

Integrated loss- and gain-of-function screens define a core network governing human embryonic stem cell behavior

Kamila Naxerova,^{1,2,3} Bruno Di Stefano,^{4,5} Jessica L. Makofske,¹ Emma V. Watson,^{1,2} Marit A. de Kort,⁴ Timothy D. Martin,^{1,2} Mohammed Dezfulian,^{1,2} Dominik Ricken,³ Eric C. Wooten,^{1,2} Mitzi I. Kuroda,¹ Konrad Hochedlinger,⁴ and Stephen J. Elledge^{1,2}

¹Division of Genetics, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA; ²Howard Hughes Medical Institute, Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA; ³Center for Systems Biology, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA; ⁴Department of Molecular Biology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA

Understanding the genetic control of human embryonic stem cell function is foundational for developmental biology and regenerative medicine. Here we describe an integrated genome-scale loss- and gain-of-function screening approach to identify genetic networks governing embryonic stem cell proliferation and differentiation into the three germ layers. We identified a deep link between pluripotency maintenance and survival by showing that genetic alterations that cause pluripotency dissolution simultaneously increase apoptosis resistance. We discovered that the chromatin-modifying complex SAGA and in particular its subunit TADA2B are central regulators of pluripotency, survival, growth, and lineage specification. Joint analysis of all screens revealed that genetic alterations that broadly inhibit differentiation across multiple germ layers drive proliferation and survival under pluripotency-maintaining conditions and coincide with known cancer drivers. Our results show the power of integrated multilayer genetic screening for the robust mapping of complex genetic networks.

[*Keywords:* human embryonic stem cells; germ layer formation; genetic screening]

Supplemental material is available for this article.

Received September 21, 2021; revised version accepted September 22, 2021.

Human embryonic stem cells (hESCs) are of central interest to developmental biology and regenerative medicine. Their unique biology—capacity for unlimited proliferation coupled to an unusual cell cycle configuration (Liu et al. 2019), ability to differentiate into the three germ layers and a wide array of derivative cell types, and distinctive epigenetic (Gaspar-Maia et al. 2011) and transcriptional (Efroni et al. 2008) characteristics—makes them a challenging and fascinating topic of study. Although great strides have been made in charting the molecular traits of hESCs (Young 2011), much remains to be learned about the genetic networks that control their most fundamental

behaviors: proliferation, survival, pluripotency, and differentiation. Powerful new methods for genome-scale loss-of-function (Shalem et al. 2014) and gain-of-function (Sack et al. 2018) genetic screening have opened up new opportunities for high-throughput identification of genes that control these processes. In murine embryonic stem cells (ESCs), screens have identified central regulators of pluripotency and differentiation (Zambrowicz et al. 1998; Aubert et al. 2002; Chambers et al. 2003; Pritsker et al. 2006; Hu et al. 2009). Genome-wide screens in hESCs, which differ from mouse ESCs in several important respects (Yu and Thomson 2008), have been conducted less

⁵Present address: Stem Cells and Regenerative Medicine Center, Baylor College of Medicine, Houston, TX 77030, USA.

Corresponding author: selledge@genetics.med.harvard.edu, naxerova.kamila@mgh.harvard.edu

Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.349048.121>.

© 2021 Naxerova et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

frequently, but have yielded many equally important insights (Chia et al. 2010; Tajonar et al. 2013; Zhang et al. 2013; Shalem et al. 2014; Gonzales et al. 2015; Yilmaz et al. 2018, 2020; Ihry et al. 2019; Li et al. 2019).

Genetic screens represent high-throughput versions of the classical “one gene—one phenotype” approach of investigating the effects of individual gene alterations on a process of interest. However, in most screens, the genetic underpinnings of only one predefined phenotype (e.g. increased fitness) are interrogated at a time through either loss- or gain-of-function alterations. This “many genes—one phenotype” design can be limiting for the robust identification of complex genetic networks that would best be probed from many angles, across different conditions, and with different kinds of perturbations. Such “many genes—many phenotypes” measurements, ideally integrated with additional genomic data, would provide a more precise and comprehensive view of the networks that control interconnected cellular behaviors such as proliferation and differentiation.

Here we describe a systematically designed collection of genome-scale screens defining the effects of gene inactivation (via CRISPR/Cas9) and overexpression (via open reading frames [ORFs]) on hESC proliferation, survival and the formation of the three germ layers: endoderm, mesoderm, and ectoderm. Integrating screen data with orthogonal genomic, epigenomic and transcriptomic information, we paint a comprehensive picture of the principal genetic networks governing hESC behavior. We identified an inverse relationship between pluripotency and survival by showing that apoptosis resistance automatically increases when the pluripotency maintenance network is perturbed. We discovered that the chromatin-modifying complex SAGA and in particular its subunit TADA2B regulate all fundamental hESC functions, including pluripotency, survival, growth and differentiation ability. We identify genes that act as universal differentiation inhibitors or universal differentiation facilitators and show that these genes coincide with known cancer drivers. Finally, we show that genetic alterations that interfere with correct germ layer formation drive proliferation under pluripotency-maintaining conditions. Thus, using a “many genes—many phenotypes” systems genetics approach, we provide a deep and detailed view of the core networks that govern hESC behavior.

Results

Genome-wide gain- and loss-of-function screens to identify regulators of hESC proliferation and germ layer formation

To identify genes that control hESC proliferation and germ layer formation, we performed parallel genome-wide loss- and gain-of-function screens in the male hESC line HUES64 using a CRISPR/Cas9 library (Wang et al. 2015; Martin et al. 2017) and a recently developed inducible barcoded ORF library (Fig. 1A; Sack et al. 2018). We established a reverse tetracycline transactivator (rtTA)-expressing nonclonal HUES64 subline for the latter

screen (Materials and Methods). We chose HUES64 because it was extensively investigated as a model system for epigenetic changes during germ layer formation in previous studies, providing us with orthogonal reference data (Bock et al. 2011; Gifford et al. 2013). Successfully transduced cells passed through four distinct screen arms per library: a “proliferation and survival” screen, and three germ layer formation screens. During proliferation screens, cells were grown in pluripotency-maintaining conditions for 13–16 population doublings (PD) while undergoing regular single-cell dissociation and replating in the presence of a ROCK inhibitor (Watanabe et al. 2007). Cell morphology remained consistent throughout this time period (Fig. 1B, top two panels), and we observed no population-level changes in pluripotency marker TRA-1-60 expression (Fig. 1B, bottom panel), indicating that the majority of cells continued to proliferate as pluripotent stem cells. We confirmed that no chromosomal abnormalities emerged during the screens by low-coverage whole-genome sequencing of all replicates at the first and last collection time points (Supplemental Fig. S1). We quantified the effect of a gene’s inactivation or overexpression on cellular fitness under these conditions by comparing the abundance of single-guide RNA (sgRNA) groups targeting the same gene, or ORF barcodes mapping to the same gene, at the beginning and end of the screens. Screens were analyzed with two different methodologies, edgeR/Camera (Robinson et al. 2010; Ritchie et al. 2015) and MAGeCK (Li et al. 2014), which yielded similar results (see Supplemental Table S1 for full results and false discovery rates). While screening in hESCs is technically challenging due to the cells’ delicacy and their sensitivity to Cas9-induced double-strand breaks (Ihry et al. 2018), we conducted rigorous quality controls and benchmarked performance in comparison with previously published screens. These results are presented in the Supplemental Note on screen performance (Supplemental Material).

For the germ layer formation screens, cells were placed into media that induced differentiation into SOX1/OTX2-positive ectoderm, Brachyury/HAND1-positive mesoderm, or GATA4/SOX17-positive definitive endoderm (Materials and Methods; Supplemental Fig. S2). After 5 d of differentiation, correctly differentiated cells were separated from undifferentiated cells via fluorescence-activated cell sorting (FACS) (Fig. 1A). We used well-established cell surface marker combinations that were previously used in a landmark study on the transcriptional and epigenetic dynamics of germ layer formation (Gifford et al. 2013). These included up-regulation of CXCR4 for endoderm (D’Amour et al. 2005), and loss of EpCAM and up-regulation of NCAM for mesoderm and ectoderm (Sundberg et al. 2009; Evseenko et al. 2010). Please see the Supplemental Note for more details on FACS strategy and experimental design. This approach enabled us to integrate our screen results with ChIP-seq and RNA-seq data from this previous work. Genes that affect differentiation efficiency were identified by comparing the abundance of sgRNAs and ORF barcodes between differentiated and undifferentiation populations with edgeR/Camera and MAGeCK (Supplemental Table S1).

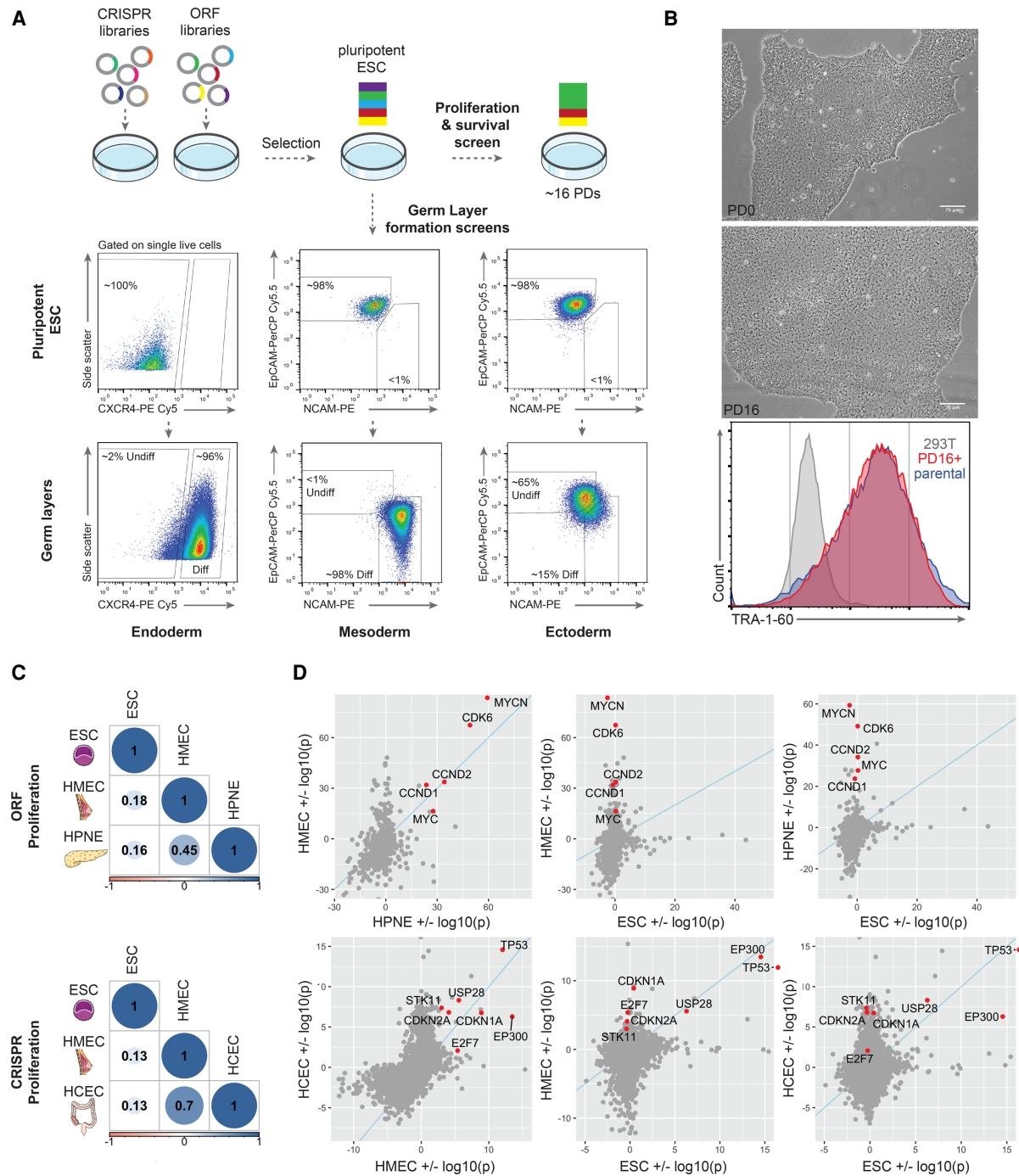


Figure 1. CRISPR/Cas9 and ORF expression screens in proliferating and differentiating hESCs. (A) Experimental design schematic. Transduced cells proliferate for 13–16 populations doublings (PD) in pluripotency-maintaining conditions (mTeSR1 media) or enter 5-d differentiation protocols to form CXCR4-positive endoderm, EpCAM-negative/NCAM-positive mesoderm, and EpCAM-negative/NCAM-positive ectoderm. (B) Morphology of transduced cells after selection with puromycin at the beginning (PD0) and end (PD16) of the proliferation screen. Scale bar, 70 μ m. The *bottom* panel shows TRA-1-60 expression in 293T cells as a negative control, parental uninfected HUES64, and transduced cells that have completed the proliferation screen (>16 PDs). All panels show cells transduced with the P1/P3 CRISPR sublibrary. (C) Pairwise correlation coefficients of \log_2 fold changes for all genes between beginning and end time points in the CRISPR and ORF proliferation screens, across all interrogated cell lines. (D) Log-log P -value plots showing the performance of genes in the proliferation screens across all cell lines. If a gene has a negative fold change (i.e., is depleted during the screen), it receives a value of $\log_{10}(P\text{-value})$, while genes with positive fold changes are plotted as $-\log_{10}(P\text{-value})$. Therefore, genes in the *top right* quadrant enrich in both compared cell lines, genes in the *bottom left* quadrant drop out in both, and genes in the *top left* and *bottom right* quadrants behave in opposite ways. Please note that axes are adjusted to allow optimal viewing of the majority of genes; some outliers with extremely low P -values (e.g. PMAIP1 in ESCs) are not plotted on this scale.

hESCs and somatic cells have distinct cell cycle regulation

We began by comparing hESC proliferation screen results with equivalent screens that we conducted with the same libraries in normal human somatic cell lines (Sack et al. 2018). For both CRISPR and ORF libraries, screen-wide correlation coefficients between hESCs and human mammary epithelial cells (HMECs), human pancreatic nestin-expressing cells (HPNEs), or human colonic epithelial cells (HCECs) were substantially lower than correlations between somatic cell lines, indicating globally distinct behavior of hESCs (Fig. 1C). We further noticed that across all of our CRISPR and ORF screens, as well as previously published data, genetic alterations generally resulted in less pronounced proliferation phenotypes in hESCs than in somatic cells (explored in detail in the [Supplemental Note on screen performance](#)), indicating that hESCs' uncommitted chromatin (Gaspar-Maia et al. 2011) and transcriptional (Efroni et al. 2008) state may make them less sensitive to perturbation of single pro- or antiproliferative pathways. Examining the genes that affected proliferation in somatic cells, but not in hESCs, we noted large discrepancies in the performance of classical cell cycle regulators. While overproduction of MYC and components of the CDK4/6–Cyclin D complex strongly drove proliferation in both HMECs and HPNEs, overexpression of these ORFs had no effect in hESCs (Fig. 1D, top panel). Conversely, CDKN1B scored as the number one growth-inhibiting ORF in both HMECs and HPNEs, but showed attenuated effects in hESCs, scoring only as the 176th most depleted gene ([Supplemental Table S1](#)). Similarly, CRISPR-mediated knockout of CDK inhibitors and E2F7, a repressor of G1/S transition genes (Westendorp et al. 2012), enhanced proliferation in HMECs and HCECs but not hESCs (Fig. 1D, bottom panel). These results are consistent with documented differences in cell cycle regulation between pluripotent and differentiated cells. ESCs go through an accelerated G1 phase and exhibit constitutively high CDK activity and RB phosphorylation, as well as constitutively high expression of E2F target genes (White and Dalton 2005). Thus, our screening results effectively capture the unique cell cycle wiring and diminished role of the restriction point in hESCs.

Heightened apoptosis readiness characterizes pluripotent cells in vitro and in the human embryo

Next, we examined genes that enriched during hESC proliferation and passaging when mutated or overexpressed. The majority of these genes had no effect on the growth of somatic cell lines (Fig. 2A,C; [Supplemental Fig. S3](#)). Genes that scored in the CRISPR screen (i.e., functioned as hESC-specific growth and survival suppressors) predominantly fell into a small number of functional categories (Fig. 2A). Multiple components of the histone-modifying Spt–Ada–Gcn5 acetyltransferase (SAGA) complex were among the highest-ranking genes, along with members of the Polycomb repressive complex 1 (PRC1) and a large number of proapoptotic genes. High-scoring

genes further included core constituents of the pluripotency network (POU5F1/OCT4, SOX2, and LIN28A) and members of differentiation pathways such as retinoic acid signaling (RARA and RXRA) and WNT signaling (TCF7L2, CTNNBIP1, and MED13). Gene ontology (GO) enrichment analysis further revealed that genes involved in the regulation of interferon- γ production had significant phenotypes (Fig. 2B). Remarkably, while different genes enriched in the ORF proliferation screen, they fell into very similar functional categories (Fig. 2C), including apoptosis-related genes, PRC1 components, and a subset of pluripotency- and differentiation-associated genes. In contrast, CRISPR and ORF screens diverged significantly with respect to depleted genes. As expected, dropouts in the CRISPR screen were universally essential genes required for fundamental aspects of cell function such as RNA processing and mitochondrial function (Fig. 2B). They furthermore overlapped highly significantly with essential genes identified in two other hESC lines ([Supplemental Note on screen performance](#)). In contrast, genes that inhibited hESC growth when overexpressed were primarily developmental transcription factors that play roles in later stages of embryogenesis and tissue-specific differentiation. These were represented by GO categories such as developmental process, cell differentiation, and anatomical structure development (Fig. 2D), potentially indicating that forced differentiation due to ectopic overexpression of developmental transcription factors interferes with hESC proliferation.

Practically the entire p53-mediated apoptosis pathway was present among the top enriched genes in both hESC screens, with proapoptotic signaling proteins and effectors scoring in the CRISPR arm, and antiapoptotic members of the BCL2 family scoring in the ORF arm (Fig. 2E; Youle and Strasser 2008). With the exception of TP53 and USP28, these genes had no effect on the growth of somatic cell lines (Fig. 2A,B; [Supplemental Fig. S3A,B](#)). hESCs have a high propensity for undergoing apoptosis in response to colony dissociation (Watanabe et al. 2007) and DNA damage (Wilson et al. 2010). This propensity is caused by high mitochondrial priming; i.e., a high baseline ratio between proapoptotic and antiapoptotic proteins (Liu et al. 2013). Our results show that lowering this ratio by overexpressing or inactivating apoptotic effectors can substantially raise hESCs' survival likelihood, as has been seen for BCL2 overexpression (Ardehali et al. 2011). Notably, the scoring of antiapoptotic genes in the ORF screen suggested that Cas9-mediated toxicity is not the reason for the enrichment of proapoptotic gene knockouts in the CRISPR screen (Ihry et al. 2018).

A key question is whether the low apoptotic threshold of hESCs is an artifact of common culture conditions, or whether it is hardwired into the pluripotent state, perhaps in order to ensure efficient elimination of compromised cells in the epiblast. To address this issue, we first analyzed RNA-seq data of HUES64 hESCs and their differentiated germ layer derivatives (Gifford et al. 2013). Genes involved in p53-mediated apoptosis were significantly up-regulated in undifferentiated hESCs (Fig. 2F, left panels), consistent with the idea that high mitochondrial

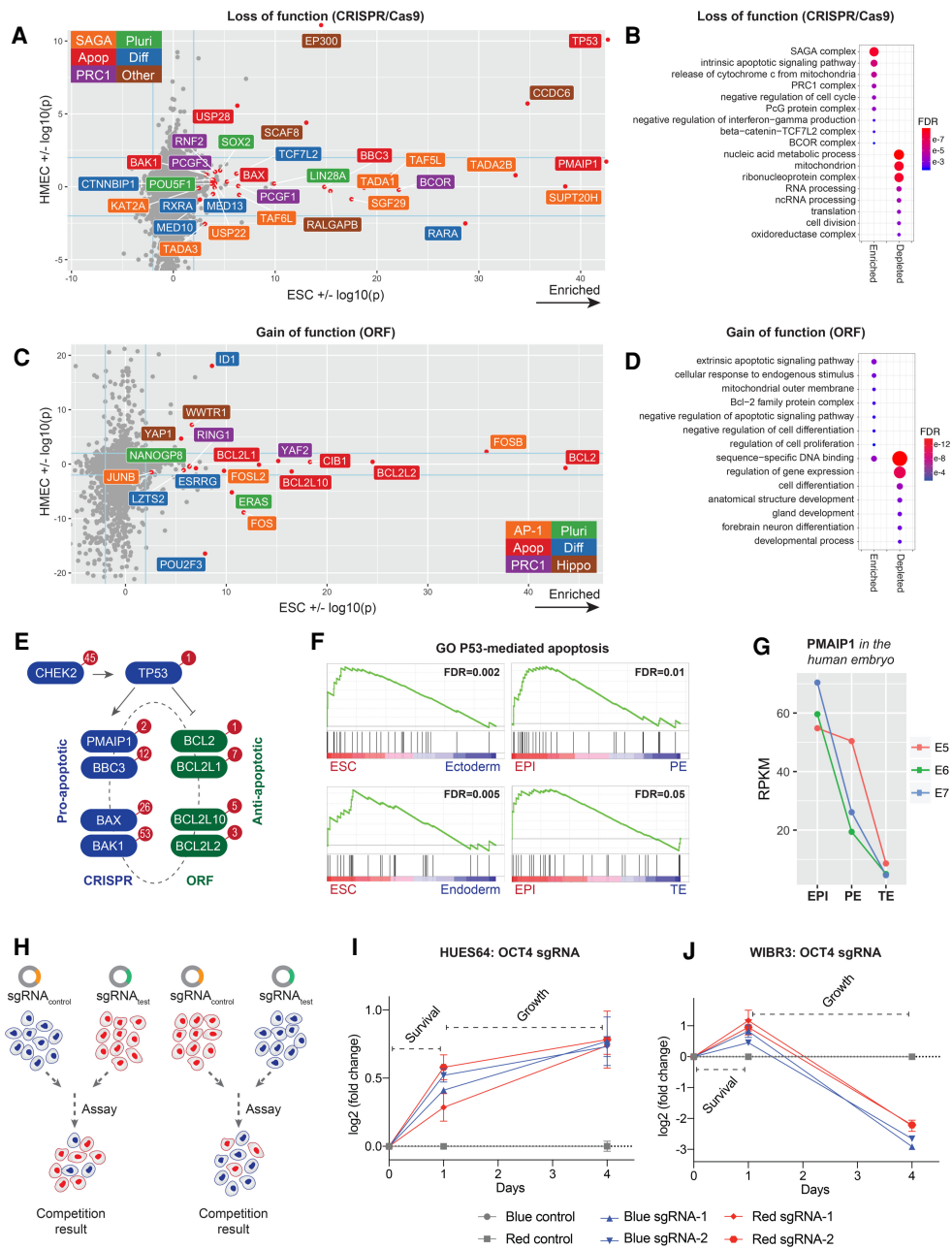


Figure 2. hESC-specific drivers and inhibitors of proliferation and survival. (A) CRISPR proliferation screen log-log P -value plot comparing the performance of genes between hESCs and HMECs. Axes are as in Figure 1D, adjusted for optimal viewing of most genes, with outliers indicated by a red point outside the coordinate system. Blue lines correspond to $P = 0.01$. (B) GO categories enriched in CRISPR proliferation screen hits, calculated by GOrilla. (C) ORF proliferation screen log-log P -value plot comparing hESCs and HMECs. Axes are as in A. (D) GO categories enriched in ORF proliferation screen hits, calculated by GOrilla. (E) Schematic representation of the mitochondrial (intrinsic) apoptosis pathway. The enrichment ranks of genes in the CRISPR (blue) and ORF (green) proliferation screens are plotted within the red circles. (F) Expression of apoptosis-related genes in hESCs. Enrichment analysis of gene set “GO intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator” in ESCs (HUES64) versus HUES64-derived ectoderm and HUES64-derived endoderm, in human epiblast (EPI) versus primitive endoderm (PE), and in EPI versus trophectoderm (TE). Genes are ordered by their \log_2 fold changes; e.g., $\log_2(\text{ESC}/\text{ectoderm})$. FDRs were calculated via GSEA gene set permutation. (G) Expression of PMAIP1 in human EPI, PE, and TE cells at embryonic days (E) 5, 6, and 7. Each data point represents the mean of dozens to hundreds of single-cell RNA-seq profiles. (H) Schematic illustrating the multicolor competition assay. (I) Survival and growth of HUES64 hESCs transduced with sgRNAs against OCT4 (blue/red) or a set of control genes (gray). The first measurement time point (day 1) is a survival assay in which single cells are plated in the presence of ROCK inhibitor and counted 24 h later. The second measurement (day 4) quantifies how much cells have proliferated since readhering after passaging. Fold changes were normalized to the control. Error bars show standard deviation of three replicates. (J) Survival and growth of WIBR3 OCT4-GFP hESCs transduced with sgRNAs against OCT4 (blue/red) or a set of control genes (gray). Error bars show standard deviation of one to four replicates.

priming is part of the pluripotent phenotype. More importantly, we found that the same effect can be observed in human embryos. We analyzed published single-cell RNA-seq data derived from human preimplantation embryos (Petropoulos et al. 2016) and found that apoptosis-related genes were highly expressed in the human epiblast (EPI) and swiftly down-regulated upon differentiation to primitive endoderm (PE) or trophectoderm (TE) (Fig. 2F, right panels). PMAIP1, the second most enriched gene in our CRISPR screen, was among the top 25 up-regulated genes in a comparison of epiblast versus trophectoderm, with a precipitous drop of expression upon differentiation (Fig. 2G). Collectively, these results show that high apoptosis tendency is an integral component of the pluripotent state *in vitro* and most likely *in vivo*, as cultured embryos typically give rise to viable offspring.

We reasoned that a hardwired connection between pluripotency and apoptosis readiness could potentially explain the counterintuitive observation that loss of core pluripotency regulators SOX2, OCT4 (POU5F1), and LIN28A resulted in a competitive advantage during the CRISPR proliferation screen (Fig. 2A). In an elegant study of hESC mitochondrial priming, Liu et al. (2013) showed that depletion of OCT4 significantly reduces hESC apoptosis in response to DNA damage. Since the ability to evade apoptosis evidently is one of the dominant selective pressures in the proliferation and survival screens, we hypothesized that increased survival of OCT4 or SOX2 mutant cells during colony dissociation and passaging is responsible for the observed enrichment. To test this hypothesis directly, we conducted validation experiments with an internally controlled multicolor competition assay (MCA). We generated HUES64 sublines that constitutively expressed either blue fluorescent protein (BFP) or the far-red fluorescent protein E2-Crimson (E2C). Blue cells were transduced with a small pool of control sgRNAs that had no phenotype in any of our CRISPR screens (Materials and Methods), while red cells were transduced with one of two sgRNAs targeting OCT4. Red and blue cells were then mixed and allowed to compete against each other in a survival and proliferation assay. To account for possible effects of BFP or E2C expression on the cells' competitive advantage, we also conducted the experiment with reversed colors (blue OCT4 mutants with red controls) (Fig. 2H). We observed a large survival advantage for OCT4 mutants 24 h after single-cell dissociation and plating in the presence of ROCK inhibitor (analogous to the proliferation screen setting) (Fig. 2I). The mutants' subsequent proliferation rate was also slightly elevated over control. To determine whether increased survival of OCT4 mutants was a general phenomenon, we repeated the MCA in WIBR3 OCT4-GFP hESCs (Hockemeyer et al. 2011). In this cell line, enhanced green fluorescent protein (eGFP) is integrated into the endogenous OCT4 locus (following the last OCT4 codon and preceded by a 2A sequence for separate translation), allowing for FACS-based evaluation of OCT4 expression and pluripotency. Expression of sgRNAs against OCT4 resulted in a loss of GFP expression in up to 90% of cells (Supplemental Fig. S3C), confirming

high Cas9 efficiency. As in HUES64, we observed a strong enrichment of OCT4 mutants during passaging (Fig. 2J), confirming our hypothesis that pluripotency loss confers a survival advantage. Effect sizes were similar in HUES64 (mean 1.36-fold enrichment in 24 h) and WIBR3 (1.68-fold). In contrast to HUES64, the postpassaging proliferation rate of WIBR3 OCT4 mutants was substantially reduced, indicating that this cell line differentiates into a more slowly proliferating cell type after pluripotency loss (Fig. 2J). Together, these results suggested that like OCT4, other CRISPR screen enrichment hits might also represent genes that are required for pluripotency maintenance.

The SAGA complex regulates survival and proliferation in hESCs

A notable result from the CRISPR proliferation screen was the strong enrichment of chromatin-modifying SAGA (or STAGA) complex components. SAGA is a large protein complex that contains a histone acetyltransferase (HAT) module, a deubiquitination (DUB) module, and two modules that interact with the transcriptional machinery (SPT and TAF) (Wang and Dent 2014). Of the 18 proteins that belong to these modules, nine (50%) scored among the top 100 enriched genes in the CRISPR screen ($P = 1.5 \times 10^{-16}$). Hits were distributed across all modules (Fig. 3A). The most striking phenotype was exhibited by TADA2B, a component of the HAT module. We validated the effect of TADA2B loss in HUES64 using the MCA. TADA2B mutants showed enhanced survival during passaging, and also proliferated significantly faster than wild-type hESCs (Fig. 3B). Increased survival and proliferation were also observed in WIBR3 TADA2B mutants (Fig. 3C). We furthermore performed a standard growth assay in a third hESC cell line, UCLA9, and again observed a large growth advantage in mutants (Fig. 3D), suggesting that, unlike OCT4, TADA2B loss uniformly enhances survival and proliferation of hESCs. To determine whether the strong proproliferative effects of TADA2B loss were specific to hESCs, as our comparison with equivalent screens in two somatic cell lines suggested (Fig. 2A; Supplemental Fig. S3A,B), we analyzed fold changes after TADA2B loss across CRISPR screens in 626 human cancer cell lines from the Cancer Dependency Map (DepMap). Log₂ fold changes and dependency scores (Meyers et al. 2017) for TADA2B were negative in a majority of DepMap lines, indicating that TADA2B loss generally does not provide a growth advantage to somatic cells. The greatest positive log₂ fold change for TADA2B observed across the DepMap was 3.12, corresponding to an approximately ninefold enrichment during the screen. In hESCs, TADA2B mutants enriched ~600-fold (log₂ fold change 9.24), highlighting the unique behavior of TADA2B in this cell type (Fig. 3E). Other common tumor suppressors like TP53 showed comparable performance between the DepMap (which contains many cell lines that are already TP53 mutant and thus insensitive to TP53 targeting sgRNAs) and our screens (Fig. 3E).

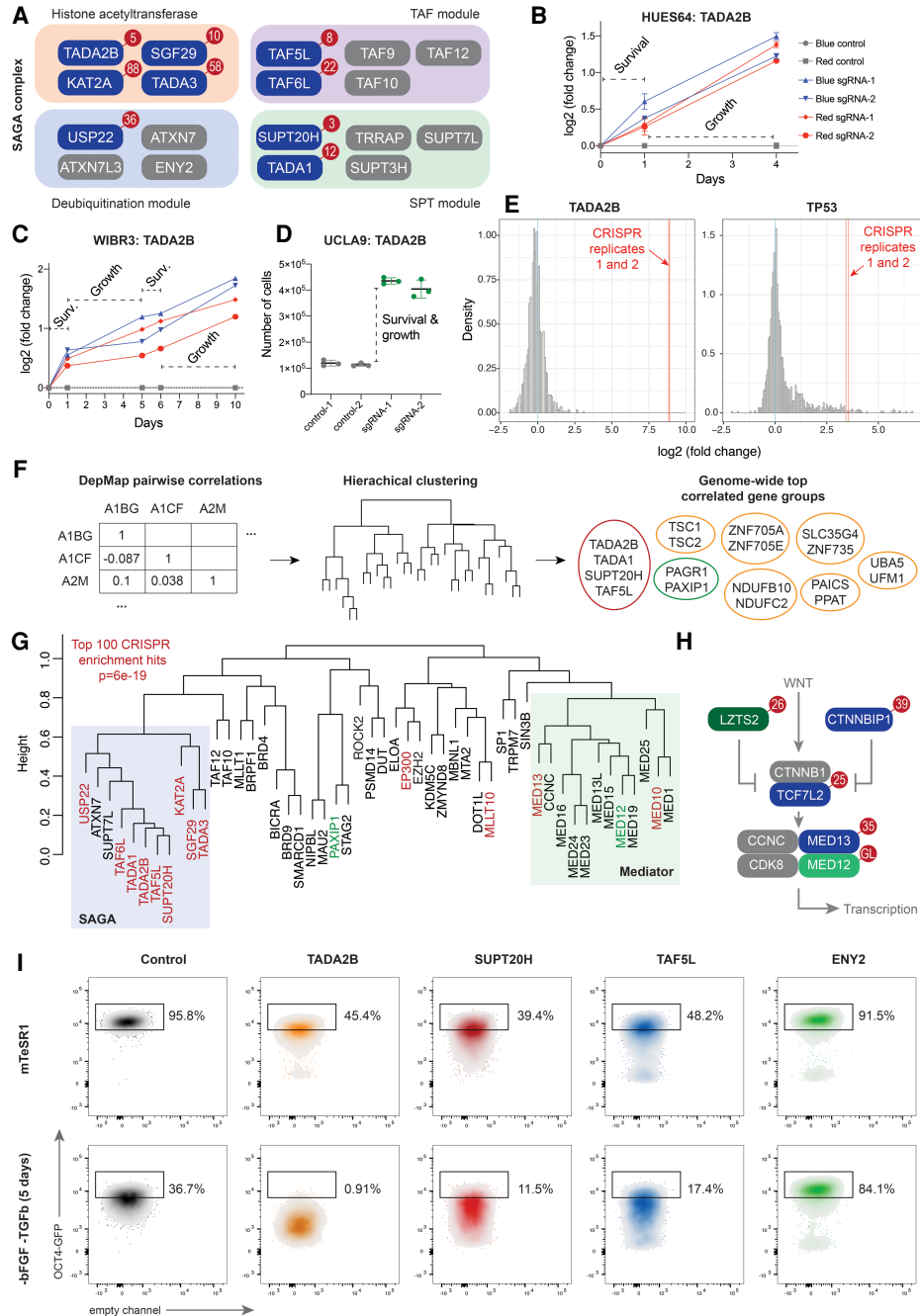


Figure 3. Role of the SAGA complex in hESC survival, proliferation, and pluripotency maintenance. (A) Schematic of the SAGA complex. Components that scored among the top 100 enriched genes in the CRISPR proliferation screen are colored blue and carry their enrichment rank within the red circle. (B) Survival and proliferation of HUES64 hESCs transduced with sgRNAs against TADA2B (blue/red) or a set of control genes (gray). (C) Survival and proliferation of WIBR3 OCT4-GFP hESCs transduced with sgRNAs against TADA2B (blue/red) or a set of control genes (gray). (D) Combined survival and proliferation assay (4 d) in UCLA9 hESCs transduced with sgRNAs against TADA2B or control sgRNAs. (E) Log₂ fold change distributions for TADA2B and TP53 across 625 cancer cell lines in the DepMap. Log₂ fold changes for TADA2B and TP53 from the hESC CRISPR screens (replicates 1 and 2) are indicated as red lines. (F) Identification of highly correlated gene groups in the DepMap via calculation of pairwise correlations and hierarchical clustering. The top eight correlated gene groups in the human genome are shown. (G) Correlation-based hierarchical clustering of the top 50 TADA2B-correlated genes in the DepMap. Top hits from the hESC CRISPR proliferation screen (red) are highly significantly enriched among this set of genes. Genes shown in light-green font scored in the germ layer formation screens (see below). (H) WNT signaling pathway schematic. The enrichment ranks of genes in the CRISPR (blue) or ORF (green) proliferation screens are plotted within the red circles. Light green for MED12 indicates that it scored as a differentiation facilitator in germ layer formation screens (see below). (I) OCT4-GFP expression in WIBR3 hESCs transduced with sgRNAs against SAGA complex components TADA2B, SUPT20H, TAF5L, and ENY2. The top row shows cells cultured in mTeSR1 media, and the bottom row shows cells after 5 d in mTeSR1 media lacking bFGF and TGF- β .

The large number of SAGA components among the top scoring hits in the proliferation screen indicated that their function was nonredundant and unusually tightly correlated. To determine whether high correlations between SAGA subunits could also be observed in non-hESC cell lines, we queried the DepMap for highly correlated gene groups. We calculated pairwise correlations between the dependency scores for all genes across all cell lines and used this correlation distance matrix for hierarchical clustering. Based on the resulting dendrogram, we identified gene groups with the strongest genome-wide associations. Only one among the top 10 correlated gene groups contained more than two genes. This was a module containing the SAGA components TADA2B, TADA1, SUPT20H, and TAF5L (Fig. 3F). The average correlation coefficient within this module was $r=0.87$, a remarkably high value that surpassed even the correlation observed among members of protein complexes as fundamental as the mitochondrial membrane respiratory chain NADH dehydrogenase/complex I (NDUFB10 and NDUFC2, $r=0.82$). To confirm beyond our screen data that other gene members of this core module behave similarly to TADA2B in hESCs, we validated the effects of SUPT20H loss with the MCA and observed enhanced survival in HUES64 and UCLA9, and also increased proliferation in HUES64 (Supplemental Fig. S3D). Hence, our screens reveal SAGA, a protein complex whose members display some of the strongest genetic relationships in the human genome, as a top regulator of survival and proliferation in hESCs.

To explore the genetic neighborhood of TADA2B in more detail and identify potential SAGA interaction partners, we extracted the top 50 TADA2B-correlated genes from the DepMap and clustered them by their pair-wise dependency score correlations (Fig. 3G). We noted that the overlap between the TADA2B-correlated genes and our top 100 hESC CRISPR screen hits was highly significant ($P=6 \times 10^{-19}$). The analysis revealed tight genetic associations between SAGA and the Mediator complex (Fig. 3G). Mediator is a general transcription coactivator that is recruited to target genes by the activation domains of transcription factors (Allen and Taatjes 2015). Since specific Mediator subunits (of which there are 26) interact with different transcription factors, its components are thought to represent the “end point of cell signaling pathways” (Jiang et al. 1998; Allen and Taatjes 2015), enabling the translation of distinct intracellular and extracellular signals into specific transcriptional output. SAGA participates in this process by modifying and opening chromatin around transcription start sites and facilitating transcript elongation (Koutelou et al. 2010). MED13, a component of the CDK8 module of Mediator that is required for relaying WNT signals (Allen and Taatjes 2015), was among the top CRISPR screen hits, in addition to several other members of the WNT signaling pathway (Fig. 3H). These orthogonal data from the DepMap confirm that our hESC CRISPR screen identified an interconnected gene network (consisting of SAGA, several Mediator subunits, and WNT pathway members) that plays a fundamental role in regulating hESC survival and/or proliferation.

The SAGA complex is involved in pluripotency maintenance

While SAGA complex function in hESCs could be pleiotropic and include direct acetylation and activation of p53 (Barlev et al. 2001), the conspicuous position of its subunits at the top of the CRISPR screen list along with OCT4 and SOX2 led us to speculate that SAGA may have a role in pluripotency maintenance. To test the effects of SAGA loss on pluripotency, we knocked out one component of each of the four SAGA modules in WIBR3 OCT4-GFP cells. We chose TADA2B from the HAT module, TAF5L from the TAF module, SUPT20H from the SPT module, and ENY2 from the DUB module. All genes but ENY2 were among the top 10 hits in our CRISPR screen, suggesting that ENY2's function may diverge from the others. Loss of TADA2B, TAF5L, and SUPT20H caused pronounced down-regulation of OCT4-GFP in WIBR3 cells in pluripotency-maintaining conditions, while ENY2 mutants largely remained OCT4-GFP-positive (Fig. 3I). TADA2B, TAF5L, and SUPT20H mutants also lost OCT4-GFP expression more rapidly than cells transduced with a control sgRNA when cultured in media that lacked bFGF and TGF- β , signals that are required for pluripotency maintenance (Fig. 3I). In contrast, ENY2 mutants were unable to down-regulate OCT4-GFP expression, confirming the divergent function of ENY2 suggested by the hESC CRISPR screen. We further validated these findings by measuring RNA levels of pluripotency markers OCT4, NANOG, and LEFTY1 in SAGA mutants and found that the results were consistent with accelerated loss of pluripotency in TADA2B, TAF5L, and SUPT20H mutants, and abnormally stable pluripotency in ENY2 mutants (Supplemental Fig. S4). The same results were obtained with shRNAs against TADA2B and ENY2 (Supplemental Fig. S5). We conclude that the SAGA complex has a hitherto unappreciated but central function regulating major aspects of hESC behavior: survival, proliferation, and pluripotency maintenance. Furthermore, SAGA has previously been implicated in the activation of some lineage-specific genes (Chen and Dent 2014), raising the interesting possibility that it also governs differentiation, a function that is discussed in more detail below.

Loss of the BCOR-PRC1.1 complex enhances hESC survival

Polycomb repressive complexes PRC1 and PRC2 co-occupy the promoters of developmental genes in ESCs and ensure the repression of lineage-specific transcriptional programs (Boyer et al. 2006; Lee et al. 2006; Morey et al. 2012). Loss of EZH2, the catalytic subunit of PRC2, leads to self-renewal and proliferation defects in hESCs (Collinson et al. 2016). Consistent with the latter observation, a gene set consisting of PRC2 core components EZH2, SUZ12, EED, and RBBP4 was significantly depleted in the CRISPR proliferation screen (Fig. 4A). Surprisingly, we found an inverse phenotype for PRC1. Several PRC1 components stood out among the highest-ranking enriched hits in the CRISPR proliferation screen (Fig. 2A),

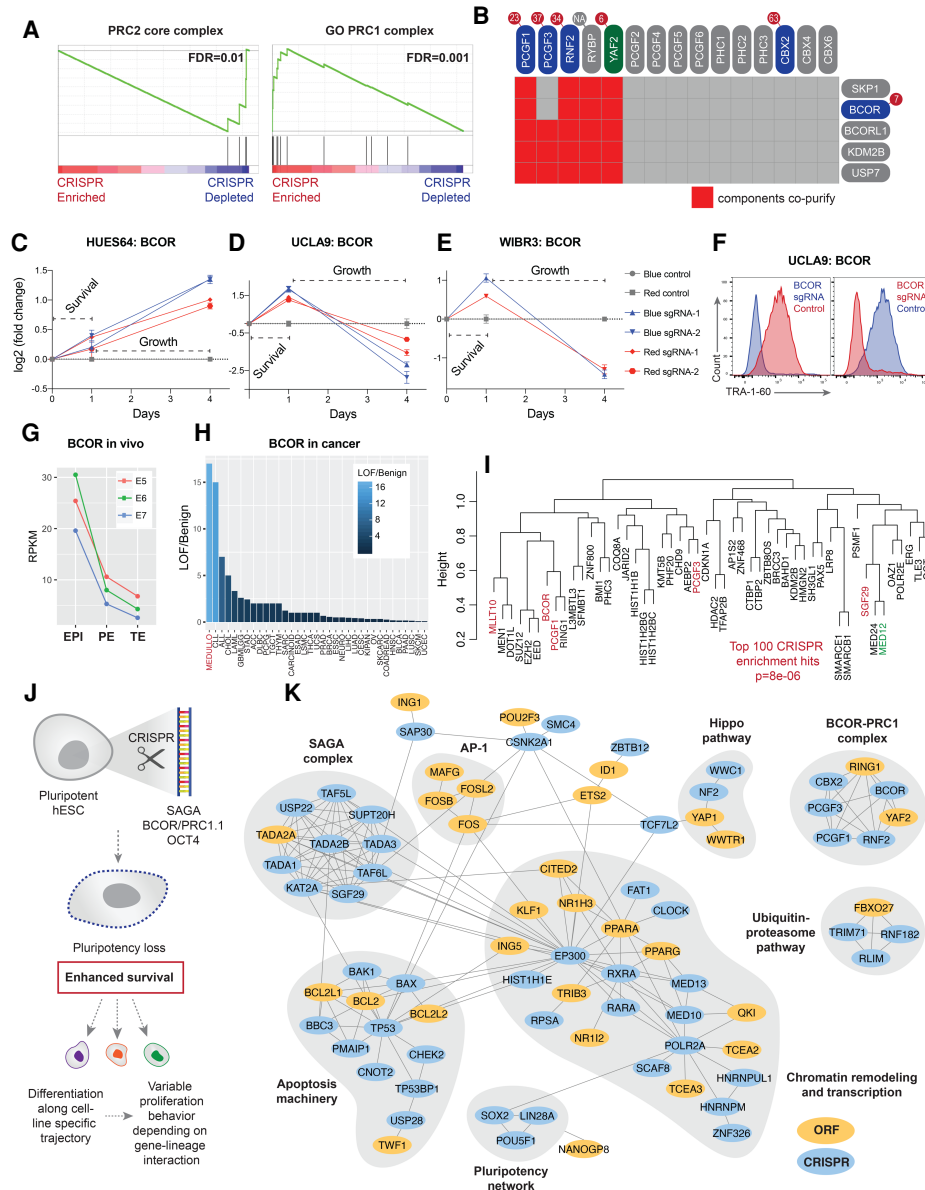


Figure 4. Role of the BCOR–PRC1.1 complex in hESC survival, proliferation, and pluripotency maintenance. (A) Gene set enrichment analysis of PRC1 and PRC2 gene sets in the hESC CRISPR proliferation screen. FDR (GSEA gene set permutation) for PRC1 was adjusted for testing of all GO cellular component gene sets (discovery phase). PRC2 FDR represents a subsequent targeted analysis. (B) Heat map of tandem affinity purification profiles based on data from Gao et al. (2012). PRC1 subunits indicated at the *top* of each column were used as bait, and interacting proteins indicated at the *right* were identified with liquid chromatography–mass spectrometry. Red fields indicate copurification, and gray fields indicate no interaction. The enrichment ranks of genes in the CRISPR (blue) and ORF (green) proliferation screens are plotted within the red circle. Genes in gray did not score among the top 100 enriched genes in either screen. (C–E) Survival and proliferation of HUES64 (C), UCLA9 (D), and WIBR3 OCT4-GFP (E) transduced with sgRNAs against BCOR (blue/red) or a set of control genes (gray). Fold changes were normalized to the control. (F) TRA-1-60 expression in UCLA9 hESCs transduced with sgRNAs against BCOR or a set of control genes. The two panels show the multicolor competition assay color swap. (G) Expression of BCOR in human EPI, PE, and TE cells at embryonic days (E) 5, 6, and 7. Each data point represents the mean of dozens to hundreds of single-cell RNA-seq profiles. (H) Ratios of loss of function (LOF) to benign mutations in BCOR across human cancers. (I) Correlation-based hierarchical clustering of the top 50 BCOR-correlated genes in the DepMap. Top hits from the hESC CRISPR proliferation screen (red) are highly significantly enriched among this set of genes. Genes shown in light-green font scored in the germ layer formation screens (see below). (J) Schematic summary of proliferation screen results. Loss of pluripotency regulators like SAGA, BCOR/PRC1.1, and OCT4 leads to pluripotency loss and increased survival. After cells have exited pluripotency, subsequent proliferation behavior depends on the interaction between the introduced genetic alteration and cell line-specific differentiation propensities. (K) StringDB protein–protein interaction (PPI) network of the top 100 hits enriched in the ORF (yellow) and CRISPR (blue) proliferation screens. Genes with fewer than three connections were removed from the network. The PPI enrichment *P*-value, provided by StringDB to measure whether the network has more interactions than expected by chance, is $<1 \times 10^{-16}$.

and this result was further confirmed by gene set enrichment analysis (Fig. 4A). PRC1 members RING1A and YAF2 were notable hits in the ORF screen, potentially due to dominant-negative effects (Fig. 2C). The highest-ranking CRISPR hit was BCOR, a transcriptional corepressor that is part of a variant PRC1 complex. BCOR mutant cells enriched ~32-fold over the course of the screen. Notably, it was recently reported that BCOR mutations accumulate and expand in induced pluripotent stem cell lines intended for clinical use (Trounson 2017), suggesting that BCOR loss confers a selective advantage in some cell lines.

PRC1 complexes are highly dynamic, and a large number of variants with different subunits exist (Gao et al. 2012). To investigate whether components of a specific PRC1 complex were targeted in our screens, we analyzed mass spectrometry data in which different PRC1 subunits were used as bait to recover interacting proteins (Gao et al. 2012). With the exception of CBX2, all screen hits mapped to the BCOR-containing variant PRC1.1 complex (Fig. 4B). To confirm these results, we investigated survival and proliferation phenotypes of BCOR mutants in three hESC lines using the MCA. We found that BCOR loss leads to significantly enhanced survival during passaging in all three cell lines (Fig. 4C–E). In UCLA9 and WIBR3, we also analyzed expression of the pluripotency marker TRA-1-60 and OCT-GFP, respectively, and found that BCOR loss led to dissolution of pluripotency (Fig. 4F; Supplemental Fig. S6), in line with recent reports (Wang et al. 2018). These data further support our proposition that many of our CRISPR screen candidates represent genes that are required for pluripotency maintenance. In the human embryo, BCOR had a similar expression profile to PMAIP1, with high expression in the epiblast and lower expression in the primitive endoderm and trophectoderm (Fig. 4G).

Interestingly, while pluripotency loss and enhanced survival were a ubiquitous result of BCOR loss, increased proliferation was only observed in HUES64 (Fig. 4C). We wondered whether, upon loss of pluripotency, HUES64 differentiates into a cell type in which BCOR acts as a tumor suppressor, while UCLA9 and WIBR3 differentiate into cell types that are not growth-limited by BCOR. This idea was informed by two lines of evidence. First, HUES64 has a strong tendency toward differentiating toward the neuroectoderm. In a “lineage scorecard” comparison of hESCs, HUES64 had the highest ectoderm propensity score and the second highest neural lineage score among 18 tested hESC lines, indicating that, upon pluripotency loss, HUES64 assumes a pronounced neural progenitor-like phenotype (Bock et al. 2011). Second, studies in ectoderm-derived *Drosophila* tissues reported a dichotomy between PRC1 and PRC2 knockout phenotypes similar to what we observed in HUES64. During larval imaginal eye disc development, loss of PRC2 leads to hypoproliferation, while PRC1 loss results in dramatic hyperproliferation and tumor formation (Martinez et al. 2009; Loubiere et al. 2016). If BCOR indeed acts as a conserved tumor suppressor in developing ectodermal tissues (potentially through restriction of MYC and YAP1) (Sup-

plemental Fig. S6), human cancers derived from the developing ectoderm should be enriched for BCOR loss-of-function (LOF) mutations. To test this hypothesis, we analyzed BCOR mutations in 20,536 samples from the TCGA and COSMIC databases. We plotted the ratio between LOF and benign mutations (an established measure of the selective pressure exerted on a gene) (Davoli et al. 2013) across 36 cancer types and found that LOF mutations in BCOR were most strongly selected in medulloblastoma (Fig. 4H), a childhood cancer that originates in the cerebellum or dorsal brain stem during embryonic development (Gibson et al. 2010). These data nominate BCOR as a candidate growth suppressor in developing ectodermal tissues. Finally, as for TADA2B, we used the DepMap to define BCOR’s genetic neighborhood. Again, there was a highly significant overlap between BCOR-related genes and the top 100 hESC CRISPR screen hits ($P = 8 \times 10^{-06}$) (Fig. 4I). The analysis revealed a surprising genetic association between BCOR complex members, SAGA component SGF29, and Mediator subunits MED12 and MED24, suggesting that at least in some cell lines, including hESCs, these proteins are likely to participate in interconnected processes.

A core gene network controlling pluripotency, hESC growth, and survival

Taken together, our results strongly support a model in which loss of components that are required for pluripotency maintenance (OCT4, BCOR, and a subset of SAGA subunits) leads to pluripotency exit and a concomitant, hardwired rise in apoptosis resistance and survival (Fig. 4J). This effect is reproducible across all tested hESC lines. In contrast, the proliferation rate of cells that have exited pluripotency is cell line- and mutation-dependent. HUES64 cells have a pronounced tendency to assume a neuroectodermal fate upon pluripotency loss, and this fate is likely to influence the genes that confer selective advantage during subsequent growth. TADA2B stands out as a gene whose loss drives increased proliferation across all tested hESC lines.

To integrate the results from CRISPR and ORF screens and thereby create a definitive reference network of genes that control pluripotency, growth, and survival in hESCs, we used the STRING database to combine top enrichment hits in the form of a protein–protein interaction (PPI) network using experimentally validated or curated interactions only (Fig. 4K). In addition to known pluripotency regulators (OCT4, SOX2, and LIN28A), three dominant interconnected subnetworks emerged: the apoptosis machinery, the SAGA complex, and many closely interacting genes that are involved in chromatin remodeling and transcriptional regulation, including EP300, POLR2A, and multiple Mediator complex subunits. Given their physical interactions and genetic associations, these proteins likely function together in controlling fundamental hESC phenotypes. Interestingly, the chromatin remodeling and transcription module also contained the retinoic acid receptors RARA and RXRA (identified in the CRISPR screen), which form a heterodimer and activate the

expression of differentiation-associated genes (Gudas and Wagner 2011). ORF hits included nuclear receptors PPARA and PPARG, which compete with RARA for binding to RXRA (van Neerven et al. 2008). RARA/RXRA signaling leads to chromatin decompaction, recruitment of the transcriptional machinery to target genes, and the initiation of lineage-specific gene expression (Gudas and Wagner 2011). Notably, KAT2A (also known as GCN5), the histone acetyltransferase component of the SAGA complex, has been shown to act as a coactivator of retinoic acid signaling (Vilhais-Neto et al. 2017). These results raised the intriguing possibility that in addition to its role in pluripotency maintenance, the SAGA complex could have a role in lineage specification.

Novel differentiation regulators emerge from germ layer formation screens

We examined genes that control germ layer formation by comparing cells collected in the “correctly differentiated” and “undifferentiated” gates (Fig. 1A; Supplemental Note on screen performance) with each other. Differentiation efficiencies varied by protocol: Endoderm and mesoderm induction media consistently induced differentiation of >95% of cells, while ectoderm induction media typically only yielded 15%–50% of differentiated cells. For the endoderm and mesoderm, we therefore focused our analyses on genes that interfere with differentiation when overexpressed or inactivated, since statistical power to detect genes that are depleted in the small number of undifferentiated cells (corresponding to enrichment in the differentiated population) (see the tables in the Supplemental Note on screen performance for exact cell numbers) is low. Differentiated and undifferentiated population sizes were much more balanced for the ectoderm (Fig. 1A), enabling us to examine both sides for that germ layer.

Genes and pathways with known functions in germ layer formation behaved as expected. In the CRISPR screens, the transcription factor T (Brachyury), a master regulator of mesodermal differentiation (Yamaguchi et al. 1999), was the number 5 gene required for mesoderm formation (Fig. 5A; Supplemental Fig. S7; Supplemental Table S1), while FOXH1, an important transcription factor mediator of Nodal signaling (Yamamoto et al. 2001), was the top gene required for endoderm formation (Fig. 5B). ZIC2, a gene whose mutation causes severe brain malformation (Nagai et al. 2000), was specifically required for ectoderm formation (Fig. 5A,B). To understand functional gene categories required for correct differentiation in greater detail, we performed GO enrichment analysis. Endoderm development, nodal/activin receptor signaling, and SMAD protein complex assembly (represented by FOXH1, SMAD2, SMAD3, and ACVR1B) were among the top enriched categories in the endoderm CRISPR screen (Fig. 5C), consistent with results in other hESC lines and the known function of the nodal signaling pathway in mediating endodermal differentiation (Li et al. 2019). Similarly, several genes that scored in the mesoderm CRISPR screen belonged to GO categories “mesoderm formation” (T, MSGN1, SNAI1, and WNT5A) and “primary germ layer formation.” Meso-

derm hits were also enriched for genes belonging to *trans*-synaptic signaling (e.g., KCNC4) (Fig. 5C). The SAGA complex scored as the most important functional category required for ectoderm formation (Fig. 5C). (Additional details on SAGA’s role in differentiation are presented below.) Notably, BCOR scored as the top gene enriched in differentiated versus undifferentiated ectoderm in the MAGeCK analysis of the CRISPR P2 sublibrary screen, further supporting the notion that BCOR acts as a growth suppressor in the developing ectoderm (Supplemental Fig. S7).

As in the proliferation setting, ORF germ layer formation screens complemented the CRISPR results and touched on many similar biological themes (Fig. 5D–F; Supplemental Fig. S7). Genes whose expression blocked ectoderm, endoderm, or mesoderm formation were predominantly developmental transcription factors represented by GO categories “regulation of transcription by RNA polymerase II,” “anatomical structure morphogenesis,” and “cell fate commitment” (Fig. 5F). The main signaling pathways known to mediate differentiation signals into the three germ layers were clearly delineated in the screen results. WNT signaling drives mesodermal differentiation (Yamaguchi et al. 1999); consequently, expression of negative WNT pathway regulators (CTNBP1, TLE1, and TLE4) strongly inhibited mesoderm formation (Fig. 5G), and WNT signaling regulation was one the significant GO categories enriched among mesoderm-inhibiting ORFs (Fig. 5F). Neuroectoderm differentiation requires the absence of bone morphogenetic protein (BMP) signals (Stern 2006), a fact that was faithfully recapitulated in the ectoderm screen, where positive regulators of BMP signaling in particular (BMP6/7), and members of the TGF- β superfamily in general (GDF7 and INHBB), acted as ectoderm formation inhibitors (Fig. 5G). GO categories “cellular response to BMP stimulus” and “regulation of MAPK cascade” were significantly enriched among ectoderm-inhibiting ORFs (Fig. 5G). In contrast, BMP/TGF- β inhibitors (GREM2, NOG, LDLRAD4, and SMAD6) enhanced ectoderm differentiation (Fig. 5G). Finally, genes implicated in the epithelial-to-mesenchymal transition (SNAI1 and SNAI2) and Hippo signaling (YAP1 and WWTR1) strongly inhibited endoderm formation (Fig. 5F). Overall, the results anecdotally suggested that proper differentiation of a germ layer is disrupted by ectopic expression of genes that are normally expressed later in development (e.g., during organogenesis) or in other lineages. For example, expression of SNAI1 and MSGN1, genes that are required for proper mesoderm development (Carver et al. 2001; Chalamalasetty et al. 2014), potentially blocked endoderm formation (Fig. 5E). Overexpression of POU3F4, a transcription factor essential to inner ear development (Phippard et al. 1999), was incompatible with mesodermal differentiation (Fig. 5D), as was expression of PAX8, a transcription factor playing a major role in kidney and urinary tract development (Sharma et al. 2015). Of note, some particularly potent mesoderm-specific transcription factors (T and MSGN1) scored as ectoderm enhancers, probably because their overexpression is sufficient to drive cells into the

EpCAM⁻/NCAM⁺ gate even in the absence of mesoderm induction media (Fig. 5D).

After confirming that known genes and pathways implicated in germ layer formation behaved as expected, we turned our attention to novel regulators identified in our screens. For validation purposes, we wanted to narrow the significant hits across germ layers down to a smaller set of genes that were most likely to have large functional importance. Therefore, we compared our results with published ENCODE data of histone modification changes during ectoderm, mesoderm, and endoderm formation in HUES64 (Gifford et al. 2013). We performed a genome-wide quantification of the relative change of the activating histone mark H3K4me3 during differentiation and visualized promoters that strongly gained H3K4me3 in particular germ layers. Known transcription factors behaved as expected in this analysis: Ectoderm-specific OTX1 and OTX2 gained H3K4me3 during ectodermal but not mesodermal differentiation, while GATA4 and SNAI1 gained H3K4me3 specifically in mesodermal cells (Fig. 5H). Among the most conspicuous genes emerging from this analysis was SPRY4, a negative regulator of receptor tyrosine kinase signaling (Felfly and Klein 2013) that massively gained H3K4me3 during mesoderm differentiation and simultaneously lost this mark during ectoderm formation (Fig. 5H). SPRY4 also was the number 2 gene required for mesoderm formation in our CRISPR screen (outperforming even T) (Fig. 5A). SPRY4 has previously been implicated in hematopoiesis (Mendenhall et al. 2004), but to our knowledge has not been shown to play a role in early mesoderm formation. We validated the effects of SPRY4 knockout using the MCA and observed a highly significant blockade of mesodermal differentiation in mutants (Supplemental Fig. S8). An analogous analysis for H3K4me3 gain during endoderm formation revealed a small group of genes that strongly gain H3K4me3 during endodermal but not ectodermal differentiation (Fig. 5I). This group included known endodermal differentiation mediators such as GATA4, SOX17, and GSC, but also MANEA, a poorly characterized endomannosidase located in the Golgi apparatus. In our CRISPR screen, MANEA was among the top 20 genes required for endoderm formation (Fig. 5B). Multicolor competition assays confirmed that MANEA mutants show impaired ability to differentiate into CXCR4-expressing endoderm (Supplemental Fig. S8). We also further investigated PAGR1 (PAXIP1-associated glutamate-rich protein 1), which scored as the number 2 gene required for endoderm formation, along with its interaction partner PAXIP1 (no. 14 required gene). The PAXIP1/PAGR1 pair already stood out in the DepMap-based analysis of tightly correlated gene groups (Fig. 3F) and has been suggested to function in the relay of TGF- β signals (Baas et al. 2018). We confirmed with the multicolor competition assay that PAGR1 mutants indeed showed severe endoderm formation defects (Supplemental Fig. S8). We also validated MED12, the number 3 gene required for ectoderm formation (after SAGA complex members SGF29 and TADA2B) and observed strong differentiation blockade in mutants (Supplemental Fig. S8). Finally, we validated the effects

of SNAI1 expression on endoderm formation, noting that as seen in our screen results, SNAI1 blocked endodermal differentiation, but enhanced mesoderm and ectoderm formation (Supplemental Fig. S8).

A novel insight emerging from the ORF differentiation screens was the potent influence of an unexpected gene family, the keratin-associated proteins (KRTAPs), on endoderm and ectoderm formation. KRTAPs were recently discovered as tissue-specific drivers of proliferation in human mammary epithelial cells (an ectodermal derivative) (Sack et al. 2018). In the ectoderm formation screen, KRTAPs scored as potent enhancers of differentiation (or possibly as proliferation drivers in differentiated ectodermal cells) (Fig. 5J). They had the opposite effect in the endoderm formation screen, and (as a group) behaved neutrally in the mesoderm screen. KRTAPs' influence on ectoderm and endoderm differentiation was large: In both cases, the GO category "keratin filament," which contains mostly KRTAPs, was the most significant out of 434 tested gene sets in the GO "cellular component" category. Among the many different KRTAP genes—the family consists of 101 members that are chiefly expressed in hair follicles (Wu et al. 2008) in mature tissues—KRTAP4 and KRTAP10 subfamily members were the most potent "lineage switches" (Fig. 5K). In conjunction with the observation that the same KRTAP family members can drive proliferation in mammary epithelium, these results suggest the interesting possibility that tissue-specific proliferation drivers found in mature cells (Sack et al. 2018) are partially established in early development already (or, alternatively, the pathways that allow them to be sensed are established).

Identification of general differentiation inhibitors and enhancers

Beyond identifying genes that control differentiation into specific cell types, our screen design allowed us to pinpoint genes that are universally required (CRISPR) or universally incompatible (ORF) with differentiation in HUES64. We found five ORFs whose expression blocked differentiation into all three germ layers (Fig. 6A). Two of the four Yamanaka reprogramming factors (KLF4 and MYC) (Takahashi and Yamanaka 2006) were in this group ($P = 3 \times 10^{-7}$). OCT4 and SOX2 were not in the ORF library. Among the five genes, interferon regulatory factor 4 (IRF4), a transcription factor that is predominantly expressed in the hematopoietic system (Nam and Lim 2016), acted as the most potent differentiation inhibitor, outperforming both MYC and KLF4. Genes involved in the regulation of interferon signaling were also significantly enriched among CRISPR proliferation screen hits (Fig. 2B). This is notable in the context of a recent study that showed that interferon-stimulated genes are highly expressed in hESCs and down-regulated upon differentiation (Wu et al. 2018). IRF4 was ranked as the number 1, 2, and 18 differentiation inhibitor in the endoderm, mesoderm, and ectoderm screens, respectively (average rank of 7 across all three screens). MYC's average rank was 16.7, and KLF4's rank was 86.3. ZNF398 (average rank

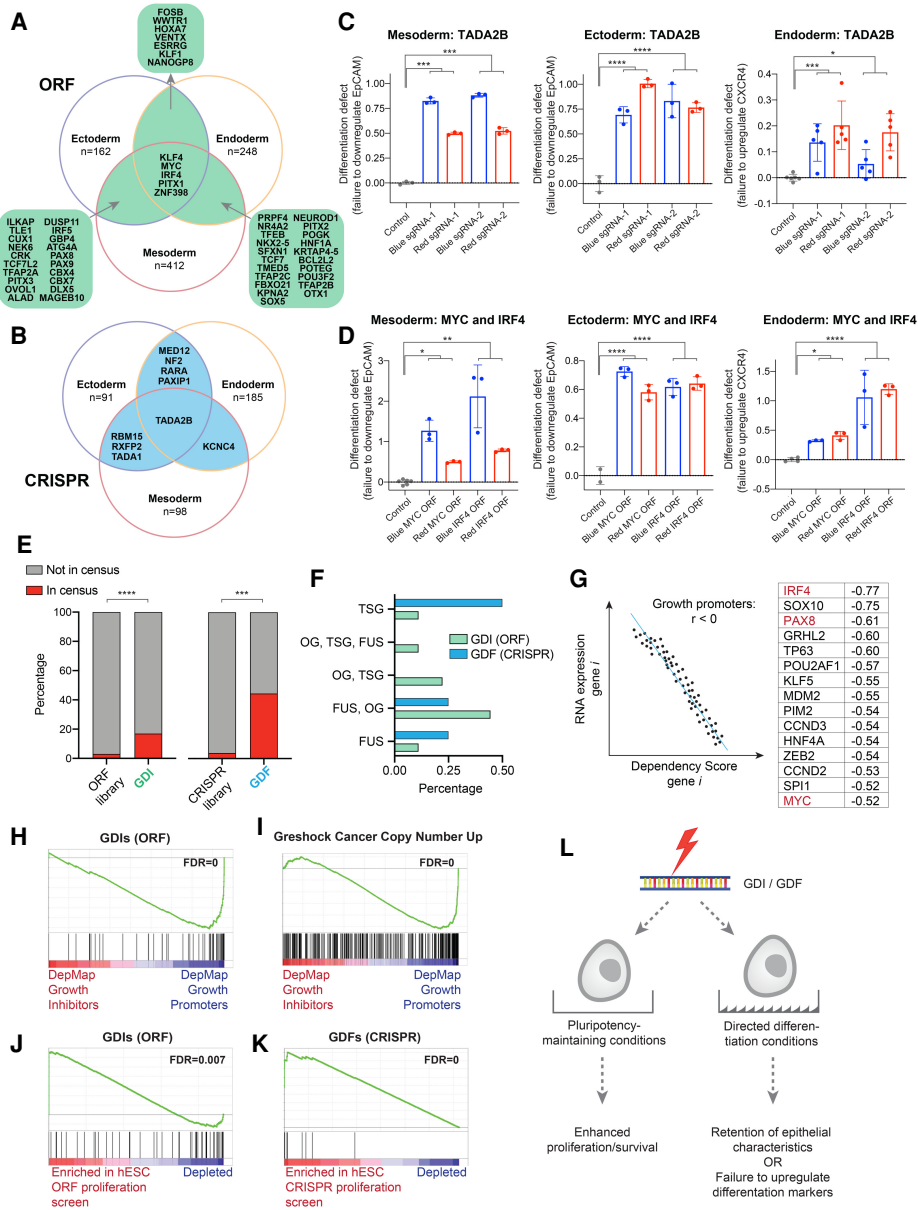


Figure 6. General differentiation facilitators and inhibitors in CRISPR and ORF germ layer formation screens. (A) Venn diagram showing the overlap between genes that scored as germ layer formation inhibitors when overexpressed ($P < 0.01$). Genes that overlap between germ layers are listed in full. (B) As in A for CRISPR. The Venn diagram shows the overlap of genes that are required for the formation of multiple germ layers; i.e., act as general differentiation facilitators. (C) Multicolor competition assays showing differentiation defects in HUES64 transduced with sgRNAs against TADA2B (blue/red) or a set of control genes (gray). The differentiation defect (Y-axis) reflects the inability of mutants to down-regulate EpCAM (mesoderm/ectoderm) or up-regulate CXCR4 (endoderm). See the Materials and Methods for details on the definition of differentiation defect values. Error bars show standard deviation. P -values were derived from one-way ANOVA, comparing all conditions against the control group and adjusted with Dunnett’s multiple comparisons test. (D) As in C for MYC and IRF4 overexpression. (E) Overlap between GDIs and GDFs and genes in the Cancer Gene Census. P -values from Fisher’s exact test. (F) Cancer Gene Census classification of GDIs (ORF) and GDFs (CRISPR). Annotation for each gene is shown exactly as in the Cancer Gene Census. For some genes, multiple functions have been reported; e.g., IRF4 is annotated as OG, TSG, and FUS. (TSG) Tumor suppressor gene, (OG) oncogene, (FUS) fusion. (G) DepMap genes ordered by the correlation between their dependency scores and expression levels. Low dependency scores indicate slower growth upon gene loss; high scores indicate improved growth upon gene loss. Negative correlations were observed for genes that tend to be more essential as their expression increases (growth-promoting genes). The scatter plot at the left represents hypothetical data to illustrate the relationship. (H) ORF general differentiation blockers are significantly enriched among growth-promoting genes. (I), Genes that are amplified in cancer are significantly enriched among growth-promoting genes. (J,K) Gene set enrichment analysis showing the behavior of GDIs (J) and GDFs (K) during ORF and CRISPR hESC proliferation screens, respectively. (L) Schematic summarizing the association between differentiation-inhibiting and growth-promoting genetic alterations in hESCs. (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$, (****) $P < 0.0001$.

216) is an ER α -interacting transcription factor (Conroy et al. 2002) with largely unexplored function. We found that it is specifically expressed in the human epiblast and down-regulated upon differentiation into trophoblast and primitive endoderm (Supplemental Fig. S9). Finally, PITX1 (average rank 15.7) is a homeodomain transcription factor whose loss has been linked to limb abnormalities (Klopocki et al. 2012). On the CRISPR side, only one gene was required for the formation of all three germ layers: TADA2B, the SAGA complex member that already stood out in the CRISPR proliferation screen (Fig. 6B). TADA1, another SAGA complex member, was among the three genes necessary for both ectoderm and mesoderm formation (Fig. 6B).

A vital role for TADA2B and the SAGA complex in differentiation processes

We performed MCAs in HUES64 to validate the effects of TADA2B loss on differentiation efficiency. TADA2B mutants had defects in the formation of all three germ layers (Fig. 6C). For the ectoderm and mesoderm, failure to down-regulate EpCAM was more pronounced than failure to up-regulate NCAM (Supplemental Fig. S9A) indicating that SAGA was less vital for the induction of differentiation-associated genes than for the correct down-regulation of the epithelial-like hESC phenotype. Consistent with this observation, the differentiation defect was least pronounced for the endoderm, where, according to established protocols (Gifford et al. 2013), we only measured CXCR4 up-regulation without monitoring EpCAM down-regulation. The requirement for TADA2B during endoderm, mesoderm, and ectoderm differentiation was further validated in WIBR3 hESCs (Supplemental Fig. S9B). Note that many other genes that are important for pluripotency maintenance (e.g., OCT4 or SOX2) did not score in any of the differentiation screens discussed above, highlighting the unique dual role of TADA2B and the SAGA complex in the regulation of both pluripotency and differentiation.

IRF4 expression blocks germ layer formation

Like TADA2B loss, IRF4 expression potently blocked differentiation into all three germ layers (Fig. 6D). We validated it side by side with MYC for comparative purposes and, as in the screens, its effects were more pronounced than MYC's in two out of three germ layers. Notably, as for TADA2B, the ectoderm and mesoderm differentiation defects were more severe with respect to EpCAM down-regulation than to NCAM up-regulation (Supplemental Fig. S10A). Given the differentiation defect similarities between TADA2B loss and IRF4 expression, we next wanted to evaluate whether elevated IRF4 levels also lead to pluripotency exit. While MYC overexpression led to a small but reproducible reduction in TRA-1-60 expression in HUES64 (Supplemental Fig. S10), IRF4 expression slightly increased overall TRA-1-60 levels, indicating that, unlike SAGA loss, IRF4 expression does not result in overt pluripotency dissolution. To understand tran-

scriptional regulation by IRF4 in greater detail, we compared published gene expression data of IRF4^{-/-} and IRF4-overexpressing primary B cells (Ochiai et al. 2013). As expected, genes up-regulated in IRF4-expressing cells were highly enriched for IRF4 targets identified by independent studies (Supplemental Fig. S10C). Notably, the third most up-regulated transcript in IRF4-expressing cells (after IRF4 itself and Rhoophilin-2) was EpCAM, suggesting that IRF4 may in part block differentiation by sustaining EpCAM expression (consistent with our flow cytometry results). Gene set enrichment analysis further revealed that MYC targets were significantly up-regulated in IRF4-expressing cells (Supplemental Fig. S10C), in line with a known role for IRF4 in MYC activation (Weilemann et al. 2015) that could partially contribute to IRF4's effects in hESCs. Interestingly, several gene sets representing pluripotency-specific gene signatures were also significantly up-regulated in IRF4-expressing cells (Supplemental Fig. S10C), further suggesting the possibility that IRF4 may drive the expression of a subset of pluripotency-associated genes and thus interfere with differentiation.

General differentiation inhibitors and facilitators are cancer drivers

We noted that like MYC, IRF4 acts as a potent oncogene in several malignancies, including multiple myeloma (Shaffer et al. 2008) and T-cell lymphoma (Boddicker et al. 2015). This raised the question of whether ORFs that broadly inhibit hESC differentiation are enriched for oncogenes. To enable a broader analysis, we defined general differentiation inhibitors (GDIs) as ORFs that blocked formation of at least two germ layers (corresponding to the overlapping fields of the Venn diagrams in Fig. 6A, $n = 53$). Similarly, we defined as general differentiation facilitators (GDFs) genes that were required for the formation of at least two germ layers (overlapping fields of the Venn diagram in Fig. 6B, $n = 9$). Both GDIs and GDFs were significantly more likely to be cancer driver genes (as defined by the COSMIC Cancer Gene Census) than expected by chance ($P = 2 \times 10^{-4}$ for GDFs and $P = 3 \times 10^{-5}$ for GDIs, Fisher's exact test) (Fig. 6E). Examining their role in the census annotation, we found that the largest fraction of GDIs was classified as fusion oncogenes (FUS and OG), while the largest fraction of GDFs was annotated as tumor suppressors (TSG) (Fig. 6F). This is consistent with the recovery of these genes from gain-of-function and loss-of function screens, respectively. We conclude that genetic alterations that broadly inhibit correct differentiation of hESCs coincide with genetic alterations that promote cancer development in somatic tissues.

To study the connection between general differentiation regulators and cancer drivers further, we returned to the DepMap. Cancer cells often depend on the expression of specific oncogenes (e.g., MYC) (Jain et al. 2002) to sustain their growth and survival, a phenomenon termed "oncogene addiction" (Weinstein 2002). Oncogenes exist in three partially overlapping classes: those that can be activated by point mutations (e.g., KRAS),

those that gain a novel function (e.g., IDH1), or those that are primarily activated by overexpression (e.g., MYC). The latter class is the most difficult to identify. We reasoned that we could identify genes that behave like oncogenes (or, more generally, growth-promoting genes) by evaluating whether dependency on a gene rises with its expression level (Fig. 6G). The advantage of extracting oncogenes from the DepMap in this manner—over the use of categorical lists as found in the Cancer Gene Census—is that the strength of the expression–dependency correlation allows for a straightforward quantitative ranking of growth-promoting genes. Remarkably, after calculating correlations between expression and dependency score across all genes, we found that IRF4 emerged as the number 1 growth-promoting gene (Fig. 6G). PAX8 was third, and MYC was number 15. To determine whether GDIs generally overlapped with growth-promoting genes, we conducted gene set enrichment analysis on the basis of genes ranked by their dependency score–expression correlation. We found a highly significant enrichment of GDIs among growth-promoting genes (FDR = 0) (Fig. 6H). Genes that are frequently amplified in cancer were also enriched among DepMap growth promoters (FDR = 0) (Fig. 6I) and furthermore significantly overlapped with GDIs ($P = 0.016$, Fisher’s exact test).

Genetic alterations that inhibit differentiation drive proliferation

These results demonstrated that ORFs whose expression widely perturbs differentiation decisions in our germ layer screens coincide with growth-promoting genes in general and oncogenes in particular. This is consistent with the widely documented observation that many cancer-causing mutations (in both oncogenes [e.g., CTNNB1] and tumor suppressors [e.g., APC or NOTCH1]) function by disturbing the balance between proliferation and differentiation in tissue stem cells and their progeny (Vogelstein et al. 2013; Alcolea et al. 2014). Under the model of a dynamic proliferation–differentiation equilibrium, genetic alterations that inhibit correct differentiation thus (directly or indirectly) also confer increased proliferative and survival capacity. To determine whether such an equilibrium was detectable in our paired proliferation–germ layer formation screens, we tested whether ORF general differentiation inhibitors were enriched among top scoring genes in the ORF proliferation screen. Indeed, we found a highly significant overlap between the 100 most enriched genes in the ORF proliferation screen and ORF general differentiation inhibitors (overlap = 6, $P = 5.5 \times 10^{-6}$, Fisher’s exact test). Gene set enrichment analysis further confirmed that expression of genes that inhibit differentiation drives proliferation (FDR = 0.007) (Fig. 6J). The significant overlap of differentiation inhibitors and proliferation enhancers was mainly driven by FOSB, BCL2L2, KLF1, ESRG, WWTR1, and NANOGP8, which scored with significant effect sizes in the proliferation screen. Similarly, in the loss-of-function setting, the overlap between the 100 most enriched genes in the CRISPR proliferation screen and GDFs was highly significant (overlap = 5,

$P = 5.7 \times 10^{-10}$, Fisher’s exact test) and could be further confirmed by gene set enrichment analysis (FDR = 0) (Fig. 6K). TADA2B, TADA1, RARA, NF2, and RBM15 were the most dominant contributors to this effect.

Together, these results show that genetic alterations that interfere with differentiation by causing abnormal retention of epithelial characteristics and/or a failure to up-regulate differentiation-associated markers are also disproportionately involved in driving hESC proliferation (Fig. 6L). The same characteristics can often be observed for genes that act as oncogenes or tumor suppressors in somatic tissues (Vogelstein et al. 2013). A prime example explored in detail here is TADA2B, a gene that controls hESC pluripotency, survival, and proliferation, and also plays a central role in differentiation decisions.

Discussion

The genetic regulation of hESC proliferation, survival, and differentiation is of immense interest, as these processes set the stage on which further human development unfolds and are critically important for regenerative medicine applications. Consequently, the epigenetic landscape of hESCs and their differentiated derivatives has been studied in great detail (Mikkelsen et al. 2007; Gifford et al. 2013; Xie et al. 2013; Tsankov et al. 2015). Here we layer on this knowledge functional gene information derived from genome-scale loss- and gain-of-function perturbations of hESC proliferation and differentiation.

Among the most interesting insights emerging from our integrated screening approach are the manifold interconnections between proliferation, survival, pluripotency, and differentiation phenotypes. These interconnections are hardwired through specific genetic networks and likely serve particular and important purposes. For example, an intimate relationship between pluripotency and survival capacity (illustrated in detail by our studies of TADA2B, SUPT20H, BCOR, and OCT4 loss) may be required to ensure that damaged cells apoptose with high efficiency in early embryonic development, thus preventing compromised progeny from contributing to the developing embryo. The evolutionary cost of failed gestation is immense and under significant evolutionary pressure.

Increased survival upon pluripotency loss appears to be a universal phenomenon. In contrast, proliferation phenotypes after loss of key pluripotency-maintaining genes are variable and cell line-dependent, as illustrated by our comparative analysis of BCOR and OCT4 loss across different cell lines. Evidently, the precise effect of gene knockout depends on the cell type into which an hESC line differentiates after pluripotency resolution. This is consistent with the highly tissue-specific nature of proliferation control (Sack et al. 2018) and has implications for future work in hESC lines, which have a reputation for excessive variability in their response to genetic perturbations. Our results indicate that this variability is a function of lineage bias and may thus be able to be predicted and controlled for prospectively. TADA2B stands out

as a gene that maintains pluripotency and limits proliferation and survival across all tested cell lines, suggesting that it holds a unique position in the hESC control network.

Another fascinating connection is the relationship between proliferation/survival and differentiation phenotypes, discernible in both the loss- and gain-of-function setting. What are the molecular underpinnings of this link? A variety of mechanisms may be at play. First, we have to consider the possibility that loss of pluripotency simply leads to reduced lineage potential and impaired differentiation. Our screening results suggest that this is not the main explanation, as loss of classical pluripotency regulators like OCT4 or SOX2 did not result in a differentiation phenotype in any of the germ layer formation screens. Loss of BCOR even enhanced ectoderm differentiation. A second possibility is that the cell cycle directly regulates differentiation and vice versa, a notion for which extensive support exists in the literature (Liu et al. 2019). ESCs have an unusually short G1 phase that lengthens upon cell fate specification (Dalton 2015). G1 has generally been found to associate with increased sensitivity to differentiation signals; its length may therefore determine the balance between self-renewal and differentiation (Hardwick and Philpott 2014). Proteins expressed at different cell cycle stages appear to have dense interactions with differentiation pathways. For example, CDK4/6-cyclin D complexes accumulating during G1 phase repress Activin/Nodal signaling and inhibit mesendodermal differentiation (Pauklin and Vallier 2013), and cyclin D interacts with transcription factors to modulate developmental gene expression directly (Pauklin et al. 2016). ATM/ATR signaling during S phase directly supports pluripotency maintenance through enhancement of TGF- β signaling (Gonzales et al. 2015). Furthermore, cell cycle exit, such as that afforded by CDK inhibitors p21 and p57, is known to be required for muscle cell differentiation (Zhang et al. 1999). Thus, it is likely that genetic alterations that drive progression through the cell cycle and shorten G1 phase will inhibit differentiation. However, most of the genes that scored as GDIs or GDFs with simultaneous proliferation phenotypes were not classical cell cycle regulators such as cyclins or CDK inhibitors. Instead, they tended to be genes that are involved in the transduction of differentiation signals. Examples include RARA (CRISPR/retinoic acid signaling), MED12 (CRISPR/WNT signaling), NF2 (CRISPR/Hippo), and WWTR1 (ORF/Hippo). This raises the interesting third possibility that some genetic networks that are involved in differentiation initiation additionally “moonlight” as regulators of proliferation and/or survival under pluripotency-maintaining conditions to coordinate cell proliferation and differentiation processes. A well-known example of such moonlighting is the aforementioned Activin/Nodal pathway, which is required for endoderm formation but also is essential for the maintenance of pluripotency (Vallier et al. 2005; Pauklin and Vallier 2015). Our results suggest that WNT, retinoic acid, and Hippo pathways, as well as pathways that function through SAGA/Mediator, may behave similarly.

The connection between proliferation/survival and differentiation is furthermore particularly interesting because it plays a large role in carcinogenesis. Hyperproliferation coupled to differentiation defects is a hallmark of most malignancies, evident on the morphological as well as molecular levels (Ben-Porath et al. 2008; Naxerova et al. 2008). We observed a significant enrichment of cancer drivers among GDFs and GDIs, with ORF screen hits matching oncogenes and CRISPR screen hits matching tumor suppressors. While not all GDFs/GDIs have proliferation phenotypes, and not all proliferation regulators affect differentiation, the significant overlap between the two categories suggests the existence of a prime pool of “dual regulators” that is subject to selection in different cancer types. For example, both WWTR1 overexpression (ORF/GDI) and NF2 loss (CRISPR/GDF) drive proliferation/survival under pluripotency-maintaining conditions, inhibit differentiation into endoderm and ectoderm, and act as cancer drivers across multiple tumor types (as an oncogene and tumor suppressor gene, respectively). It is conceivable that alteration of such dual regulators is particularly useful from the perspective of an incipient tumor because in addition to gaining a proliferative or survival advantage through the alteration, the simultaneous built-in differentiation defect supports continued rapid cell cycle progression through additional independent mechanisms (such as maintenance of a short G1 phase upon failure to differentiate) (see the discussion above), resulting in particularly robust hyperproliferation phenotypes. Further dissection of the genetic regulation underpinning these linked cellular phenotypes will hopefully contribute to a deeper understanding of development and carcinogenesis alike.

The eight gain-of-function and loss-of-function screens performed in this study have provided the most extensive examination of gene functionality in hESCs to date. Altogether, they have implicated hundreds of genes in critical aspects of hESC proliferation, survival, pluripotency maintenance, and control of mesoderm, ectoderm, and endoderm differentiation. The pairing of loss- and gain-of-function analyses has been particularly powerful in identifying rate-limiting steps in processes controlling hESC homeostasis and differentiation, with important connections between fundamental developmental roles and tumorigenesis. The extensive nature of these analyses provides a genetic blueprint underpinning early decisions in development and will act as a foundation on which to guide future exploration of these factors in processes critical to human health and disease.

Materials and methods

Libraries, lentiviral transduction, and genome-wide screening in HUES64

Both CRISPR and ORF libraries have been described previously (Wang et al. 2015; Martin et al. 2017; Sack et al. 2018). The CRISPR library comprised 18,166 genes with five gRNAs per gene and was subdivided into three pools (P1, P2, and P3). To keep the screen size manageable, we screened P1/P3 and P2 separately. The doxycycline-inducible ORF library corresponded to

“Library 2” in our previously published study (Sack et al. 2018) and comprised 13,255 unique genes. To generate an rtTA-expressing population suitable for screening with the ORF library, HUES64 cells were transduced with pInducer20-EGFP lentivirus (Meerbrey et al. 2011) and induced with 1 $\mu\text{g}/\text{mL}$ doxycycline (dox) for 2 d. Subsequently, 40%–45% of the cells with the highest EGFP signal were sorted and expanded. This population was then used for transduction with the ORF library.

CRISPR and ORF libraries were screened at a low multiplicity of infection (0.2–0.3) and at a representation of 500. All screens were performed in duplicate (i.e., consisted of two separately transduced pools). Our protocol for virus production in 293T has been described elsewhere (Sack et al. 2018). One day before lentiviral transduction, HUES64 cells were dissociated into single cells with Accutase (Thermo Fisher) and plated in 10-cm dishes at a density of 60,000–80,000/cm² in the presence of 10 μM ROCK inhibitor Y-27632 (Stem Cell Technologies). The next day, lentivirus was added along with 4 $\mu\text{g}/\text{mL}$ polybrene to facilitate infection. Twenty-four hours later, cells were placed into media containing 1 $\mu\text{g}/\text{mL}$ puromycin and maintained in these media until an uninfected control plate was dead. Puromycin-free media were used for the remainder of the screens. For CRISPR screens, we waited 6–10 d after transduction before collecting the first reference time point (PD0) to allow sufficient time for introduction and repair of double-strand breaks by Cas9. For the ORF screen, cells were allowed to recover for 3 d after puromycin selection was complete and were then placed into media containing 1 $\mu\text{g}/\text{mL}$ dox for 1 d before collecting the PD0 time point. Subsequently, cells were maintained for 13 (ORF) to 16 (CRISPR) population doublings in mTeSR1 media (supplemented with 1 $\mu\text{g}/\text{mL}$ dox in the ORF screen), dissociated with Accutase every 4 d, and replated as single cells with 10 μM ROCK inhibitor at a density of $\sim 18,000/\text{cm}^2$. During every passage, cell pellets were collected and immediately frozen at -80°C . Methods for preparation of genomic DNA from cell pellets, PCR amplification of sgRNAs, and ORF barcodes and sequencing have been described previously (Martin et al. 2017; Sack et al. 2018).

Germ layer differentiation and fluorescence-activated cell sorting

For ectoderm formation screens, transduced single cells were plated at a density of $\sim 55,000/\text{cm}^2$ in the presence of ROCK inhibitor. On the next day (day 1), cells were exposed to previously described ectoderm induction media (Gifford et al. 2013) containing 2 μM BMP inhibitor dorsomorphin (Tocris 3093), Wnt inhibitor PNU 74654 (Tocris 3534), and TGF- β inhibitor A 83-01 (Tocris 2939). For mesoderm formation, cells were plated at the same density as for ectoderm and exposed to commercially available mesoderm induction media (Stem Cell Technologies 05221). For endoderm differentiation, cells were plated at a high density ($\sim 220,000/\text{cm}^2$) and exposed to commercially available endoderm induction media (Stem Cell Technologies 05110). The correct performance of these protocols was confirmed by staining for lineage-specific transcription factors with a germ layer antibody kit (R&D SC022), according to the manufacturer's instructions. On the fifth day of differentiation, cells were dissociated with Accutase and sorted on a FACSria instrument (BD Biosciences) in the presence of ROCK inhibitor. Gates were as shown in Figure 1A, with adjustments depending on the differentiation efficiencies achieved for a specific library. We used the following antibodies: PerCP-Cy 5.5 mouse antihuman EpCAM (BD Biosciences 347199), CD56 PE clone NCAM16.2 (BD Biosciences 340724), and PE-Cy 5 mouse antihuman CD184 (BD Biosciences 555975). For TRA-1-60 staining, we used antihuman TRA-1-60 antibody and clone TRA-1-60R (Stem Cell Technologies

60064PE). To stain differentiated germ layers for immunofluorescence imaging, we used antibodies from the human three-germ layer three-color immunocytochemistry kit (R&D SC022).

Additional Materials and Methods are available in the Supplemental Material.

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We thank Laurence Daheron, Veronica Akopian, and Ramona Pop for advice regarding hESC culture, germ layer differentiation, and FACS methodology, and Danny Reinberg, James Stafford, and Zhonghua Gao for providing PRC1 mass spectrometry data. This work was supported by National Institutes of Health grants to S.J.E. (AG11085 and CA234600), M.I.K. (GM126944), and K.N. (R37CA225655 and P01HL142494). S.J.E. is an Investigator with the Howard Hughes Medical Institute.

Author contributions: K.N. and S.J.E. designed the experiments. K.N., B.D.S., E.V.W., M.A.d.K., M.D., and T.D.M. performed experiments. K.N., J.L.M., B.D.S., E.V.M., D.R., and E.C.W. analyzed data. K.N., B.D.S., J.L.M., M.I.K., K.H., and S.J.E. discussed and interpreted results. K.N. wrote the manuscript which was read and edited by all coauthors. S.J.E. supervised and directed the study.

References

- Alcolea MP, Greulich P, Wabik A, Frede J, Simons BD, Jones PH. 2014. Differentiation imbalance in single oesophageal progenitor cells causes clonal immortalization and field change. *Nat Cell Biol* **16**: 612–619. doi:10.1038/ncb2963
- Allen BL, Taatjes DJ. 2015. The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol* **16**: 155–166. doi:10.1038/nrm3951
- Ardehali R, Inlay MA, Ali SR, Tang C, Drukker M, Weissman IL. 2011. Overexpression of BCL2 enhances survival of human embryonic stem cells during stress and obviates the requirement for serum factors. *Proc Natl Acad Sci* **108**: 3282–3287. doi:10.1073/pnas.1019047108
- Aubert J, Dunstan H, Chambers I, Smith A. 2002. Functional gene screening in embryonic stem cells implicates Wnt antagonism in neural differentiation. *Nat Biotechnol* **20**: 1240–1245. doi:10.1038/nbt763
- Baas R, van Teeffelen HAAM, Tjalsma SJD, Timmers HTM. 2018. The mixed lineage leukemia 4 (MLL4) methyltransferase complex is involved in transforming growth factor β (TGF- β)-activated gene transcription. *Transcription* **9**: 67–74. doi:10.1080/21541264.2017.1373890
- Barlev NA, Liu L, Chehab NH, Mansfield K, Harris KG, Halazonetis TD, Berger SL. 2001. Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. *Mol Cell* **8**: 1243–1254. doi:10.1016/S1097-2765(01)00414-2
- Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, Weinberg RA. 2008. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* **40**: 499–507. doi:10.1038/ng.127
- Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, Ziller M, Croft GF, Amoroso MW, Oakley DH, et al. 2011. Reference maps of human ES and iPSC cell variation enable high-

- throughput characterization of pluripotent cell lines. *Cell* **144**: 439–452. doi:10.1016/j.cell.2010.12.032
- Boddicker RL, Kip NS, Xing X, Zeng Y, Yang ZZ, Lee JH, Almada LL, Elsawa SF, Knudson RA, Law ME, et al. 2015. The oncogenic transcription factor IRF4 is regulated by a novel CD30/NF- κ B positive feedback loop in peripheral T-cell lymphoma. *Blood* **125**: 3118–3127. doi:10.1182/blood-2014-05-578575
- Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, et al. 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**: 349–353. doi:10.1038/nature04733
- Carver EA, Jiang R, Lan Y, Oram KF, Gridley T. 2001. The mouse snail gene encodes a key regulator of the epithelial-mesenchymal transition. *Mol Cell Biol* **21**: 8184–8188. doi:10.1128/MCB.21.23.8184-8188.2001
- Chalamalasetty RB, Garriock RJ, Dunty WC, Kennedy MW, Jailwala P, Si H, Yamaguchi TP. 2014. Mesogenin 1 is a master regulator of paraxial presomitic mesoderm differentiation. *Development* **141**: 4285–4297. doi:10.1242/dev.110908
- Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**: 643–655. doi:10.1016/S0092-8674(03)00392-1
- Chen T, Dent SYR. 2014. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat Rev Genet* **15**: 93–106. doi:10.1038/nrg3607
- Chia N-Y, Chan Y-S, Feng B, Lu X, Orlov YL, Moreau D, Kumar P, Yang L, Jiang J, Lau M-S, et al. 2010. A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* **468**: 316–320. doi:10.1038/nature09531
- Collinson A, Collier AJ, Morgan NP, Sienerth AR, Chandra T, Andrews S, Rugg-Gunn PJ. 2016. Deletion of the Polycomb-group protein EZH2 leads to compromised self-renewal and differentiation defects in human embryonic stem cells. *Cell Rep* **17**: 2700–2714. doi:10.1016/j.celrep.2016.11.032
- Conroy AT, Sharma M, Holtz AE, Wu C, Sun Z, Weigel RJ. 2002. A novel zinc finger transcription factor with two isoforms that are differentially repressed by estrogen receptor- α . *J Biol Chem* **277**: 9326–9334. doi:10.1074/jbc.M107702200
- Dalton S. 2015. Linking the cell cycle to cell fate decisions. *Trends Cell Biol* **25**: 592–600. doi:10.1016/j.tcb.2015.07.007
- D'Amour KA, Agulnick AD, Eliazar S, Kelly OG, Kroon E, Baetge EE. 2005. Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat Biotechnol* **23**: 1534–1541. doi:10.1038/nbt1163
- Davoli T, Xu AW, Mengwasser KE, Sack LM, Yoon JC, Park PJ, Elledge SJ. 2013. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* **155**: 948–962. doi:10.1016/j.cell.2013.10.011
- Efroni S, Duttagupta R, Cheng J, Dehghani H, Hoepfner DJ, Dash C, Bazett-Jones DP, Le Grice S, McKay RDG, Buetow KH, et al. 2008. Global transcription in pluripotent embryonic stem cells. *Cell Stem Cell* **2**: 437–447. doi:10.1016/j.stem.2008.03.021
- Evseenko D, Zhu Y, Schenke-Layland K, Kuo J, Latour B, Ge S, Scholes J, Dravid G, Li X, MacLellan WR, et al. 2010. Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. *Proc Natl Acad Sci* **107**: 13742–13747. doi:10.1073/pnas.1002077107
- Felfly H, Klein OD. 2013. Sprouty genes regulate proliferation and survival of human embryonic stem cells. *Sci Rep* **3**: 2277. doi:10.1038/srep02277
- Gao Z, Zhang J, Bonasio R, Strino F, Sawai A, Parisi F, Kluger Y, Reinberg D. 2012. PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol Cell* **45**: 344–356. doi:10.1016/j.molcel.2012.01.002
- Gaspar-Maia A, Alajem A, Meshorer E, Ramalho-Santos M. 2011. Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol* **12**: 36–47. doi:10.1038/nrm3036
- Gibson P, Tong Y, Robinson G, Thompson MC, Currie DS, Eden C, Kranenburg TA, Hogg T, Poppleton H, Martin J, et al. 2010. Subtypes of medulloblastoma have distinct developmental origins. *Nature* **468**: 1095–1099. doi:10.1038/nature09587
- Gifford CA, Ziller MJ, Gu H, Trapnell C, Donaghey J, Tsankov A, Shalek AK, Kelley DR, Shishkin AA, Issner R, et al. 2013. Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* **153**: 1149–1163. doi:10.1016/j.cell.2013.04.037
- Gonzales KAU, Liang H, Lim Y-SS, Chan Y-SS, Yeo J-CC, Tan C-PP, Gao B, Le B, Tan Z-YY, Low K-YY, et al. 2015. Deterministic restriction on pluripotent state dissolution by cell-cycle pathways. *Cell* **162**: 564–579. doi:10.1016/j.cell.2015.07.001
- Gudas LJ, Wagner JA. 2011. Retinoids regulate stem cell differentiation. *J Cell Physiol* **226**: 322–330. doi:10.1002/jcp.22417
- Hardwick LJA, Philpott A. 2014. Nervous decision-making: to divide or differentiate. *Trends Genet* **30**: 254–261. doi:10.1016/j.tig.2014.04.001
- Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, et al. 2011. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* **29**: 731–734. doi:10.1038/nbt.1927
- Hu G, Kim J, Xu Q, Leng Y, Orkin SH, Elledge SJ. 2009. A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev* **23**: 837–848. doi:10.1101/gad.1769609
- Ihry RJ, Worringer KA, Salick MR, Frias E, Ho D, Theriault K, Kommineni S, Chen J, Sondey M, Ye C, et al. 2018. P53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nat Med* **24**: 939–946. doi:10.1038/s41591-018-0050-6
- Ihry RJ, Salick MR, Ho DJ, Sondey M, Kommineni S, Paula S, Raymond J, Henry B, Frias E, Wang Q, et al. 2019. Genome-scale CRISPR screens identify human pluripotency-specific genes. *Cell Rep* **27**: 616–630.e6. doi:10.1016/j.celrep.2019.03.043
- Jain M, Arvanitis C, Chu K, Dewey W, Leonhardt E, Trinh M, Sundberg CD, Bishop JM, Felsher DW. 2002. Sustained loss of a neoplastic phenotype by brief inactivation of MYC. *Science* **297**: 102–104. doi:10.1126/science.1071489
- Jiang YW, Veschambre P, Erdjument-Bromage H, Tempst P, Conaway JW, Conaway RC, Kornberg RD. 1998. Mammalian mediator of transcriptional regulation and its possible role as an end-point of signal transduction pathways. *Proc Natl Acad Sci* **95**: 8538–8543. doi:10.1073/pnas.95.15.8538
- Klopocki E, Kähler C, Foulds N, Shah H, Joseph B, Vogel H, Lüttgen S, Bald R, Besoke R, Held K, et al. 2012. Deletions in PITX1 cause a spectrum of lower-limb malformations including mirror-image polydactyly. *Eur J Hum Genet* **20**: 705–708. doi:10.1038/ejhg.2011.264
- Koutelou E, Hirsch CL, Dent SY. 2010. Multiple faces of the SAGA complex. *Curr Opin Cell Biol* **22**: 374–382. doi:10.1016/j.ceb.2010.03.005
- Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, et al. 2006. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**: 301–313. doi:10.1016/j.cell.2006.02.043
- Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, Irizarry RA, Liu JS, Brown M, Liu XS. 2014. MAGECK enables robust

- identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* **15**: 554. doi:10.1186/s13059-014-0554-4
- Li QV, Dixon G, Verma N, Rosen BP, Gordillo M, Luo R, Xu C, Wang Q, Soh CL, Yang D, et al. 2019. Genome-scale screens identify JNK–JUN signaling as a barrier for pluripotency exit and endoderm differentiation. *Nat Genet* **51**: 999–1010. doi:10.1038/s41588-019-0408-9
- Liu JC, Guan X, Ryan JA, Rivera AG, Mock C, Agarwal V, Letai A, Lerou PH, Lahav G. 2013. High mitochondrial priming sensitizes hESCs to DNA-damage-induced apoptosis. *Cell Stem Cell* **13**: 483–491. doi:10.1016/j.stem.2013.07.018
- Liu L, Michowski W, Kolodziejczyk A, Sicinski P. 2019. The cell cycle in stem cell proliferation, pluripotency and differentiation. *Nat Cell Biol* **21**: 1060–1067. doi:10.1038/s41556-019-0384-4
- Loubiere V, Delest A, Thomas A, Bonev B, Schuettengruber B, Sati S, Martinez A-M, Cavalli G. 2016. Coordinate redeployment of PRC1 proteins suppresses tumor formation during *Drosophila* development. *Nat Genet* **48**: 1436–1442. doi:10.1038/ng.3671
- Martin TD, Cook DR, Choi MY, Li MZ, Haigis KM, Elledge SJ. 2017. A role for mitochondrial translation in promotion of viability in K-Ras mutant cells. *Cell Rep* **20**: 427–438. doi:10.1016/j.celrep.2017.06.061
- Martinez A-M, Schuettengruber B, Sakr S, Janic A, Gonzalez C, Cavalli G. 2009. Polyhomeotic has a tumor suppressor activity mediated by repression of Notch signaling. *Nat Genet* **41**: 1076–1082. doi:10.1038/ng.414
- Meerbrey KL, Hu G, Kessler JD, Roarty K, Li MZ, Fang JE, Herschowitz JI, Burrows AE, Ciccio A, Sun T, et al. 2011. The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc Natl Acad Sci* **108**: 3665–3670. doi:10.1073/pnas.1019736108
- Mendenhall EM, Eckfeldt CE, Ekker SC, Verfaillie CM. 2004. The role of Sprouty family members in hematopoiesis in zebrafish and mammals. *Blood* **104**: 137. doi:10.1182/blood.V104.11.137.137
- Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, Dharia NV, Montgomery PG, Cowley GS, Pantel S, et al. 2017. Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. *Nat Genet* **49**: 1779–1784. doi:10.1038/ng.3984
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**: 553–560. doi:10.1038/nature06008
- Morey L, Pascual G, Cozzuto L, Roma G, Wutz A, Benitah SA, Di Croce L. 2012. Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. *Cell Stem Cell* **10**: 47–62. doi:10.1016/j.stem.2011.12.006
- Nagai T, Aruga J, Minowa O, Sugimoto T, Ohno Y, Noda T, Mikoshiba K. 2000. Zic2 regulates the kinetics of neurulation. *Proc Natl Acad Sci* **97**: 1618–1623. doi:10.1073/pnas.97.4.1618
- Nam S, Lim JS. 2016. Essential role of interferon regulatory factor 4 (IRF4) in immune cell development. *Arch Pharm Res* **39**: 1548–1555. doi:10.1007/s12272-016-0854-1
- Naxerova K, Bult CJ, Peaston A, Fancher K, Knowles BB, Kasif S, Kohane IS. 2008. Analysis of gene expression in a developmental context emphasizes distinct biological leitmotifs in human cancers. *Genome Biol* **9**: R108. doi:10.1186/gb-2008-9-7-r108
- Ochiai K, Maienschein-Cline M, Simonetti G, Chen J, Rosenthal R, Brink R, Chong AS, Klein U, Dinner AR, Singh H, et al. 2013. Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity* **38**: 918–929. doi:10.1016/j.immuni.2013.04.009
- Pauklin S, Vallier L. 2013. The cell-cycle state of stem cells determines cell fate propensity. *Cell* **155**: 135–147. doi:10.1016/j.cell.2013.08.031
- Pauklin S, Vallier L. 2015. Activin/nodal signalling in stem cells. *Development* **142**: 607–619. doi:10.1242/dev.091769
- Pauklin S, Madrigal P, Bertero A, Vallier L. 2016. Initiation of stem cell differentiation involves cell cycle-dependent regulation of developmental genes by Cyclin D. *Genes Dev* **30**: 421–433. doi:10.1101/gad.271452.115
- Petropoulos S, Edsgård D, Reinius B, Deng Q, Panula SP, Code-luppi S, Plaza Reyes A, Linnarsson S, Sandberg R, Lanner F. 2016. Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell* **165**: 1012–1026. doi:10.1016/j.cell.2016.03.023
- Phippard D, Lu L, Lee D, Saunders JC, Crenshaw EB. 1999. Targeted mutagenesis of the POU-domain gene *Brn4/Pou3f4* causes developmental defects in the inner ear. *J Neurosci* **19**: 5980–5989. doi:10.1523/JNEUROSCI.19-14-05980.1999
- Pritsker M, Ford NR, Jenq HT, Lemischka IR. 2006. Genomewide gain-of-function genetic screen identifies functionally active genes in mouse embryonic stem cells. *Proc Natl Acad Sci* **103**: 6946–6951. doi:10.1073/pnas.0509861103
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**: e47. doi:10.1093/nar/gkv007
- Robinson MD, McCarthy DJ, Smyth GK. 2010. Edger: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140. doi:10.1093/bioinformatics/btp616
- Sack LM, Davoli T, Li MZ, Li Y, Xu Q, Naxerova K, Wooten EC, Bernardi RJ, Martin TD, Chen T, et al. 2018. Profound tissue specificity in proliferation control underlies cancer drivers and aneuploidy patterns. *Cell* **173**: 499–514.e23. doi:10.1016/j.cell.2018.02.037
- Shaffer AL, Emre NCT, Lamy L, Ngo VN, Wright G, Xiao W, Powell J, Dave S, Yu X, Zhao H, et al. 2008. IRF4 addiction in multiple myeloma. *Nature* **454**: 226–231. doi:10.1038/nature07064
- Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG, et al. 2014. Genome-scale CRISPR–Cas9 knockout screening in human cells. *Science* **343**: 84–87. doi:10.1126/science.1247005
- Sharma R, Sanchez-Ferraz O, Bouchard M. 2015. Pax genes in renal development, disease and regeneration. *Semin Cell Dev Biol* **44**: 97–106. doi:10.1016/j.semdb.2015.09.016
- Stern CD. 2006. Neural induction: 10 years on since the ‘default model.’ *Curr Opin Cell Biol* **18**: 692–697. doi:10.1016/j.ceb.2006.09.002
- Sundberg M, Jansson L, Ketolainen J, Pihlajamäki H, Suuronen R, Skottman H, Inzunza J, Hovatta O, Narkilahti S. 2009. CD marker expression profiles of human embryonic stem cells and their neural derivatives, determined using flow-cytometric analysis, reveal a novel CD marker for exclusion of pluripotent stem cells. *Stem Cell Res* **2**: 113–124. doi:10.1016/j.scr.2008.08.001
- Tajonar A, Maehr R, Hu G, Sneddon JB, Rivera-Feliciano J, Cohen DE, Elledge SJ, Melton DA. 2013. Brief report: VGLL4 is a novel regulator of survival in human embryonic stem cells. *Stem Cells* **31**: 2833–2841. doi:10.1002/stem.1445

- Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**: 663–676. doi:10.1016/j.cell.2006.07.024
- Trounson A. 2017. Potential pitfall of pluripotent stem cells. *N Engl J Med* **377**: 490–491. doi:10.1056/NEJMcibr1706906
- Tsankov AM, Gu H, Akopian V, Ziller MJ, Donaghey J, Amit I, Gnirke A, Meissner A. 2015. Transcription factor binding dynamics during human ES cell differentiation. *Nature* **518**: 344–349. doi:10.1038/nature14233
- Vallier L, Alexander M, Pedersen RA. 2005. Activin/nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci* **118**: 4495–4509. doi:10.1242/jcs.02553
- van Neerven S, Kampmann E, Mey J. 2008. RAR/RXR and PPAR/RXR signaling in neurological and psychiatric diseases. *Prog Neurobiol* **85**: 433–451. doi:10.1016/j.pneurobio.2008.04.006
- Vilhais-Neto GC, Fournier M, Plassat J-L, Sardu ME, Saraf A, Garnier J-M, Maruhashi M, Florens L, Washburn MP, Pourquié O. 2017. The WHHERE coactivator complex is required for retinoic acid-dependent regulation of embryonic symmetry. *Nat Commun* **8**: 728. doi:10.1038/s41467-017-00593-6
- Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. 2013. Cancer genome landscapes. *Science* **339**: 1546–1558. doi:10.1126/science.1235122
- Wang L, YR Dent S. 2014. Functions of SAGA in development and disease. *Epigenomics* **6**: 329–339. doi:10.2217/epi.14.22
- Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, Sabatini DM. 2015. Identification and characterization of essential genes in the human genome. *Science* **350**: 1096–1101. doi:10.1126/science.aac7041
- Wang Z, Gearhart MD, Lee YW, Kumar I, Ramazanov B, Zhang Y, Hernandez C, Lu AY, Neuenkirchen N, Deng J, et al. 2018. A non-canonical BCOR–PRC1.1 complex represses differentiation programs in human ESCs. *Cell Stem Cell* **22**: 235–251.e9. doi:10.1016/j.stem.2017.12.002
- Watanabe K, Ueno M, Kamiya D, Nishiyama A, Matsumura M, Wataya T, Takahashi JB, Nishikawa S, Nishikawa S, Muguruma K, et al. 2007. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat Biotechnol* **25**: 681–686. doi:10.1038/nbt1310
- Weilemann A, Grau M, Erdmann T, Merkel O, Sobhifshar U, Anagnostopoulos I, Hummel M, Siegert A, Hayford C, Madle H, et al. 2015. Essential role of IRF4 and MYC signaling for survival of anaplastic large cell lymphoma. *Blood* **125**: 124–132. doi:10.1182/blood-2014-08-594507
- Weinstein IB. 2002. Cancer: enhanced: addiction to oncogenes—the Achilles heel of cancer. *Science* **297**: 63–64. doi:10.1126/science.1073096
- Westendorp B, Mokry M, Groot Koerkamp MJA, Holstege FCP, Cuppen E, de Bruin A. 2012. E2f7 represses a network of oscillating cell cycle genes to control S-phase progression. *Nucleic Acids Res* **40**: 3511–3523. doi:10.1093/nar/gkr1203
- White J, Dalton S. 2005. Cell cycle control of embryonic stem cells. *Stem Cell Rev* **1**: 131–138. doi:10.1385/SCR:1:2:131
- Wilson KD, Sun N, Huang M, Zhang WY, Lee AS, Li Z, Wang SX, Wu JC. 2010. Effects of ionizing radiation on self-renewal and pluripotency of human embryonic stem cells. *Cancer Res* **70**: 5539–5548. doi:10.1158/0008-5472.CAN-09-4238
- Wu D-D, Irwin DM, Zhang Y-P. 2008. Molecular evolution of the keratin associated protein gene family in mammals, role in the evolution of mammalian hair. *BMC Evol Biol* **8**: 241. doi:10.1186/1471-2148-8-241
- Wu X, Dao Thi VL, Huang Y, Billerbeck E, Saha D, Hoffmann HH, Wang Y, Silva LAV, Sarbanes S, Sun T, et al. 2018. Intrinsic immunity shapes viral resistance of stem cells. *Cell* **172**: 423–438.e25. doi:10.1016/j.cell.2017.11.018
- Xie W, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, Whitaker JW, Tian S, Hawkins RD, Leung D, et al. 2013. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* **153**: 1134–1148. doi:10.1016/j.cell.2013.04.022
- Yamaguchi TP, Takada S, Yoshikawa Y, Wu N, McMahon AP. 1999. T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes Dev* **13**: 3185–3190. doi:10.1101/gad.13.24.3185
- Yamamoto M, Meno C, Sakai Y, Shiratori H, Mochida K, Ikawa Y, Saijoh Y, Hamada H. 2001. The transcription factor FoxH1 (FAST) mediates nodal signaling during anterior–posterior patterning and node formation in the mouse. *Genes Dev* **15**: 1242–1256. doi:10.1101/gad.883901
- Yilmaz A, Peretz M, Aharony A, Sagi I, Benvenisty N. 2018. Defining essential genes for human pluripotent stem cells by CRISPR–Cas9 screening in haploid cells. *Nat Cell Biol* **20**: 610–619. doi:10.1038/s41556-018-0088-1
- Yilmaz A, Braverman-Gross C, Bialer-Tsypin A, Peretz M, Benvenisty N. 2020. Mapping gene circuits essential for germ layer differentiation via loss-of-function screens in haploid human embryonic stem cells. *Cell Stem Cell* **27**: 679–691.e6. doi:10.1016/j.stem.2020.06.023
- Youle RJ, Strasser A. 2008. The BCL-2 protein family: opposing activities that mediate cell death. *Nat Rev Mol Cell Biol* **9**: 47–59. doi:10.1038/nrm2308
- Young RA. 2011. Control of the embryonic stem cell state. *Cell* **144**: 940–954. doi:10.1016/j.cell.2011.01.032
- Yu J, Thomson JA. 2008. Pluripotent stem cell lines. *Genes Dev* **22**: 1987–1997. doi:10.1101/gad.1689808
- Zambrowicz BP, Friedrich GA, Buxton EC, Lilleberg SL, Person C, Sands AT. 1998. Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. *Nature* **392**: 608–611. doi:10.1038/33423
- Zhang P, Wong C, Dou L, Finegold M, Harper JW, Elledge SJ. 1999. P21(CIP1) and p57(KIP2) control muscle differentiation at the myogenin step. *Genes Dev* **13**: 213–224. doi:10.1101/gad.13.2.213
- Zhang Y, Schulz VP, Reed BD, Wang Z, Pan X, Mariani J, Euskirchen G, Snyder MP, Vaccarino FM, Ivanova N, et al. 2013. Functional genomic screen of human stem cell differentiation reveals pathways involved in neurodevelopment and neurodegeneration. *Proc Natl Acad Sci* **110**: 12361–12366. doi:10.1073/pnas.1309725110