
d
UD Endoderm Mesoderm Neuroectoderm
hESC Day 1 Day 2 Day 3
p16 p21 p27 p57 p18 p15 p14
e
Endoderm Mesoderm Neuroectoderm
UD D1 D2 D3
p14 p15 p16 p18 p21 p27 p57 Actin
f
Nanog
p57
Nanog + p57
DAPI
g
Late G1
G1/S
S/G2/M
Early G1
Cdt1-mKO2 intensity
Geminin-mAG intensity
h
hESC
Primitive streak
Definitive endoderm
Oct4 Nanog Sox2
Mixl1
EOMES
Sox17
p18
p15
p57
p14
p16
p27
p21
Relative expression normalised to UD cells
Time (hours)

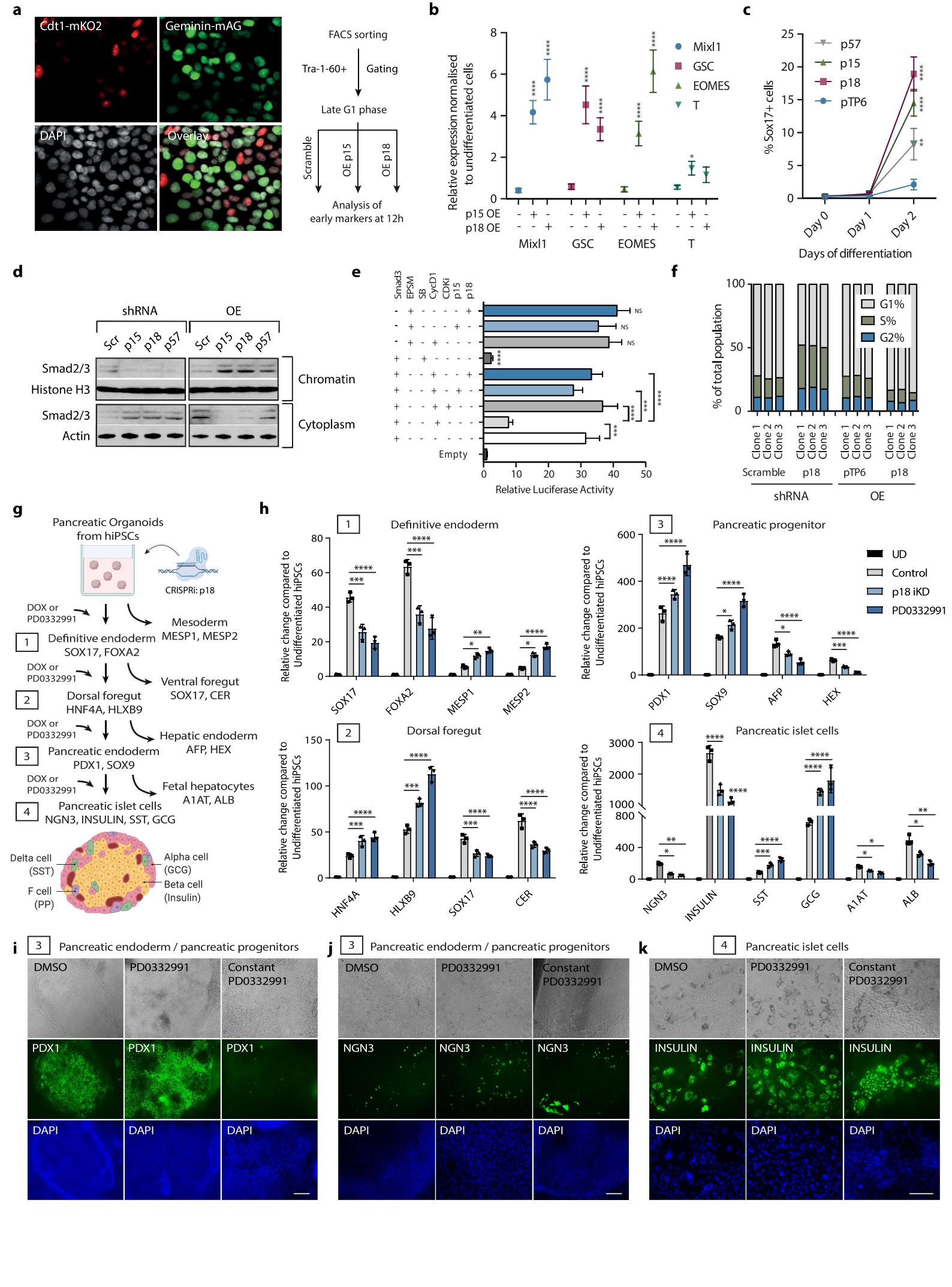

a
Compound screening
Small molecule inhibitors of epigenetic enzymes
hPSCs
Differentiated cell
Incubation with biologically active compounds
hPSCs
5 days
Pluripotency marker expression:
OCT4 / CD133 / SSEA4
b
Compound Library
26.06% Bromodomains
14.08% HDAC
11.27% Histone methyltransferase
9.86% Tudor domain
9.15% Arginine methyltransferase
8.45% Lysine demethylases
3.52% Histone demethylase
2.82% Kinase inhibitor
2.82% Methyl Lysine Binder
2.11% Histone acetyltransferase
1.41% DNA methyltransferase
1.41% Peptidyl arginine deiminase
1.41% Poly ADP ribose polymerase
1.41% Proteasome
0.70% Dehydrogenase
0.70% Prolyl-Hydroxylases
0.70% Proteasome and HDAC
2.11% Unspecified
c
Compounds (1-146)
10 5 0
SSEA4-
CD133-
OCT4-
CD133-/OCT4-/SSEA4-
GSK343
UNC1999
CPI-169
BAY-598
LLY-507
PFI-5
d
Compound screen of hPSCs
Normalised OCT4-/CD133-/SSEA4- cells compared to DMSO treatement
GSK343 (EZH2 inh)
UNC1999 (EZH2 inh)
CPI-169 (EZH2 inh)
PFI-5 (SMYD2 inh)
BAY-598 (SMYD2 inh)
LLY-507 (SMYD2 inh)
Small molecule compounds
e
DMSO
GSK343
UNC1999
PFI-5
LLY-507
Scramble
EZH2 KD1
EZH2 KD2
SMYD2 KD1
SMYD2 KD2
f
Relative expression compared to DMSO treatment
DMSO
GSK343
UNC1999
PFI-5
LLY-507
OCT4
NANOG
SOX2
T
EOMES
SOX17
MESP1
MESP2
SOX1
PAX6
g
Relative expression compared to Scramble treatment
DMSO
EZH2 KD1
EZH2 KD2
h
DMSO
GSK343
UNC1999
EZH2 KD1
EZH2 KD2
NANOG
EOMES
DAPI
i
Relative expression compared to Scramble treatment
Undifferentiated
ENDO + DMSO
ENDO + PFI-5
ENDO + LLY-507
ENDO + SMYD2 KD1
ENDO + SMYD2 KD2
OCT4
NANOG
SOX2
T
EOMES
SOX17
j
DMSO
PFI-5
LLY-507
SMYD2 KD1
SMYD2 KD2
T
EOMES
DAPI
k
Relative expression compared to DMSO treatment
DMSO
GSK343
UNC1999
PFI-5
LLY-507
p14
p15
p16
p18
p21
p27
p57
l
Relative expression compared to Scramble treatment
DMSO
EZH2 KD1
EZH2 KD2
m
Relative expression compared to Scramble treatment
Undifferentiated
ENDO + DMSO
ENDO + PFI-5
ENDO + LLY-507
ENDO + SMYD2 KD1
ENDO + SMYD2 KD2
n
Regulation of hPSC Pluripotency vs Differentiation
SMYD2
EZH2
Differentiation
CDKI induction
Pluripotency
CDKI repression


a
48h
24h
hPSC
Activin
Bmp4
Fgf
PI3K
- Activin
- Bmp4
- Fgf
+ PI3K
p14
p15
p16
p18
p21
p27
p57
Gene Activin Bmp4 Fgf PI3K
p14 = ++ = =
p15 +++ ++ = ++
p16 = - = =
p18 +++ = ++ --
p21 = = – =
p27 ++ = = =
p57 +++ + = +++
Relative expression normalised to undifferentiated cells
UD
All cytokines
No Activin
No Fgf2
No Bmp4
Active PI3K
b
p14/p16 and p15 locus
p18 locus
p19 locus
p21 locus
p27 locus
p57 locus
Oct4 hESC
NANOG hESC
SKIL Leukemia
SMAD2/3 hESC
SMAD2/3 Endoderm
EOMES Endoderm
EZH2 hESC
H3K27me3 hESC
H3K4me3 hESC
c
d
Relative expression normalised to IgG CHIP
p15
p18
p27
p57
1st IgG
1st OCT4
1st NANOG
2nd SMAD2/3
e
p15 locus
p18 locus
p27 locus
p57 locus
Relative enrichment normalised to IgG CHIP
IgG
OCT4
SMAD2/3
SNON
EZH2
Scramble
OCT4 KD
NS
f
SCRAMBLE KD
NANOG KD
OCT4 KD
NANOG
p18
NANOG + p18
DAPI
OCT4
OCT4 + p18
g
Relative expression compared to Scramble cells
Scramble
OCT4 KD
h
NANOG KD
i
Relative binding normalised to IgG CHIP
Scr H3K4me3
OCT4 KD H3K4me3
Scr H3K27me3
OCT4 KD H3K27me3
j
Empty
GFP OE
NANOG OE
SNON OE
OCT4 OE
Relative Luciferase Activity


a
Gene TSS
p15 430 / 3537 / 211 (dns)
p18 -372 / -1621 / 7578
p57 -6995
b
p15 p18 p57
Relative binding normalised to IgG CHIP
1st IgG
1st EOMES
2nd Smad2/3
c
Scramble
EOMES KD
EOMES
p18
EOMES + p18
DAPI
d
Relative expression normalised to undifferentiated cells
UD
Scramble
EOMES KD
p14 p15 p16 p18 p21 p27 p57
e
p57
p15
p18
Empty
GFP OE
EOMES OE
Relative Luciferase Activity
f
EOMES CHIP
SMAD2/3 CHIP
Relative enrichment normalised to IgG CHIP
EOMES KD
p15 p18 p57 p15 p18 p57
g
SMAD2/3 CHIP
EOMES CHIP
Relative enrichment normalised to IgG CHIP
SB
p15 p18 p57 p15 p18 p57
h
p18 locus
SMAD2/3 CHIP
EOMES CHIP
NANOG CHIP
Relative binding normalised to IgG CHIP
UD 12h 24h 36h 48h 60h 72h
i
p15 locus
SMAD2/3 CHIP
EOMES CHIP
NANOG CHIP
Relative binding normalised to IgG CHIP
UD 12h 24h 36h 48h 60h 72h
j
p57 locus
SMAD2/3 CHIP
EOMES CHIP
NANOG CHIP
Relative binding normalised to IgG CHIP
UD 12h 24h 36h 48h 60h 72h
k
p15 p18 p57 p15 p18 p57
H3K4me3
H3K27me3
Relative enrichment normalised to IgG CHIP
EOMES KD
l
p15 p18 p21 p57
IgG
SMYD2
Relative binding normalised to IgG CHIP
ns
SB431542
m
p15 p18 p57
Relative binding normalised to IgG CHIP
1st IgG
1st EOMES
2nd SMYD2
n
SMYD2 CHIP
Relative enrichment normalised to IgG CHIP
EOMES KD
p15 p18 p57
o
Puripotency
OCT4
NANOG
SnoN
EZH2
Smad2
Smad3
RNA pol II
No CDKI expression
Differentiation
EOMES
SMYD2
Smad2
Smad3
RNA pol II
CDKIs: p18, p15, p57
Endoderm TFs


a
H9
Scramble KD
p15 KD
p14/p16 KD
p18 KD
p21 KD
p27 KD
p57 KD
b
Pluripotency
Endoderm
Relative expression normalised to Scramble cells
Oct4
Nanog
Sox2
EOMES
Sox17
Pluripotency
Neuroectoderm
Oct4
Nanog
Sox2
Sox1
Pax6
UD
Scr
p15
p16
p18
p21
p27
p57
c
Endoderm markers
Endoderm differentiation
CDKI Clones
p57.3
p57.2
p57.1
p27.3
p27.2
p27.1
p21.3
p21.2
p21.1
p18.3
p18.2
p18.1
p16.3
p16.2
p16.1
p15.3
p15.2
p15.1
Scr
Neg
EOMES+
Sox17+
CXCR4+
% of positive cells in total population
d
Neuroectoderm markers
Neuroectoderm differentiation
CDKI Clones
Pax6+
Sox1+
% of positive cells in total population
e
Undifferentiated
Endoderm differentiation (low Activin)
Scramble KD
Scramble KD
p15 KD
p18 KD
p57 KD
SOX2
EOMES
SOX2 + EOMES
DAPI
f
Scramble hESCs
p18 KD hESCs
p21 KD hESCs
p57 KD hESCs
Injection of shRNA hESCs
Tumor growth
3 months
Teratoma analysis:
histological examination
or immunohistochemistry
g
shRNA Scramble
shRNA p21
shRNA p18
shRNA p57


a
Overexpression of CDKIs
Number of colonies
400
300
200
100
0
GFP p14 p15 p16 p18 p21 p27 p57
b
GFP
OE p14
OE p15
OE p16
OE p18
OE p21
OE p27
OE p57
c
Germ layer differentiation
CDKI OE
Cyclin D Double KD
EdU incorporation
Cell cycle analysis
d
Percentage of total population
100
50
0
G1%
S%
G2%
Clone 1
Clone 2
Clone 3
pTP6
p14
p15
p16
p18
p21
p27
p57
e
OE Control
OE p14
OE p15
OE p16
OE p18
OE p21
OE p27
OE p57
Tra-1-60
Sox17
Tra-1-60 + Sox17
DAPI
f
Pluripotent
Endoderm
Mesoderm
Neuroderm
GFP
p14
p15
p16
p18
p21
p27
p57
Relative expression normalised to GFP OE cells
30
20
10
0
Pou5f1
Nanog
Sox2
EOMES
Sox17
GSC
T
Mesp1
Sox1
Pax6
g
Endoderm differentiation
OE CDKI Clones
p57.3
p57.2
p57.1
p27.3
p27.2
p27.1
p21.3
p21.2
p21.1
p18.3
p18.2
p18.1
p16.3
p16.2
p16.1
p15.3
p15.2
p15.1
Vector
Neg
EOMES+
Sox17+
CXCR4+
0
20
40
60
80
100
% of positive cells in total population
h
Neuroectoderm differentiation
OE CDKI Clones
Sox1+
Pax6+
0
20
40
60
80
100
% of positive cells in total population


a
Cdt1-mKO2
Geminin-mAG
DAPI
Overlay
FACS sorting
Tra-1-60+
Gating
Late G1 phase
Scramble
OE p15
OE p18
Analysis of early markers at 12h
b
Relative expression normalised to undifferentiated cells
Mixl1
GSC
EOMES
T
p15 OE
p18 OE
c
% Sox17+ cells
p57
p15
p18
pTP6
Day 0
Day 1
Day 2
Days of differentiation
d
shRNA
OE
Scr
p15
p18
p57
Smad2/3
Histone H3
Chromatin
Smad2/3
Actin
Cytoplasm
e
Smad3
EPSM
SB
CycD1
CDKi
p15
p18
NS
Empty
Relative Luciferase Activity
f
% of total population
G1%
S%
G2%
Clone 1
Clone 2
Clone 3
Scramble
p18
pTP6
shRNA
OE
g
Pancreatic Organoids from hiPSCs
CRISPRi: p18
DOX or PD0332991
Mesoderm
MESP1, MESP2
1
Definitive endoderm
SOX17, FOXA2
Ventral foregut
SOX17, CER
2
Dorsal foregut
HNF4A, HLXB9
Hepatic endoderm
AFP, HEX
3
Pancreatic endoderm
PDX1, SOX9
Fetal hepatocytes
A1AT, ALB
4
Pancreatic islet cells
NGN3, INSULIN, SST, GCG
Delta cell (SST)
F cell (PP)
Alpha cell (GCG)
Beta cell (Insulin)
h
1
Definitive endoderm
Relative change compared to Undifferentiated hiPSCs
SOX17
FOXA2
MESP1
MESP2
3
Pancreatic progenitor
PDX1
SOX9
AFP
HEX
UD
Control
p18 iKD
PD0332991
2
Dorsal foregut
HNF4A
HLXB9
SOX17
CER
4
Pancreatic islet cells
NGN3
INSULIN
SST
GCG
A1AT
ALB
i
3
Pancreatic endoderm / pancreatic progenitors
DMSO
PD0332991
Constant PD0332991
PDX1
DAPI
j
3
Pancreatic endoderm / pancreatic progenitors
DMSO
PD0332991
Constant PD0332991
NGN3
DAPI
k
4
Pancreatic islet cells
DMSO
PD0332991
Constant PD0332991
INSULIN
DAPI

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.pdf