Stem Cell Reports
Article
OPEN ACCESS

GCN5 Regulates FGF Signaling and Activates Selective MYC Target Genes during Early Embryoid Body Differentiation

Li Wang,[1,2,3,5] Evangelia Koutelou,[1,2,5] Calley Hirsch,[1,2] Ryan McCarthy,[1,2,6] Andria Schibler,[1,2,4,7] Kevin Lin,[1,2] Yue Lu,[1,2] Collene Jeter,[1] Jianjun Shen,[1] Michelle C. Barton,[1,2,3,4] and Sharon Y.R. Dent[1,2,3,4,*]
[1]Department of Epigenetics and Molecular Carcinogenesis
[2]Center for Cancer Epigenetics
The University of Texas MD Anderson Cancer Center, Smithville, TX 78957, USA
[3]Program in Epigenetics and Molecular Carcinogenesis
[4]Program in Genes and Development
The Graduate School of Biomedical Sciences (GSBS) of the University of Texas Health Science Center at Houston, Houston, TX 77030, USA
[5]Co-first author
[6]Present address: Department of Cell and Developmental Biology, University of Pennsylvania, Philadelphia, PA 19104, USA
[7]Present address: National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA
*Correspondence: sroth@mdanderson.org
https://doi.org/10.1016/j.stemcr.2017.11.009

SUMMARY

Precise control of gene expression during development is orchestrated by transcription factors and co-regulators including chromatin modifiers. How particular chromatin-modifying enzymes affect specific developmental processes is not well defined. Here, we report that GCN5, a histone acetyltransferase essential for embryonic development, is required for proper expression of multiple genes encoding components of the fibroblast growth factor (FGF) signaling pathway in early embryoid bodies (EBs). *Gcn5*$^{-/-}$ EBs display deficient activation of ERK and p38, mislocalization of cytoskeletal components, and compromised capacity to differentiate toward mesodermal lineage. Genomic analyses identified seven genes as putative direct targets of GCN5 during early differentiation, four of which are cMYC targets. These findings established a link between GCN5 and the FGF signaling pathway and highlighted specific GCN5-MYC partnerships in gene regulation during early differentiation.

INTRODUCTION

Embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of blastocysts, which are early-stage preimplantation embryos. ESCs have the ability to remain pluripotent and self-renew or to differentiate into multiple cell lineages. ESC identity and subsequent differentiation are controlled by intricate networks of transcription factors and signaling pathways that drive precise gene expression programs. Diverse chromatin regulators play important roles in these networks (Chen and Dent, 2014; Lessard and Crabtree, 2010), but the roles of specific histone modifying enzymes in ESC self-renewal or lineage specification are poorly understood.

GCN5 was the first histone lysine acetyltransferase (HAT) to be linked to active gene transcription (Brownell et al., 1996). GCN5 functions within multimember protein complexes, including SAGA and ATAC in multicellular organisms, to coactivate transcription (Baker and Grant, 2007; Koutelou et al., 2010; Timmers and Tora, 2005). In yeast, Gcn5-containing complexes are recruited to target genes via interactions with specific DNA-binding factors, but only a few such partners, such as cMYC and nMYC, have been defined in mammalian cells (Hirsch et al., 2015; Martinez-Cerdeno et al., 2012; Zhang et al., 2008a). SAGA has also been suggested to act as a general transcription factor in yeast, widely enhancing expression of active genes (Baptista et al., 2017; Bonnet et al., 2014).

Genetic studies in mice revealed that both GCN5 and its catalytic activity are essential for normal development and embryo survival. *Gcn5*$^{-/-}$ embryos die soon after gastrulation and exhibit increased apoptosis in mesodermal lineages (Xu et al., 2000). *Gcn5* catalytic mutant mice survive until mid-gestation but develop cranial neural tube closure defects (Bu et al., 2007) due to abnormal retinoic acid signaling involving a nonhistone substrate of GCN5 (Wilde et al., 2017). These findings indicate that GCN5-containing complexes have both HAT-dependent and -independent functions in early development. The phenotypes of *Gcn5* mutant mice also support a selective role for this HAT in gene regulation, as loss of general transcription factors often leads to early lethality prior to embryo implantation (Tudor et al., 1999).

Our previous studies defined GCN5 as an important coactivator for MYC and E2F family transcription factors in the regulation of cell-cycle genes involved in ESC self-renewal (Hirsch et al., 2015) and pointed to the involvement of GCN5 in early ESC differentiation (Lin et al., 2007). However, early developmental processes and associated signaling pathways modulated by GCN5 have not yet been defined.

Fibroblast growth factor (FGF) signaling is required for multiple stages of early embryonic development, from segregation of trophectoderm and primitive endoderm from the ICM (Chazaud et al., 2006; Georgiades and Rossant, 2006; Kang et al., 2017; Yamanaka et al., 2010) to specification of primitive ectoderm and the primitive streak (Ciruna and Rossant, 2001), which in turn determine the fate of epiblast. ESC-based studies indicate that FGF signaling is involved in both pluripotency maintenance and lineage specification *in vitro* (Kunath et al., 2007; Lanner and Rossant, 2010; Ying et al., 2008). However, how epigenetic factors contribute to FGF-mediated gene regulation during early development is unclear.

Here, we use ESCs bearing a floxed allele of *Gcn5* to define GCN5 functions in embryoid body (EB) differentiation. Morphological and molecular analyses of $Gcn5^{fx/fx}$ and $Gcn5^{-/-}$ EBs reveal an important role for GCN5 in the regulation of FGF signaling during early differentiation of EBs and confirm the importance of GCN5 for proper expression of select MYC target genes.

RESULTS

Gcn5 Loss Leads to Epiblast Disorganization *In Vitro*

The early lethality of $Gcn5^{-/-}$ embryos poses challenges for detailed molecular studies, so we developed ESCs that carry a floxed allele of *Gcn5* (Hirsch et al., 2015) to define pathways regulated by GCN5 during ESC differentiation *in vitro*. Before Cre-mediated recombination, the *Gcn5* floxed allele behaves as wild-type, and mice homozygous for this allele ($Gcn5^{fx/fx}$) show no overt phenotypes (Lin et al., 2008). After Cre exposure, exons 3–18 of *Gcn5* are deleted, creating a null allele ($Gcn5^{-/-}$).

ESCs readily aggregate when cultured in suspension without inhibitors of differentiation (leukemia inhibitory factor [LIF] and 2i) and undergo stepwise morphological changes to form distinct three-dimensional structures. Hallmark events of this process include visceral endoderm differentiation (VE) and basement membrane (BM) assembly (days 2–3), followed by polarized epiblast formation from the inner cells and the clearance of a central cavity (days 4–6) (Li et al., 2003). These sequential events recapitulate transitions from formation of the ICM through embryonic gastrulation, thereby providing opportunities to define molecular events *in vitro* that might contribute to the death of $Gcn5^{-/-}$ embryos shortly after gastrulation *in vivo* (Xu et al., 2000).

We initially observed that $Gcn5^{-/-}$ EBs were much smaller than $Gcn5^{fx/fx}$ EBs at day 9, when all three germ layers should be formed (Figure 1A). However, at day 3, $Gcn5^{-/-}$ EBs were not obviously different from controls. By day 5, $Gcn5^{-/-}$ EBs began to suffer more breakage than the control EBs (Figure 1A), suggesting that null EBs were structurally fragile or that they were in the process of dying. However, cleaved Lamin A levels were not increased in day 5 $Gcn5^{-/-}$ EBs (Figure S1A), indicating that *Gcn5* loss did not induce apoptosis. Immunoblots of H3S10p (Figure S1A), a marker for mitotic cells, also indicated that deletion of *Gcn5* did not inhibit cell proliferation. Global levels of H3K9ac were comparable in $Gcn5^{fx/fx}$ and $Gcn5^{-/-}$ EBs at day 5 (Figure S1B), consistent with the presence of other HATs (e.g., PCAF) in the EBs. Together these data indicate that GCN5 is required for normal morphology, but not proliferation or survival, in early stages of EB differentiation.

To investigate whether the abnormal morphology of $Gcn5^{-/-}$ EBs reflects defective formation of VE or BM at day 5, we performed immunostaining of GATA4 (for VE) and Laminin (for BM). Since a majority of the inner cells express early neuroectodermal markers by day 5, we also used immunostaining of SOX1 to visualize the organization of epiblast cells. Laminin staining of the BM showed little difference between the null and control EBs (Figure 1B). GATA4 staining was also very similar in $Gcn5^{fx/fx}$ and $Gcn5^{-/-}$ EBs, indicating normal formation and organization of the VE (Figure 1C). In contrast, SOX1 staining revealed severe disorganization of the columnar epithelial morphology of the epiblast in $Gcn5^{-/-}$ EBs, indicating that GCN5 is required for normal epiblast formation and organization.

Compromised Differentiation Potential of *Gcn5* Null EBs

The epiblast gives rise to all three germ layers (Rivera-Perez and Hadjantonakis, 2014), so we asked whether *Gcn5* loss affects lineage differentiation. We examined marker genes that normally exhibit peak expression in the embryonic epiblast, including *Fgf5* and *Otx2* (Kamiya et al., 2011; Kurokawa et al., 2004; Sumi et al., 2013; Yamanaka et al., 2010), and confirmed lower expression levels in the null epiblasts at day 5 (Figure 2A). We then used mass cytometry to more fully assess the composition of cell populations within the control and null EBs (Bendall and Nolan, 2012) (Figure S2 and Table S1). Day 5 EBs were dissociated and stained with antibodies specific for markers for either pluripotent cells or germ layer progenitors (Table S1). The antibodies were labeled with nonradioactive isotopes of rare earth metals with distinct atomic masses, which are distinguished by mass cytometry and serve as reporters for the labeled cells. This approach allowed us to delineate multiple lineages simultaneously and compare the compositions of heterogeneous populations in day 5 EBs. SPADE (spanning-tree progression analysis of density-normalized events) analysis was used to visualize and categorize populations based on combinatorial marker expression (Table S1) and to quantify the mass cytometry data.

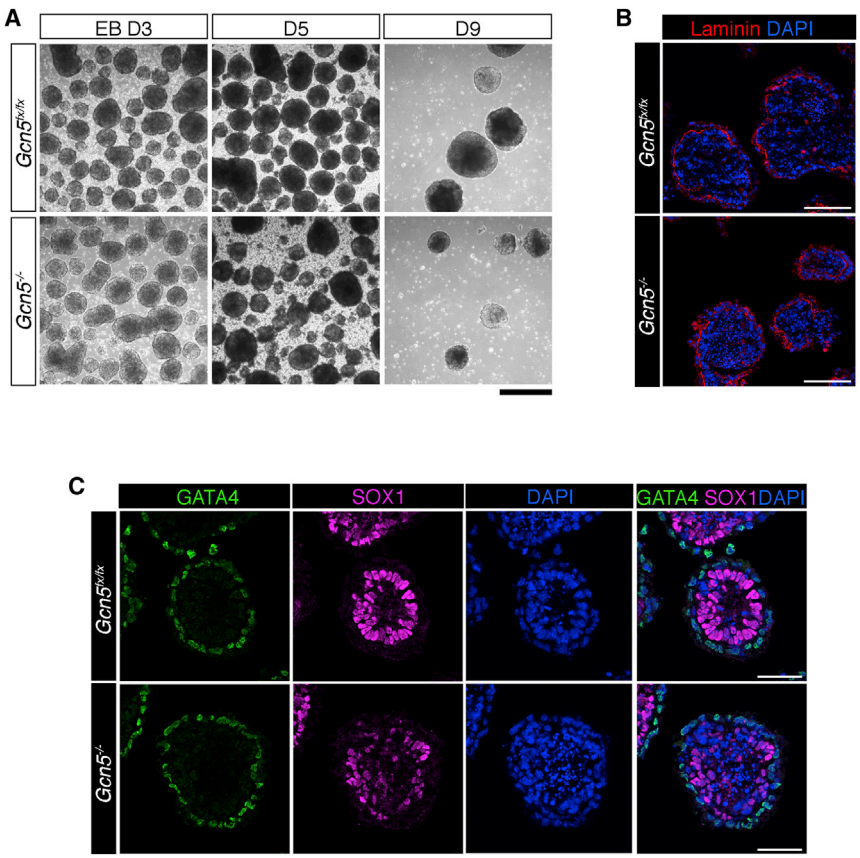


Figure 1. Loss of *Gcn5* Leads to Disorganization of Epiblast Cells during Early Differentiation

(A) Bright-field images of differentiating EBs at days 3, 5, and 9, comparing EB morphology between $Gcn5^{fx/fx}$ and $Gcn5^{-/-}$ EBs. Scale bar, 200 μM.

(B and C) Representative immunofluorescence/confocal images of EB architecture in day 5 $Gcn5^{fx/fx}$ and $Gcn5^{-/-}$ EBs, with (B) Laminin (red) staining for BM and (C) GATA4 (green) staining for VE, and SOX1 (magenta) and DAPI (blue) staining for nuclei of the epiblast. Scale bars, 100 μM in (B) and 50 μM in (C).

See also Figure S1.

The results indicated that $Gcn5^{-/-}$ EBs contained significantly lower numbers of endodermal and mesodermal cells relative to $Gcn5^{fx/fx}$ EBs, whereas comparable numbers of cells were observed for pluripotent populations (Figures 2B, S3A, and S3B). Decreased mesoderm populations in the $Gcn5^{-/-}$ EBs were confirmed across four replicate experiments (Figures S3A and S3B). The ectoderm population in the null EBs was unchanged in cell number (Figure 2B, right), although the ectoderm region in the SPADE tree showed changes in some nodes (Figure 2B, left), likely due to specificity limitations of the antibodies used to define ectoderm (Figure S2). Decreased expression of all germ layer markers was observed in late-stage $Gcn5^{-/-}$ EBs (days 9 and 12) (Figure 2C), likely stemming from the earlier defects observed at day 5.

To further confirm the effects of *Gcn5* loss on mesoderm and endoderm formation, we utilized a monolayer differentiation protocol to direct ESCs toward these lineages (Villegas et al., 2013). Again we observed significantly decreased expression of mesoderm-specific genes (*T, Flk1,* and *Pdgfrb*), but in contrast to the results in EBs, expression of endoderm-specific genes was unaffected (*Gata4*) or upregulated (*Sox17* and *FoxA2*) in $Gcn5^{-/-}$ cells differentiated in monolayer (Figure S3C). Altogether these results indicate that GCN5 is most important for mesoderm formation during ESC differentiation, reminiscent of our previous findings in $Gcn5^{-/-}$ embryos (Xu et al., 2000).

Expression Profiling Reveals a Regulatory Role for GCN5 in FGF Signaling

GCN5 acts as a transcriptional coactivator in the context of the SAGA and ATAC complexes (Koutelou et al., 2010; Spedale et al., 2012; Suganuma et al., 2008). To better understand the molecular basis underlying the defects caused by *Gcn5* loss, we performed RNA sequencing (RNA-seq) to compare gene expression profiles in day 3 and day 5 control and null EBs. Total RNA from three technical replicates of $Gcn5^{fx/fx}$ and $Gcn5^{-/-}$ EBs were sequenced, and key gene expression changes were confirmed using a second biological replicate (EBs generated from a separate matched pair of $Gcn5^{fx/fx}$ and $Gcn5^{-/-}$ ESCs) by quantitative real-time polymerase chain reaction (PCR). These time points were chosen to define both early events (day 3) and events that coincide with the onset of the abnormal phenotype of $Gcn5^{-/-}$ EBs (day 5).

Principal component analysis revealed significant differences in gene expression profiles between day 3 and day 5 EBs, consistent with developmental progression over time.

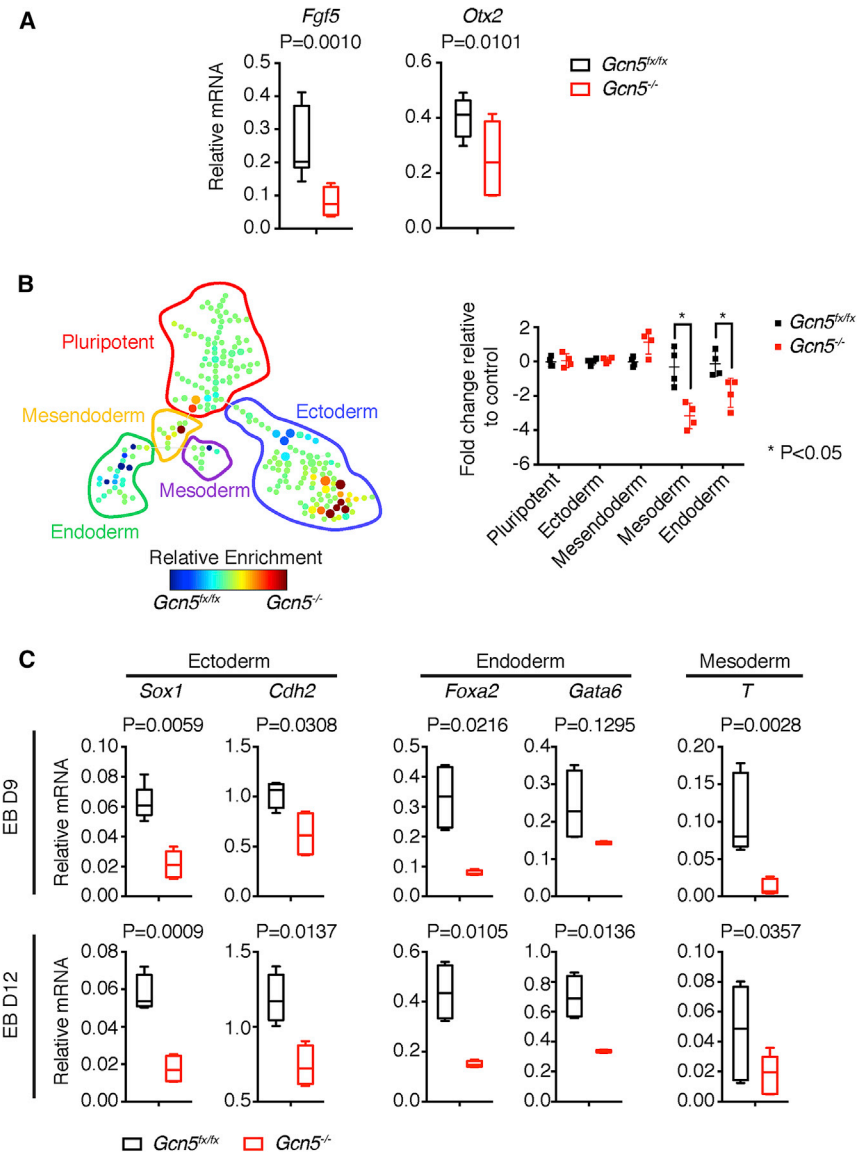


Figure 2. Compromised Differentiation Potentials of *Gcn5* Null EBs

(A) Gene expression analysis by quantitative real-time PCR for epiblast marker genes.

(B) Changes in population composition in *Gcn5* null EBs at epiblast stage. Left: SPADE tree plot showing decreased endoderm and mesoderm populations in the *Gcn5* null EBs at day 5; Right: quantitation of fold changes in cell numbers of a given population in *Gcn5* null EBs relative to control.

(C) Quantitative real-time PCR plots showing decreased expression of marker genes for ectoderm (*Sox1, Cdh2*), endoderm (*Foxa2, Gata6*), and mesoderm (*T*) in late-stage EBs (days 9 and 12).

Data are presented as means ± SD from three (A and C) or four (B) independent experiments.

See also Figures S2 and S3 and Table S1.

Of note, differences in the gene expression profiles between null and control EBs were less pronounced at day 3 than at day 5 (Figure 3A), with 754 genes exhibiting altered expression in day 5 null EBs, whereas 158 genes were affected in day 3 nulls (Figure S4A). These data indicate that *Gcn5* loss more significantly affects gene expression programs at day 5, consistent with the timing of the onset of *Gcn5* null morphological phenotype. Therefore, we focused further detailed analyses on differences in gene expression observed at day 5.

Gene set enrichment analysis revealed significant downregulation of several biological processes in the null EBs at day 5, with multicellular organismal development (MOD) and cell surface receptor linked signal transduction (CSRLST) among the most affected (Figure 3B). In the MOD category, 15 of 50 genes were among the core enrichment group in the null EBs, including *Pax5, Msx2, Gli1, Spry2*, and *Mest*, which are all important for early development or ESC differentiation (Lee et al., 2016; Szabo et al., 2009; Tefft et al., 1999; Urbanek et al., 1997; Wu et al., 2015) (Figure S4B and Table S2). In the CSRLST category, 7 of 30 genes were in the core enrichment group, including *Grb10*, the most downregulated gene in the null EBs (Figure S4C and Table S3). Consistent with these results, Ingenuity Pathway Analysis (IPA) identified a number of signaling pathways to be significantly altered in the null EBs. Interestingly, four of the seven top-ranked affected pathways were intimately linked to FGF signaling, including regulation of epithelial-to-mesenchymal transition, STAT3, FGF, and growth hormone signaling pathways (Figure 3C).

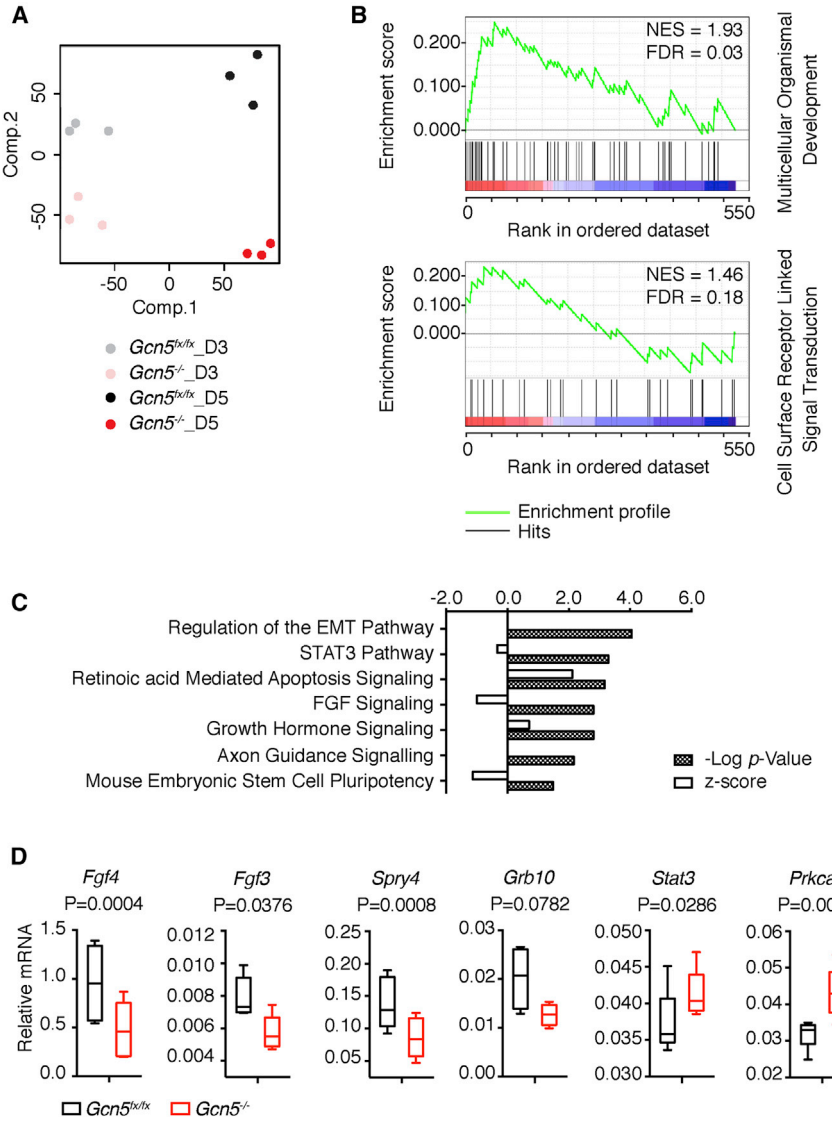


Figure 3. Expression Profiling Points to a Regulatory Role of *Gcn5* in FGF Signaling in Day 5 EBs

(A) Principal component analysis plots showing variance of expression profiles among replicative samples and increased differences in profiles between $Gcn5^{fx/fx}$ and $Gcn5^{-/-}$ EBs at day 5 compared with day 3 ($n = 3$).

(B) Example gene set enrichment analysis enrichment plots showing top processes enriched in the control EBs compared with $Gcn5^{-/-}$ EBs at day 5. Color code: red, positively correlated with $Gcn5^{fx/fx}$; blue, negatively correlated with $Gcn5^{-/-}$.

(C) Significantly altered pathways in the day 5 *Gcn5* null EBs identified by IPA canonical pathway analysis.

(D) Quantitative real-time PCR validating the key genes in the FGF signaling identified by the RNA-seq using a second biological sample. Data are presented as means ± SD from three independent experiments.

FDR, false discovery rate; NES, normalized enrichment score. See also Figure S4, Tables S2 and S3.

FGF ligands and their receptor tyrosine kinases control multiple developmental processes, including cell proliferation, survival, differentiation, and migration (Brewer et al., 2016). We confirmed that expression of key genes in the FGF pathway, including *Fgf3, Fgf4,* and *Spry4*, was altered in the $Gcn5^{-/-}$ EBs using quantitative real-time PCR. We also confirmed that *Grb10*, an effector of insulin signaling (Desbuquois et al., 2013; Yu et al., 2011), was strongly downregulated. Conversely, effector genes such as *Stat3* and *Prkca* were upregulated (Figure 3D).

Connections between Defective FGF Signaling and Morphological Abnormalities of $Gcn5^{-/-}$ EBs at Day 5

To link the above changes in gene expression to the status of FGF signaling intermediates in $Gcn5^{-/-}$ EBs, we examined levels of FGFR1 and phosphorylation of ERK, p38, c-RAF, and AKT by immunoblotting (Figures 4A and 4B). FGFR1 protein levels and activated, phosphorylated forms of both ERK and p38 (P-ERK and P-p38, respectively) were notably decreased in $Gcn5^{-/-}$ EBs, indicating deficient activation of the RAS/MAPK pathway. In contrast, levels of activated AKT (P-AKT) and P-c-RAF-Ser259 were unchanged or slightly increased. Levels of phospholipase C-γ did not change (data not shown), even though *Prkca* mRNA was upregulated in the null EBs.

To relate these molecular changes back to the morphological phenotypes of the $Gcn5^{-/-}$ EBs, we examined the organization of cytoskeletal components regulated by ERK and p38 signaling (Huang et al., 2004) using Airyscan (Zeiss) laser-scanning confocal microscopy. We first assessed filamentous actin (F-actin) in day 5 EBs using phalloidin staining. In epiblasts of EBs or early embryonic

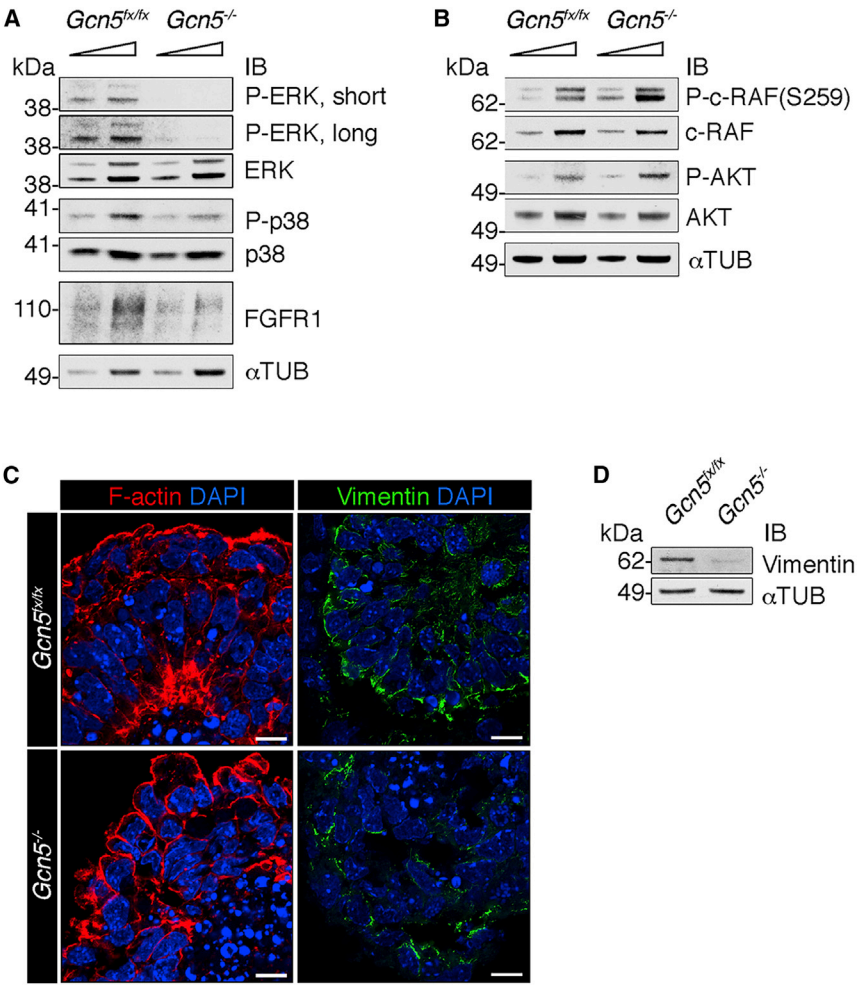


Figure 4. Abnormal Activities of the FGF Signaling Pathway in Day 5 *Gcn5*$^{-/-}$ EBs

(A) Representative immunoblots showing decreased phosphorylated forms of ERK and p38 in the *Gcn5* null EBs at day 5.

(B) Representative immunoblots showing the AKT pathway is affected to a lesser degree in the null EBs at day 5.

(C) Zeiss LSM 880 confocal Airyscan images showing disorganization of F-actin (red) and vimentin (green) in day 5 null EBs. Scale bars, 10 μM.

(D) Representative immunoblots showing decreased vimentin in *Gcn5*$^{-/-}$ EBs at day 5.

epithelial structures, F-actin localizes to the periphery of epithelial cells and is particularly enriched at apical sites of columnar epithelial cells (Loebel et al., 2011; Sakai et al., 2003). This staining pattern was observed as expected in *Gcn5*$^{fx/fx}$ EBs. In contrast, *Gcn5*$^{-/-}$ EBs displayed reduced apical distribution of F-actin, primarily in the inner cells (epiblast) (Figure 4C). Airyscan images of vimentin staining also indicated altered localization and decreased staining intensity of this intermediate filament protein in *Gcn5*$^{-/-}$ EBs (Figure 4C). Depletion of vimentin was further confirmed by immunoblotting (Figure 4D). These findings suggest that insufficient activation of ERK and p38 upon *Gcn5* loss affects F-actin and vimentin organization during early differentiation, in concordance with disorganized epithelial architectures of the inner cell layer observed in the null EBs (Figure 1C). Defective ERK signaling is also consistent with decreased mesoderm differentiation (Binetruy et al., 2007) in *Gcn5*$^{-/-}$ EBs (Figure 2) and embryos (Xu et al., 2000).

Decreased Gene Expression and Reduced H3K9ac at Gene Promoters Identified Likely Direct Targets of GCN5

To identify genes directly regulated by GCN5 in EBs at the epiblast stage, we attempted chromatin immunoprecipitation sequencing (ChIP-seq), but we were not successful using either commercially available antibodies for GCN5 or a biotin-tagging system (Hirsch et al., 2015) in differentiating ESCs. Since histone H3 lysine 9 (H3K9) is a well-characterized substrate for GCN5 (Jin et al., 2011; Kuo et al., 1996), we performed H3K9ac ChIP, using the same control and *Gcn5*$^{-/-}$ EBs used for RNA-seq. We reasoned that localized decreases in H3K9ac in the *Gcn5*$^{-/-}$ EBs might identify genes that uniquely require the presence of this HAT for proper regulation. As global levels of H3K9ac were not affected by *Gcn5* loss (Figure S1B), regions identified by this approach will likely provide an underestimate of GCN5 targets due to redundancies with PCAF and possibly other HATs.

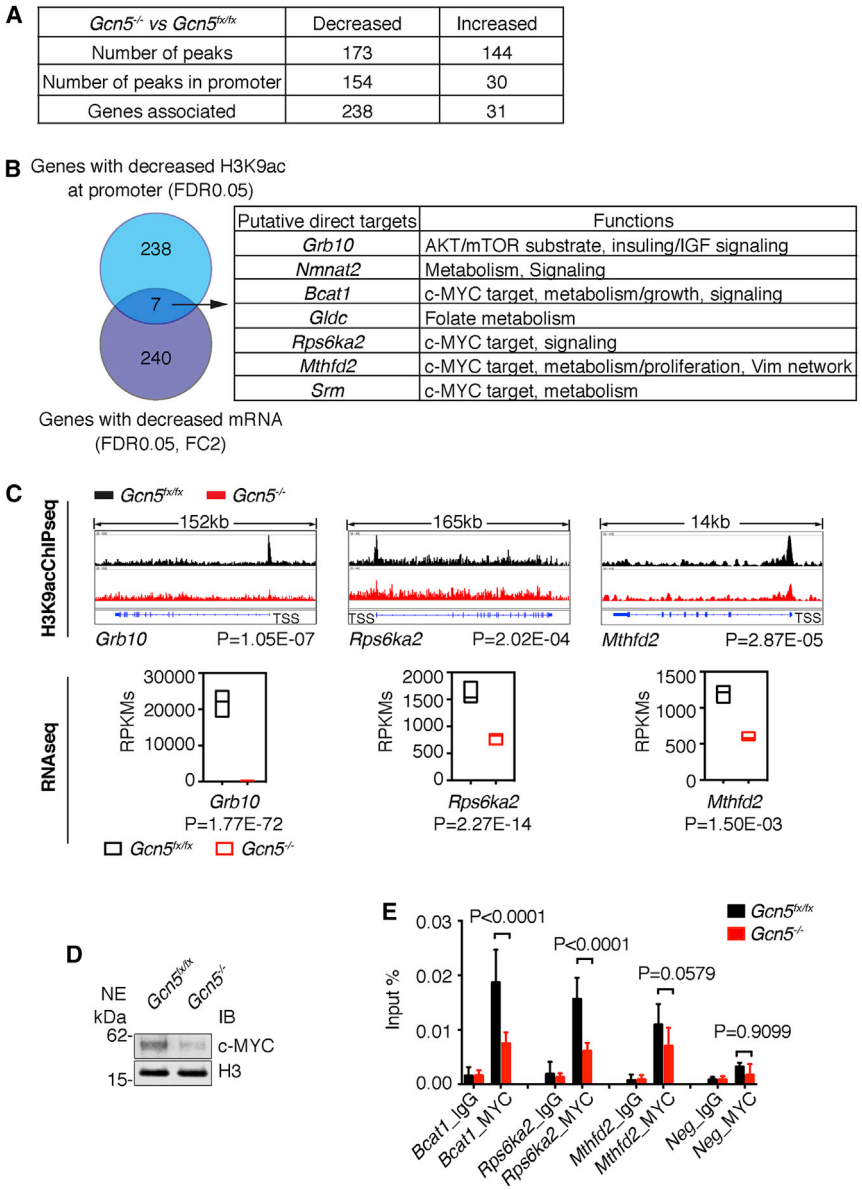


Figure 5. Identification of Direct GCN5 Target Genes at the Early Stage of Differentiation

(A) Comparing profiles of H3K9ac peaks and associated genes in $Gcn5^{fx/fx}$ and $Gcn5^{-/-}$ EBs at day 5. Loss of *Gcn5* resulted in decreased H3K9ac peaks mostly in the promoter region. Most of the genes (238 of 269) with altered H3K9ac peaks exhibited a decreased level of this mark.

(B) Putative target genes directly regulated by GCN5. Venn diagram showing overlap between H3K9ac decreased genes and GCN5 induced genes. Overlapping genes are listed in the table with a general description of the reported functions.

(C) Examples of H3K9ac peak profiles (n = 4) and RNA transcripts (n = 3) for top targets of GCN5 in day 5 EBs. RNA transcript levels are presented as means ± SD.

(D) Representative immunoblots showing reduced levels of cMYC isolated from nuclei of day 5 $Gcn5^{-/-}$ EBs. NE, nuclear extracts.

(E) ChIPs for MYC reveal decreased MYC binding at the promoters of GCN5 target genes in day 5 $Gcn5^{-/-}$EBs. Enrichment of MYC relative to input at each locus is presented as the mean ± SD from three independent experiments.

See also Table S4.

We identified 173 H3K9ac peaks that were decreased in the $Gcn5^{-/-}$ EBs, the majority (154) of which were located near gene promoters. Gene annotation identified 238 genes likely driven by these promoters, with occasions where H3K9ac peaks fell in putative promoter regions associated with more than one gene. We also uncovered 144 H3K9ac peaks that increased upon *Gcn5* loss, yet only 30 of those peaks were near the promoters, associated with 31 genes. These data are consistent with the gene-specific coactivator role of GCN5 in gene transcription (Figure 5A). Comparison of genes identified in our analyses as having decreased H3K9ac promoter peaks with ENCODE ChIP-seq data (Auerbach et al., 2013) identified candidate transcription factors that might recruit GCN5 to these regions (Table S4). Top TFs identified by this approach include HCFC1 (Q value 1.25×10^{-77}), a nuclear protein known to associate with GCN5-containing complexes (Wang et al., 2008), as well as TBP (Q value 1.69×10^{-59}) and CTCF (Q value 3.77×10^{-59}). Strikingly, a number of MYC family members, including MAX, MXI1, and cMYC, were also identified by this approach, consistent with our previous work connecting GCN5 to MYC functions in both ESCs and during somatic cell reprogramming.

To better determine which of these regions might reflect genes directly activated by GCN5, we compared genes with decreased H3K9ac at their promoters with genes identified

as downregulated more than 2-fold by RNA-seq. Only seven genes were both downregulated and decreased in H3K9ac. Three of these genes, *Grb10*, *Gldc*, and *Nmnat*, further reinforce the link among GCN5, metabolism, and signaling (Figures 5B and 5C). The other four genes (*Rps6ka2*, *Mthfd2*, *Bcat1*, and *Srm*) are directly regulated by cMYC (Ben-Yosef et al., 1998; Pikman et al., 2016; Snezhkina et al., 2016) (Figure 5B). *Myc* is induced by most mitogenic factors, including FGFs (Grandori et al., 2000). We observed decreased cMYC protein levels upon loss of *Gcn5* (Figure 5D), consistent with a previous report in a different cellular system (Patel et al., 2004). ChIP-qPCR confirmed MYC recruitment to the promoter regions of *Rps6ka2*, *Mthfd2*, and *Bcat1* in $Gcn5^{fx/fx}$ EBs and loss of MYC from these regions in $Gcn5^{-/-}$ EBs at day 5 (Figures 5C and 5E). These findings further highlight the GCN5-MYC partnership in gene regulation during early differentiation of EBs.

DISCUSSION

Our findings reveal that loss of *Gcn5* strongly affects FGF signaling at multiple levels during early differentiation of EBs, with decreased expression of *Fgf*s and FGFR1 and deficient activation of ERK and p38. Moreover, *Gcn5* loss leads to direct downregulation of specific genes involved in signaling and metabolism as well as discrete MYC gene targets (Figure 6).

FGF signaling regulates both migration and patterning of mesoderm during gastrulation (Ciruna and Rossant, 2001). Failures in the execution of this pathway are consistent with the abnormal cytoskeletal organization and defective mesoderm formation observed in $Gcn5^{-/-}$ EBs. In addition, identification of *Bcat1*, *Rps6ka2*, *Mthfd2*, and *Srm* as likely direct targets of GCN5, which exhibit decreased expression upon *Gcn5* loss along with reduced MYC binding and diminished H3K9ac at their promoters, provides further evidence that GCN5 is an important coactivator for MYC during early differentiation. These genes are known to be immediate-response, MYC target genes induced by FGF signaling.

Another notable common function of the direct targets of GCN5 identified here is that they regulate metabolism and growth downstream of AKT and/or ERK signaling, either positively (*Nmnat2*, *Bcat1*, *Gldc*, *Mthfd2*, and *Srm*) or negatively (*Grb10*, *Bcat1*, and *Rps6ka2*) (Figure 5B) (Gerdts et al., 2016; Pai et al., 2015; Pikman et al., 2016; Serra et al., 2013; She et al., 2007; Shi et al., 2012; Yu et al., 2011). Altered carbon metabolism induced by these gene expression changes could further derail differentiation (Garcia-Prat et al., 2017; Hu et al., 2016).

In addition to histones, GCN5 acetylates nonhistone proteins that may indirectly regulate gene transcription or function outside of gene regulation (Conacci-Sorrell et al., 2010; Jin et al., 2014; Wilde et al., 2017). For example, GCN5 pairs with MYC-nick, a CALPAIN-cleaved

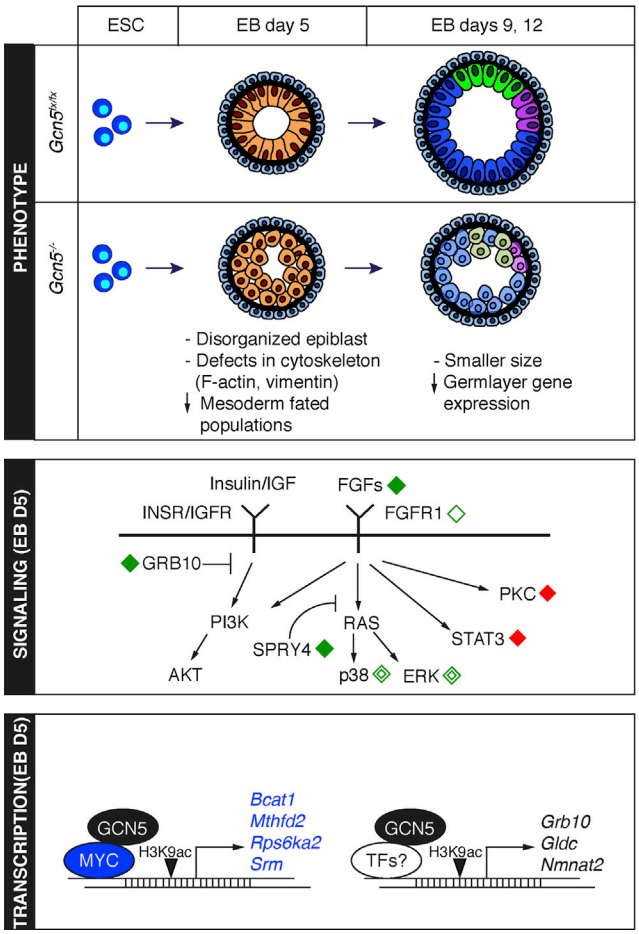


Figure 6. GCN5 Affects Multiple Components of the FGF Signaling Pathway and Activates Selective Targets during Early Differentiation

Top: the abnormal phenotype of $Gcn5^{-/-}$ EBs becomes evident early during EB differentiation (day 5). Loss of *Gcn5* leads to disorganization of the epiblast architecture that is associated with defective cytoskeleton networks (F-actin and vimentin) and decreases in progenitors fated for mesoderm. At later stages of differentiation (days 9 and 12), $Gcn5^{-/-}$ EBs are smaller in size and express lower levels of marker genes for ectoderm (blue), endoderm (green), and mesoderm (magenta), compared with the controls. Lighter shades indicate decreased expression levels of marker genes for each population. Middle: at day 5, GCN5 affects expression of multiple genes encoding critical components in FGF signaling and for proper activation of ERK and p38 pathways. Solid diamond, up (red) or down (green) regulated transcripts; open diamond (green), decreased protein expression; double open diamond (green), decreased protein phosphorylation. Bottom: at day 5, GCN5 is required for activation of genes important for signaling through promoter-associated H3K9ac, including four cMYC targets (blue).

cytoplasmic derivative of MYC, to acetylate alpha-tubulin (ac-αTUB) and to regulate cytoskeleton organization and differentiation of myoblasts (Conacci-Sorrell et al., 2010). We found that total ac-αTUB levels in *Gcn5* null EBs at day 5 were unchanged (data not shown), but we cannot exclude contributions of changes in acetylation of other targets to the developmental phenotypes we observe.

Our previous genetic studies showed that $Gcn5^{-/-};Pcaf^{-/-}$ embryos die earlier than $Gcn5^{-/-}$ embryos, even though *Pcaf* deletion on its own causes no abnormal phenotype (Xu et al., 2000; Yamauchi et al., 2000). These findings indicate that *Gcn5* and *Pcaf* have shared functions during early development. The lack of global changes in H3K9ac and limited changes in H3K9ac at gene promoters observed in day 5 $Gcn5^{-/-}$ EBs is consistent with functional redundancy with PCAF and possibly other HATs. Nonetheless, compensatory H3K9ac by PCAF did not prevent defective morphology and signaling in the early $Gcn5^{-/-}$ EBs, suggesting that GCN5 and H3K9ac are uniquely required for regulation of specific genes during early differentiation.

GCN5 is most active *in vivo* when incorporated into SAGA and ATAC in mammals (Koutelou et al., 2010; Wang et al., 2008). Our studies do not differentiate the functions of the two complexes, and the phenotypes we observe may reflect loss of activity of both. However, knockout of *Atac2,* a component of ATAC but not SAGA, did not cause defects in lineage differentiation (Suganuma et al., 2008), as was observed in $Gcn5^{-/-}$ embryos, suggesting SAGA may be most important for these early developmental events.

Interestingly, more genes were upregulated upon loss of *Gcn5* (at both days 3 and 5) than were downregulated, in contrast to the role of GCN5 as a coactivator of transcription. Many of these events are likely indirect effects of *Gcn5* loss, although we cannot exclude the possibility that GCN5 may be involved in gene repression during early differentiation. Indeed, a recent study revealed that another HAT, TIP60, acts as a transcriptional repressor in ESCs (Fazzio et al., 2008). Notably, many of the genes upregulated upon *Gcn5* loss are involved in acute phase response and interferon signaling (data not shown), consistent with our previous work in fibroblasts that indicated GCN5 and PCAF repress interferon-β expression by targeting a nonhistone substrate, TBK1 (Jin et al., 2014).

Abnormal regulation of growth-factor-driven pathways drive oncogenesis (Giancotti, 2014). Our findings here suggest that GCN5 may also be important in cancers associated with deregulation of FGFs. Future work will explore this possibility, as well as the therapeutic potential of targeting GCN5 to inhibit growth or progression of these cancers.

EXPERIMENTAL PROCEDURES

ES Cell Culture and Differentiation

$Gcn5^{fx/fx}$ and $Gcn5^{-/-}$ ESC lines were generated and characterized previously (Hirsch et al., 2015). ESCs were routinely grown on gelatin-coated plates in Dulbecco's modified Eagle's medium (DMEM)/high glucose (HyClone, SH3002201) medium supplemented with 15% (v/v) ESC-screened fetal bovine serum (HyClone, SH3007003E), 2 mM L-glutamine (HyClone, SH3003401), 0.1 mM nonessential amino acids (Corning, MT25025CI), 1% (v/v) penicillin/streptomycin (HyClone, SV30010), 0.1 mM β-mercaptoethanol (BME) (Fisher, 03446I-100), 1000 U/mL LIF/ESGRO (Millipore, ESG1107), 1 μM PD0325901 (Sigma, PZ0162) and 3 μM 1-azakenpaullone (Sigma, A3734), and passaged every 2–3 days.

For EB differentiation, 3×10^5 cells/well were plated in ultra-low attachment six-well plates in differentiation medium without LIF or 2i-s. Media were replaced every other day by settling the EBs at low speed centrifugation (100 × *g* for 1 min). The differentiation medium was DMEM/high glucose:F12 (Cellgro, MT10080CV):neurobasal medium (Gibco, 21103049) (1:1:2) supplemented with 10% KnockOut Serum Replacement medium (Gibco, A3181502), 2 mM L-glutamine, 1% (v/v) penicillin/streptomycin, and 0.1 mM BME.

Immunofluorescence and Confocal Imaging

EBs were washed once in PBS/1% BSA and fixed in 4% paraformaldehyde (PFA) for 30 min at room temperature. The fixed EBs were incubated in 7.5% sucrose/PBS for at least 1 hr at room temperature, then in 15% sucrose/PBS at 4°C overnight. The EBs were then embedded in Tissue-Tek OCT compound (Electron Microscopy Sciences, 62550-12) and incubated for 10 min at room temperature with agitation before they were frozen in liquid nitrogen (LN2). Frozen sections (8 μM) were fixed in 2% PFA for 2 min then blocked with PBS containing 0.1% Tween 20 (Fisher, BP337-500) (PBT) and 5% normal donkey serum (Millipore, S30-100ML) for 30 min at room temperature. The blocked sections were incubated with primary antibodies diluted in blocking buffer overnight at 4°C. The slides were washed with PBT three times for a total 15 min, followed by incubations with fluorescence-conjugated secondary antibodies for 40 min at room temperature. DAPI staining was performed after washing off the secondary antibodies. The slides then were washed with PBT and mounted with coverslips using ProLong Gold Antifade mounting medium (Invitrogen, P36930). The slides were imaged on a Zeiss LSM 880 laser-scanning microscope. Airyscan detector array was used to image the cytoskeletons. Standard pinhole was used to image the markers for lineages, proliferation, and apoptosis. The antibodies used are listed in the Supplemental Experimental Procedures.

Expression Analysis

EBs were harvested and total RNA was isolated using an RNeasy mini plus kit (QIAGEN, 74134) following the manufacturer's instructions. Total RNA (10–20 ng) was used per reaction, and quantitative real-time PCRs were performed on a 7500 Fast Real-Time PCR System (Applied Biosystems, 4351107) using the Power SYBR Green RNA-to-CT1-Step Kit (Life Technologies, 4389986). Three technical replicates were performed for each gene target

tested and *Pbgd* was used as a reference gene for quantification. A two-tailed Student's t test was used for pairwise comparisons. Primers used are listed in Supplemental Experimental Procedures.

RNA Sequencing

Total RNA libraries were prepared from three independent experiments using the Illumina TruSeq stranded total RNA kit according to the manufacturer's protocol, except that the PCR amplification was reduced to eight cycles. Each library (10 pM) was sequenced on an Illumina HiSeq 2500. The reads were mapped to the mouse genome (mm10) by TopHat (version 2.0.10) (Kim et al., 2013) with an overall mapping rate of 84%–94%. DESeq (Anders and Huber, 2010) was used for differential gene expression analysis. Details of RNA-seq analysis are described in Supplemental Experimental Procedures.

ChIP, qPCR, and Deep Sequencing

Day 5 EBs (*Gcn5*$^{fx/fx}$ and *Gcn5*$^{-/-}$) were washed once with PBS and dissociated using Accutase cell dissociation reagent (Gibco, A1110501). ChIPs were performed as previously described (Wen et al., 2014) with modifications in chromatin sonication. Details are included in Supplemental Experimental Procedures.

qPCRs were performed on a 7500 Fast Real-Time PCR System using ChIP DNA from three replicative experiments and Power SYBR Green PCR Master Mix (Life Technologies, 4367659). Student's t test was used for pairwise comparisons. Primers used are listed in Supplemental Experimental Procedures.

ChIP DNA from four independent experiments, including two replicates for each isolate pair (*Gcn5*$^{fx/fx}$ and *Gcn5*$^{-/-}$), were used for deep sequencing. Detailed ChIP library preparations are described in Supplemental Experimental Procedures. Each library (10 pM) was sequenced on an Illumina HiSeq 3000. Raw reads were aligned to mouse genome mm10 using bowtie (version 1.1.2) (Langmead et al., 2009). H3K9ac peaks were called by MACS (version 1.4.2) (Zhang et al., 2008b) using H3 as control. Detailed ChIP-seq data analysis is included in Supplemental Experimental Procedures.

Mass Cytometry and Data Analysis

Sample preparation for mass cytometry was performed as previously described (McCarthy et al., 2017a). Briefly, EBs were dissociated using Accutase, stained with cisplatin for viability at 25 μM for 1 min at room temperature, then quenched using PBS containing 1% BSA. The cells were fixed in 1.5% PFA for 10 min at room temperature, then permeabilized with methanol (1 mL per million cells) and incubated overnight at 4°C. Samples were barcoded (McCarthy et al., 2017b), pooled, and immunostained with the panel of antibodies shown in Supplemental Experimental Procedures. Cells were stained with 1:2000$^{191/193}$ iridium (Ir) DNA intercalator (Fluidigm), 62.5 nM final, for 10 min at room temperature. The samples were combined with EQ Four Element Calibration Beads (Fluidigm) then diluted with water to a concentration of 5×10^5 cells/mL and run at 45 μL/min on a CyTOF 2 mass cytometer (Fluidigm). Data were normalized on bead passport using CyTOF software (v6.0.626; Fluidigm).

Initial data processing and gating was done with FlowJo vX10.0. EQ Four Element Calibration Beads were removed, and data were gated on singlets by Ir193 and Event Length parameters. Removal of dead cells was done in the Pt198 channel. SPADE analysis of the data was performed using SPADE V3.0 (Qiu et al., 2011) in MATLAB r2015b (Mathworks). SPADE tree construction was performed using agglomerative clustering on all markers listed in Table S1. Annotation of SPADE tree regions was done according to marker distribution as shown in Figure S2, and cell percentages in each region were calculated for all samples. Percentages were normalized relative to mean *Gcn5*$^{fx/fx}$, and statistical significance was determined by the Wilcoxon rank-sum test performed in MATLAB using the ranksum function.

Immunoblotting

Whole-cell lysates (WCL) were prepared from D5 EBs using RIPA buffer. Nuclear extracts (NEs) were prepared following the Dignam-Roeder protocol (Dignam et al., 1983). All buffers contained protease inhibitor cocktail (Sigma, P8340) and phosphatase inhibitor cocktail (Roche, 04906845001). WCL or NE (10–20 μg) was resolved on 4%–12% Bis-Tris protein gels (WG1402BOX). Proteins were transferred to nitrocellulose membranes (Invitrogen, IB23001) using an iBlot2 (Life Technologies, IB21001), then blocked and incubated with primary antibodies following standard procedures. Primary antibodies were detected with peroxidase-conjugated secondary antibodies (1:8,000) and Amersham ECL prime western blotting detection reagent (GE Healthcare, RPN2232) following the manufacturer's instructions. The antibodies used are listed in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the RNA-seq and H3K9ac ChIP-seq data reported in this paper is GEO: GSE104454.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and four tables and can be found with this article online at https://doi.org/10.1016/j.stemcr.2017.11.009.

AUTHOR CONTRIBUTIONS

L.W., E.K., and S.Y.R.D. designed the study, analyzed the data, and wrote the paper. L.W. and E.K. planned all the molecular, cellular, and genomic studies, which were carried out by L.W.; C.H. provided the ESCs and performed some ChIP experiments; R.M. assisted with the mass cytometry experiments and performed data analysis; A.S. assisted with immunofluorescence/confocal experiments; K.L. and Y.L. performed bioinformatics analysis; C.J. and J.S. provided technical assistance for imaging and sequencing; M.C.B. supervised the cytometry experiments and provided expert advice; S.Y.R.D. supervised the overall research.

ACKNOWLEDGMENTS

This work was largely supported by NIH R01 grant GM067718 to S.Y.R.D. The Science Park NGS Core performed the deep sequencing supported by CPRIT Core Facility Support Grants RP120348 and RP170002. Bioinformatics was made possible in

part by support from the MDACC Center for Cancer Epigenetics (CCE). Immunofluorescence/confocal analysis was performed at the Science Park IF/imaging core, partially supported by CCE. We thank Andrew Salinger (A.S.) and Amanda Martin for critical material support in the lab, Hsueh-Ping Chao for help in bioinformatics, and Yoko Takata and Luis Coletta for help in NGS. We appreciate A.S. and Dr. Lisa Mustachio for helpful comments on the manuscript.

Received: July 14, 2017
Revised: November 14, 2017
Accepted: November 14, 2017
Published: December 14, 2017

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Auerbach, R.K., Chen, B., and Butte, A.J. (2013). Relating genes to function: identifying enriched transcription factors using the ENCODE ChIP-Seq significance tool. Bioinformatics *29*, 1922–1924.

Baker, S.P., and Grant, P.A. (2007). The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. Oncogene *26*, 5329–5340.

Baptista, T., Grunberg, S., Minoungou, N., Koster, M.J.E., Timmers, H.T.M., Hahn, S., Devys, D., and Tora, L. (2017). SAGA is a general cofactor for RNA polymerase II transcription. Mol. Cell *68*, 130–143.e5.

Ben-Yosef, T., Eden, A., and Benvenisty, N. (1998). Characterization of murine BCAT genes: Bcat1, a c-Myc target, and its homolog, Bcat2. Mamm. Genome *9*, 595–597.

Bendall, S.C., and Nolan, G.P. (2012). From single cells to deep phenotypes in cancer. Nat. Biotechnol. *30*, 639–647.

Binetruy, B., Heasley, L., Bost, F., Caron, L., and Aouadi, M. (2007). Concise review: regulation of embryonic stem cell lineage commitment by mitogen-activated protein kinases. Stem Cells *25*, 1090–1095.

Bonnet, J., Wang, C.Y., Baptista, T., Vincent, S.D., Hsiao, W.C., Stierle, M., Kao, C.F., Tora, L., and Devys, D. (2014). The SAGA coactivator complex acts on the whole transcribed genome and is required for RNA polymerase II transcription. Genes Dev. *28*, 1999–2012.

Brewer, J.R., Mazot, P., and Soriano, P. (2016). Genetic insights into the mechanisms of Fgf signaling. Genes Dev. *30*, 751–771.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell *84*, 843–851.

Bu, P., Evrard, Y.A., Lozano, G., and Dent, S.Y. (2007). Loss of Gcn5 acetyltransferase activity leads to neural tube closure defects and exencephaly in mouse embryos. Mol. Cell. Biol. *27*, 3405–3416.

Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev. Cell *10*, 615–624.

Chen, T., and Dent, S.Y. (2014). Chromatin modifiers and remodellers: regulators of cellular differentiation. Nat. Rev. Genet. *15*, 93–106.

Ciruna, B., and Rossant, J. (2001). FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. Dev. Cell *1*, 37–49.

Conacci-Sorrell, M., Ngouenet, C., and Eisenman, R.N. (2010). Myc-nick: a cytoplasmic cleavage product of Myc that promotes alpha-tubulin acetylation and cell differentiation. Cell *142*, 480–493.

Desbuquois, B., Carre, N., and Burnol, A.F. (2013). Regulation of insulin and type 1 insulin-like growth factor signaling and action by the Grb10/14 and SH2B1/B2 adaptor proteins. FEBS J. *280*, 794–816.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. *11*, 1475–1489.

Fazzio, T.G., Huff, J.T., and Panning, B. (2008). An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. Cell *134*, 162–174.

Garcia-Prat, L., Sousa-Victor, P., and Munoz-Canoves, P. (2017). Proteostatic and metabolic control of stemness. Cell Stem Cell *20*, 593–608.

Georgiades, P., and Rossant, J. (2006). Ets2 is necessary in trophoblast for normal embryonic anteroposterior axis development. Development *133*, 1059–1068.

Gerdts, J., Summers, D.W., Milbrandt, J., and DiAntonio, A. (2016). Axon self-destruction: new links among SARM1, MAPKs, and NAD+ metabolism. Neuron *89*, 449–460.

Giancotti, F.G. (2014). Deregulation of cell signaling in cancer. FEBS Lett. *588*, 2558–2570.

Grandori, C., Cowley, S.M., James, L.P., and Eisenman, R.N. (2000). The Myc/Max/Mad network and the transcriptional control of cell behavior. Annu. Rev. Cell Dev. Biol. *16*, 653–699.

Hirsch, C.L., Coban Akdemir, Z., Wang, L., Jayakumaran, G., Trcka, D., Weiss, A., Hernandez, J.J., Pan, Q., Han, H., Xu, X., et al. (2015). Myc and SAGA rewire an alternative splicing network during early somatic cell reprogramming. Genes Dev. *29*, 803–816.

Hu, C., Fan, L., Cen, P., Chen, E., Jiang, Z., and Li, L. (2016). Energy metabolism plays a critical role in stem cell maintenance and differentiation. Int. J. Mol. Sci. *17*, 253.

Huang, C., Jacobson, K., and Schaller, M.D. (2004). MAP kinases and cell migration. J. Cell Sci. *117*, 4619–4628.

Jin, Q., Yu, L.R., Wang, L., Zhang, Z., Kasper, L.H., Lee, J.E., Wang, C., Brindle, P.K., Dent, S.Y., and Ge, K. (2011). Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. EMBO J. *30*, 249–262.

Jin, Q., Zhuang, L., Lai, B., Wang, C., Li, W., Dolan, B., Lu, Y., Wang, Z., Zhao, K., Peng, W., et al. (2014). Gcn5 and PCAF negatively regulate interferon-beta production through HAT-independent inhibition of TBK1. EMBO Rep. *15*, 1192–1201.

Kamiya, D., Banno, S., Sasai, N., Ohgushi, M., Inomata, H., Watanabe, K., Kawada, M., Yakura, R., Kiyonari, H., Nakao, K.,

et al. (2011). Intrinsic transition of embryonic stem-cell differentiation into neural progenitors. Nature *470*, 503–509.

Kang, M., Garg, V., and Hadjantonakis, A.K. (2017). Lineage establishment and progression within the inner cell mass of the mouse blastocyst requires FGFR1 and FGFR2. Dev. Cell *41*, 496–510.e5.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Koutelou, E., Hirsch, C.L., and Dent, S.Y. (2010). Multiple faces of the SAGA complex. Curr. Opin. Cell Biol. *22*, 374–382.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development *134*, 2895–2902.

Kuo, M.H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. Nature *383*, 269–272.

Kurokawa, D., Takasaki, N., Kiyonari, H., Nakayama, R., Kimura-Yoshida, C., Matsuo, I., and Aizawa, S. (2004). Regulation of Otx2 expression and its functions in mouse epiblast and anterior neuroectoderm. Development *131*, 3307–3317.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lanner, F., and Rossant, J. (2010). The role of FGF/Erk signaling in pluripotent cells. Development *137*, 3351–3360.

Lee, H.J., Choi, N.Y., Lee, S.W., Ko, K., Hwang, T.S., Han, D.W., Lim, J., Scholer, H.R., and Ko, K. (2016). Epigenetic alteration of imprinted genes during neural differentiation of germline-derived pluripotent stem cells. Epigenetics *11*, 177–183.

Lessard, J.A., and Crabtree, G.R. (2010). Chromatin regulatory mechanisms in pluripotency. Annu. Rev. Cell Dev. Biol. *26*, 503–532.

Li, S., Edgar, D., Fassler, R., Wadsworth, W., and Yurchenco, P.D. (2003). The role of laminin in embryonic cell polarization and tissue organization. Dev. Cell *4*, 613–624.

Lin, W., Srajer, G., Evrard, Y.A., Phan, H.M., Furuta, Y., and Dent, S.Y. (2007). Developmental potential of Gcn5(-/-) embryonic stem cells in vivo and in vitro. Dev. Dyn. *236*, 1547–1557.

Lin, W., Zhang, Z., Srajer, G., Chen, Y.C., Huang, M., Phan, H.M., and Dent, S.Y. (2008). Proper expression of the Gcn5 histone acetyltransferase is required for neural tube closure in mouse embryos. Dev. Dyn. *237*, 928–940.

Loebel, D.A., Studdert, J.B., Power, M., Radziewic, T., Jones, V., Coultas, L., Jackson, Y., Rao, R.S., Steiner, K., Fossat, N., et al. (2011). Rhou maintains the epithelial architecture and facilitates differentiation of the foregut endoderm. Development *138*, 4511–4522.

Martinez-Cerdeno, V., Lemen, J.M., Chan, V., Wey, A., Lin, W., Dent, S.R., and Knoepfler, P.S. (2012). N-Myc and GCN5 regulate significantly overlapping transcriptional programs in neural stem cells. PLoS One *7*, e39456.

McCarthy, R.L., Duncan, A.D., and Barton, M.C. (2017a). Sample preparation for mass cytometry analysis. J. Vis. Exp. https://doi.org/10.3791/54394.

McCarthy, R.L., Mak, D.H., Burks, J.K., and Barton, M.C. (2017b). Rapid monoisotopic cisplatin based barcoding for multiplexed mass cytometry. Sci. Rep. *7*, 3779.

Pai, Y.J., Leung, K.Y., Savery, D., Hutchin, T., Prunty, H., Heales, S., Brosnan, M.E., Brosnan, J.T., Copp, A.J., and Greene, N.D. (2015). Glycine decarboxylase deficiency causes neural tube defects and features of non-ketotic hyperglycinemia in mice. Nat. Commun. *6*, 6388.

Patel, J.H., Du, Y., Ard, P.G., Phillips, C., Carella, B., Chen, C.J., Rakowski, C., Chatterjee, C., Lieberman, P.M., Lane, W.S., et al. (2004). The c-MYC oncoprotein is a substrate of the acetyltransferases hGCN5/PCAF and TIP60. Mol. Cell. Biol. *24*, 10826–10834.

Pikman, Y., Puissant, A., Alexe, G., Furman, A., Chen, L.M., Frumm, S.M., Ross, L., Fenouille, N., Bassil, C.F., Lewis, C.A., et al. (2016). Targeting MTHFD2 in acute myeloid leukemia. J. Exp. Med. *213*, 1285–1306.

Qiu, P., Simonds, E.F., Bendall, S.C., Gibbs, K.D., Jr., Bruggner, R.V., Linderman, M.D., Sachs, K., Nolan, G.P., and Plevritis, S.K. (2011). Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. Nat. Biotechnol. *29*, 886–891.

Rivera-Perez, J.A., and Hadjantonakis, A.K. (2014). The dynamics of morphogenesis in the early mouse embryo. Cold Spring Harb. Perspect. Biol. *7*. https://doi.org/10.1101/cshperspect.a015867.

Sakai, T., Li, S., Docheva, D., Grashoff, C., Sakai, K., Kostka, G., Braun, A., Pfeifer, A., Yurchenco, P.D., and Fassler, R. (2003). Integrin-linked kinase (ILK) is required for polarizing the epiblast, cell adhesion, and controlling actin accumulation. Genes Dev. *17*, 926–940.

Serra, V., Eichhorn, P.J., Garcia-Garcia, C., Ibrahim, Y.H., Prudkin, L., Sanchez, G., Rodriguez, O., Anton, P., Parra, J.L., Marlow, S., et al. (2013). RSK3/4 mediate resistance to PI3K pathway inhibitors in breast cancer. J. Clin. Invest. *123*, 2551–2563.

She, P., Reid, T.M., Bronson, S.K., Vary, T.C., Hajnal, A., Lynch, C.J., and Hutson, S.M. (2007). Disruption of BCATm in mice leads to increased energy expenditure associated with the activation of a futile protein turnover cycle. Cell Metab. *6*, 181–194.

Shi, C., Welsh, P.A., Sass-Kuhn, S., Wang, X., McCloskey, D.E., Pegg, A.E., and Feith, D.J. (2012). Characterization of transgenic mice with overexpression of spermidine synthase. Amino Acids *42*, 495–505.

Snezhkina, A.V., Krasnov, G.S., Lipatova, A.V., Sadritdinova, A.F., Kardymon, O.L., Fedorova, M.S., Melnikova, N.V., Stepanov, O.A., Zaretsky, A.R., Kaprin, A.D., et al. (2016). The dysregulation of polyamine metabolism in colorectal cancer is associated with overexpression of c-Myc and C/EBPbeta rather than enterotoxigenic *Bacteroides fragilis* infection. Oxid. Med. Cell. Longev. *2016*, 2353560.

Spedale, G., Timmers, H.T., and Pijnappel, W.W. (2012). ATAC-king the complexity of SAGA during evolution. Genes Dev. *26*, 527–541.

Suganuma, T., Gutierrez, J.L., Li, B., Florens, L., Swanson, S.K., Washburn, M.P., Abmayr, S.M., and Workman, J.L. (2008). ATAC

is a double histone acetyltransferase complex that stimulates nucleosome sliding. Nat. Struct. Mol. Biol. *15*, 364–372.

Sumi, T., Oki, S., Kitajima, K., and Meno, C. (2013). Epiblast ground state is controlled by canonical Wnt/beta-catenin signaling in the postimplantation mouse embryo and epiblast stem cells. PLoS One *8*, e63378.

Szabo, N.E., Zhao, T., Cankaya, M., Theil, T., Zhou, X., and Alvarez-Bolado, G. (2009). Role of neuroepithelial Sonic hedgehog in hypothalamic patterning. J. Neurosci. *29*, 6989–7002.

Tefft, J.D., Lee, M., Smith, S., Leinwand, M., Zhao, J., Bringas, P., Jr., Crowe, D.L., and Warburton, D. (1999). Conserved function of mSpry-2, a murine homolog of *Drosophila* sprouty, which negatively modulates respiratory organogenesis. Curr. Biol. *9*, 219–222.

Timmers, H.T., and Tora, L. (2005). SAGA unveiled. Trends Biochem. Sci. *30*, 7–10.

Tudor, M., Murray, P.J., Onufryk, C., Jaenisch, R., and Young, R.A. (1999). Ubiquitous expression and embryonic requirement for RNA polymerase II coactivator subunit Srb7 in mice. Genes Dev. *13*, 2365–2368.

Urbanek, P., Fetka, I., Meisler, M.H., and Busslinger, M. (1997). Cooperation of Pax2 and Pax5 in midbrain and cerebellum development. Proc. Natl. Acad. Sci. USA *94*, 5703–5708.

Villegas, S.N., Rothova, M., Barrios-Llerena, M.E., Pulina, M., Hadjantonakis, A.K., Le Bihan, T., Astrof, S., and Brickman, J.M. (2013). PI3K/Akt1 signalling specifies foregut precursors by generating regionalized extra-cellular matrix. Elife *2*, e00806.

Wang, Y.L., Faiola, F., Xu, M., Pan, S., and Martinez, E. (2008). Human ATAC Is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein. J. Biol. Chem. *283*, 33808–33815.

Wen, H., Li, Y., Xi, Y., Jiang, S., Stratton, S., Peng, D., Tanaka, K., Ren, Y., Xia, Z., Wu, J., et al. (2014). ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. Nature *508*, 263–268.

Wilde, J.J., Siegenthaler, J.A., Dent, S.Y., and Niswander, L.A. (2017). Diencephalic size is restricted by a novel interplay between gcn5 acetyltransferase activity and retinoic acid signaling. J. Neurosci. *37*, 2565–2579.

Wu, Q., Zhang, L., Su, P., Lei, X., Liu, X., Wang, H., Lu, L., Bai, Y., Xiong, T., Li, D., et al. (2015). MSX2 mediates entry of human pluripotent stem cells into mesendoderm by simultaneously suppressing SOX2 and activating NODAL signaling. Cell Res. *25*, 1314–1332.

Xu, W., Edmondson, D.G., Evrard, Y.A., Wakamiya, M., Behringer, R.R., and Roth, S.Y. (2000). Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development. Nat. Genet. *26*, 229–232.

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

Yamauchi, T., Yamauchi, J., Kuwata, T., Tamura, T., Yamashita, T., Bae, N., Westphal, H., Ozato, K., and Nakatani, Y. (2000). Distinct but overlapping roles of histone acetylase PCAF and of the closely related PCAF-B/GCN5 in mouse embryogenesis. Proc. Natl. Acad. Sci. USA *97*, 11303–11306.

Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–523.

Yu, Y., Yoon, S.O., Poulogiannis, G., Yang, Q., Ma, X.M., Villen, J., Kubica, N., Hoffman, G.R., Cantley, L.C., Gygi, S.P., et al. (2011). Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. Science *332*, 1322–1326.

Zhang, X.Y., Varthi, M., Sykes, S.M., Phillips, C., Warzecha, C., Zhu, W., Wyce, A., Thorne, A.W., Berger, S.L., and McMahon, S.B. (2008a). The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression. Mol. Cell *29*, 102–111.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008b). Model-based analysis of ChIP-seq (MACS). Genome Biol. *9*, R137.

Stem Cell Reports, Volume *10*

Supplemental Information

GCN5 Regulates FGF Signaling and Activates Selective MYC Target Genes during Early Embryoid Body Differentiation

Li Wang, Evangelia Koutelou, Calley Hirsch, Ryan McCarthy, Andria Schibler, Kevin Lin, Yue Lu, Collene Jeter, Jianjun Shen, Michelle C. Barton, and Sharon Y.R. Dent

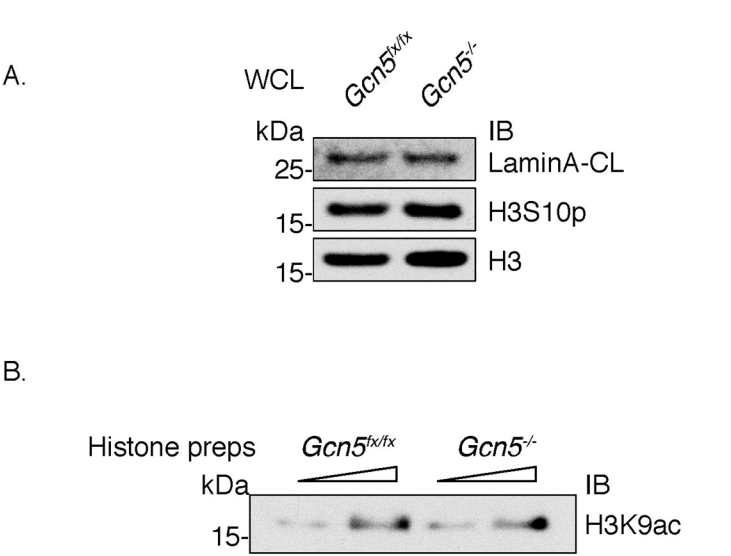


Figure S1 Loss of *Gcn5* does not cause overt defect in proliferation or apoptosis. Related to Figure 1

(A) Immunoblots of mitotic marker (H3S10p) and apoptosis marker (cleaved-LaminA) showed no differences between *Gcn5*$^{fx/fx}$ and *Gcn5*$^{-/-}$ EBs at day 5. WCL, whole cell lysates.

(B) Immunoblot of H3K9ac in day 5 EBs demonstrated no global changes upon loss of *Gcn5* at early stage of differentiation.

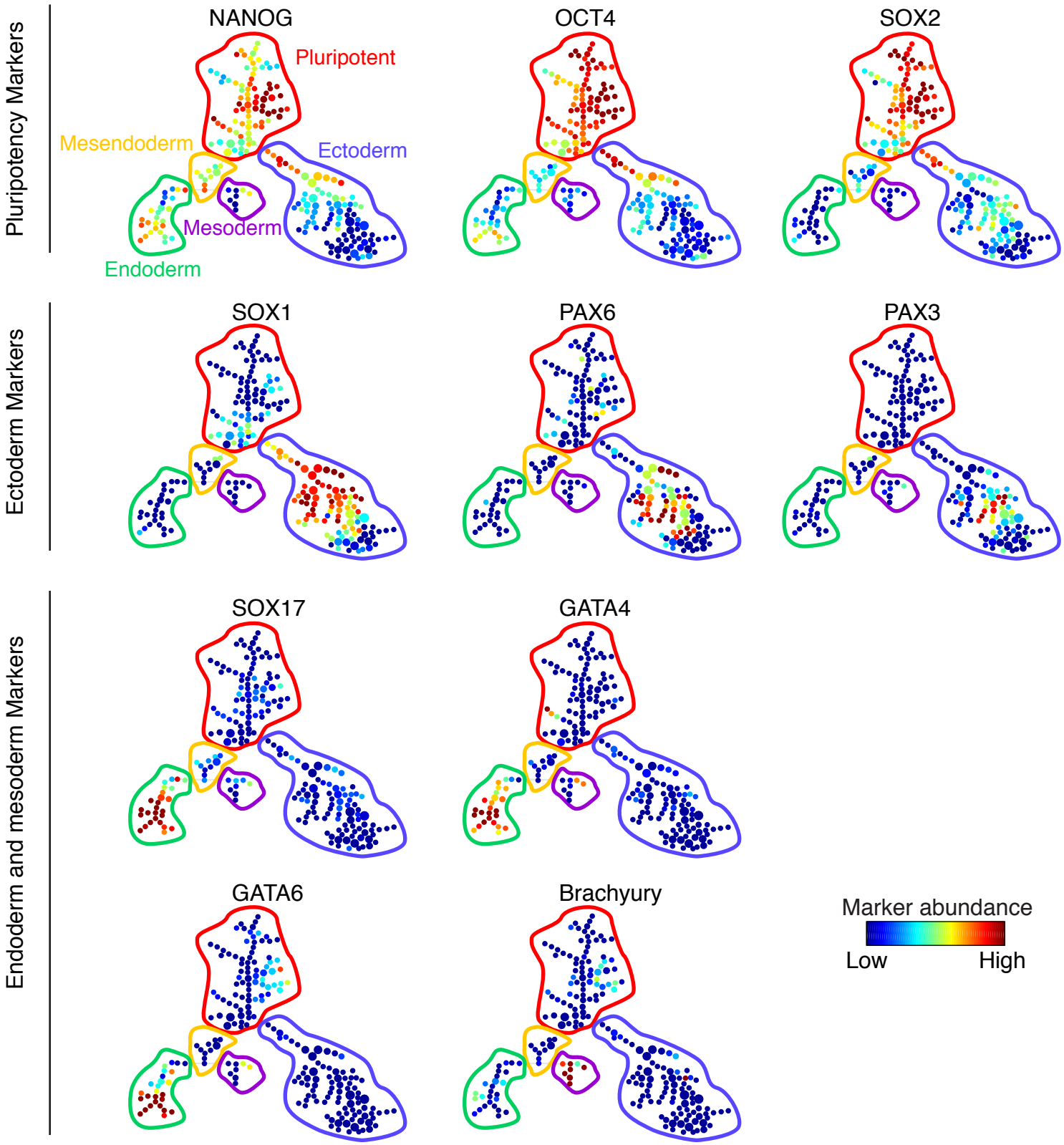


Figure S2 Mass cytometry to delineate heterogeneous cell populations during ESC differentiation. Related to Figure 2.

Proof-of-principle experiment showing lineage markers are enriched for corresponding cell populations. Upper panels: ES cells, NANOG, OCT4, and SOX2 enriched in the pluripotent region defined in red. Middle panels: differentiated EBs, SOX1, PAX6 and PAX3 enriched in the ectoderm region (blue). Lower panels: differentiated EBs, SOX17, GATA4, GATA6 enriched for the endoderm region (green) and Brachyury enriched in the mesoderm region (purple).

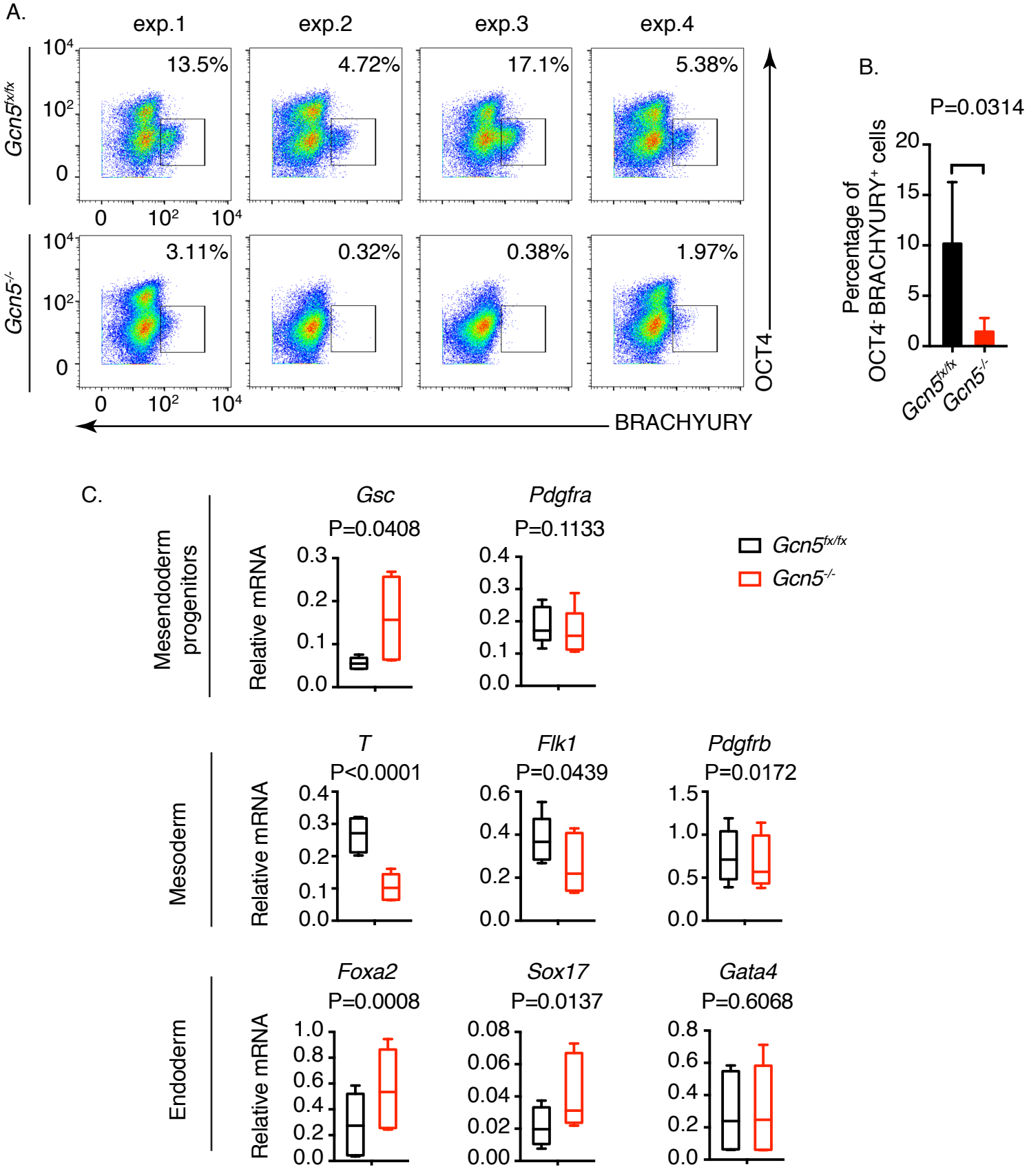
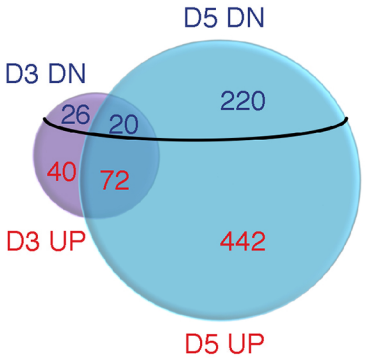
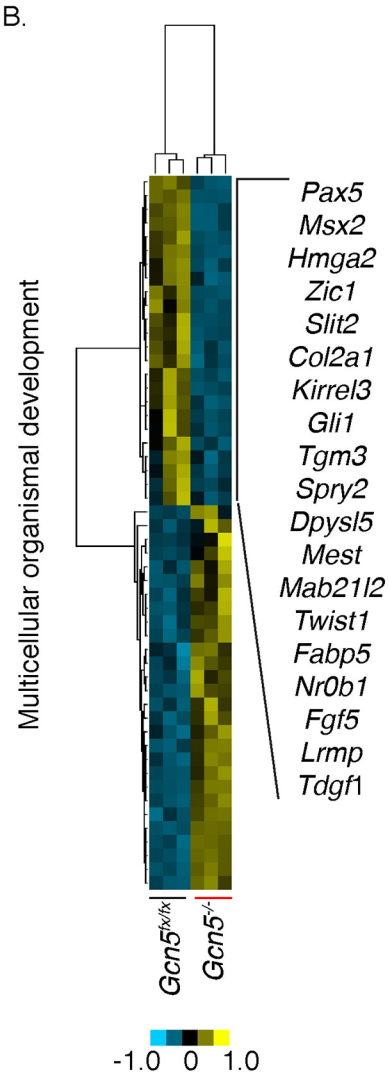


Figure S3 Loss of *Gcn5* impedes mesoderm differentiation. Related to Figure 2
(A) Gated mesoderm population of day 5 EBs from 4 independent mass cytometry experiments.
(B) Quantification of (A).
(C) qRT-PCR analysis of marker genes for indicated populations derived from monolayer mesoderm/endoderm differentiation of control and *Gcn5*$^{-/-}$ ES cells (n=3).
Data are presented as Mean ± SD, and student t-test was used for pair-wise comparisons.

A.

DE gene No. / Samples	Total	DN	UP
D3	158	46	112
D5	754	240	514
Overlap of D3^D5	104	20	72





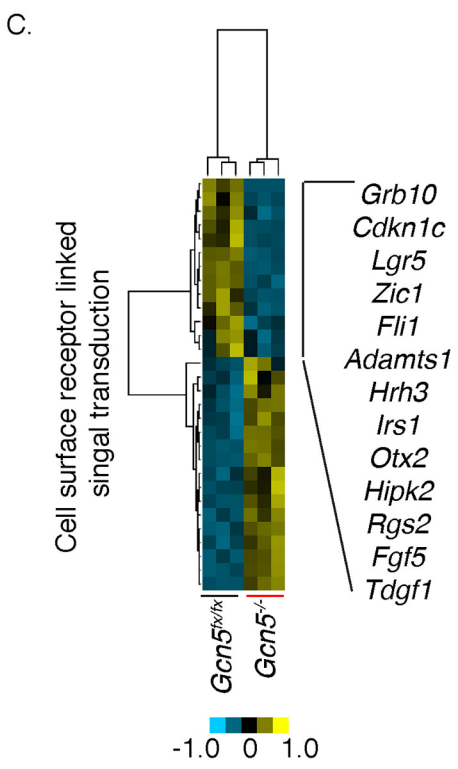


Figure S4 Loss of *Gcn5* impacted genes critical for development and signaling. Related to Figure 3

(A) Break down of the numbers of genes altered upon *Gcn5* loss in day 5 EBs.

(B and C) Heatmaps showing the top enriched genes in MOD (B) and CSRLST (C).

Color bars, normalized RPMK counts (False discovery rate 0.05, Fold change >=2)

Table S1 Lineage markers used for mass cytometry. Related to Figure 2

Antibodies	Expression	Isotope Label
Anti-NANOG	Pluripotency	163Dy
Anti-OCT4	Pluripotency	146Nd
Anti-SOX2	Pluripotency	147Sm
Anti-SOX1	Ectoderm	176Yb
Anti-PAX3	Ectoderm	170Er
Anti-PAX6	Ectoderm	153Eu
Anti-FOXA2	Endoderm	150Nd
Anti-GATA6	Endoderm	142Nd
Anti-GATA4	Endoderm	171Yb
Anti-SOX17	Endoderm	175Lu
Anti-Brachyury (T)	Mesoderm	156Gd
Anti-HAND1	Mesoderm	158Gd

Table S2 List of genes enriched in the Multicellular Organismal Development category identified by GSEA. Related to Figure 3

NAME	PROBE	GENE SYMBOL	GENE_TITLE	RANK IN GENE LIST	RANK METRIC SCORE	RUNNING ES	CORE ENRICHMENT
row_0	PAX5	PAX5	paired box gene 5 (B-cell lineage specific activator)	2	3.143970966	0.026952207	Yes
row_1	MSX2	MSX2	msh homeobox homolog 2 (Drosophila)	6	2.793720722	0.048327506	Yes
row_2	HMGA2	HMGA2	high mobility group AT-hook 2	10	2.60679245	0.067849986	Yes
row_3	ZIC1	ZIC1	Zic family member 1 (odd-paired homolog, Drosophila)	13	2.445233583	0.08787638	Yes
row_4	SLIT2	SLIT2	slit homolog 2 (Drosophila)	14	2.438150406	0.11204308	Yes
row_5	COL2A1	COL2A1	collagen, type II, alpha 1 (primary osteoarthritis, spondyloepiphyseal dysplasia, congenital)	20	2.313369513	0.12444666	Yes
row_6	KIRREL3	KIRREL3	kin of IRRE like 3 (Drosophila)	23	2.268698215	0.14272325	Yes
row_7	GLI1	GLI1	glioma-associated oncogene homolog 1 (zinc finger protein)	25	2.187872887	0.16230397	Yes
row_8	TGM3	TGM3	transglutaminase 3 (E polypeptide, protein-glutamine-gamma-glutamyltransferase)	27	2.169377327	0.18170136	Yes
row_9	SPRY2	SPRY2	sprouty homolog 2 (Drosophila)	31	2.100628138	0.1962068	Yes
row_10	DPYSL5	DPYSL5	dihydropyrimidinase-like 5	33	2.077706337	0.21469556	Yes
row_11	MEST	MEST	mesoderm specific transcript homolog (mouse)	44	1.956736565	0.21303791	Yes
row_12	MAB21L2	MAB21L2	mab-21-like 2 (C. elegans)	54	1.908952236	0.21301188	Yes
row_13	TWIST1	TWIST1	twist homolog 1 (acrocephalosyndactyly 3; Saethre-Chotzen syndrome) (Drosophila)	55	1.892674208	0.23177189	Yes
row_14	FABP5	FABP5	fatty acid binding protein 5 (psoriasis-associated)	57	1.876633048	0.24826762	Yes
row_15	OTX2	OTX2	orthodenticle homolog 2 (Drosophila)	74	1.794868827	0.23237398	No
row_16	PTCH2	PTCH2	patched homolog 2 (Drosophila)	81	1.774591208	0.23733197	No
row_17	NDP	NDP	Norrie disease (pseudoglioma)	101	1.670649171	0.21389128	No
row_18	FGF17	FGF17	fibroblast growth factor 17	117	1.568956375	0.19786368	No
row_19	EPHA2	EPHA2	EPH receptor A2	127	1.409094691	0.19288312	No
row_20	NR0B1	NR0B1	nuclear receptor subfamily 0, group B, member 1	134	1.361806631	0.19374964	No
row_21	FGF5	FGF5	fibroblast growth factor 5	160	1.005194783	0.15108146	No
row_22	LRMP	LRMP	lymphoid-restricted membrane protein	162	1.003890395	0.15892665	No
row_23	TDGF1	TDGF1	teratocarcinoma-derived growth factor 1	166	0.946586847	0.16199334	No
row_24	ALOX12B	ALOX12B	arachidonate 12-lipoxygenase, 12R type	174	-1.041075826	0.15757555	No
row_25	TRIM14	TRIM14	tripartite motif-containing 14	181	-1.18687892	0.1567082	No
row_26	ANXA2	ANXA2	annexin A2	198	-1.373691678	0.13663988	No
row_27	SGCD	SGCD	sarcoglycan, delta (35kDa dystrophin-associated glycoprotein)	201	-1.397668123	0.1462829	No
row_28	ETS1	ETS1	v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)	217	-1.476131439	0.12933522	No
row_29	ERG	ERG	v-ets erythroblastosis virus E26 oncogene homolog (avian)	222	-1.50989151	0.13588007	No
row_30	RAPGEFL1	RAPGEFL1	Rap guanine nucleotide exchange factor (GEF)-like 1	227	-1.535297155	0.14267674	No
row_31	IGFBP3	IGFBP3	insulin-like growth factor binding protein 3	249	-1.666120291	0.11498063	No
row_32	SGCE	SGCE	sarcoglycan, epsilon	251	-1.675953865	0.12948726	No

row_33	JAG2	JAG2	jagged 2	253	-1.683297992	0.14406668	No
row_34	SGCG	SGCG	sarcoglycan, gamma (35kDa dystrophin-associated glycoprotein)	282	-1.763670683	0.10260065	No
row_35	MYH3	MYH3	myosin, heavy chain 3, skeletal muscle, embryonic	290	-1.779546738	0.10550249	No
row_36	IGFBP4	IGFBP4	insulin-like growth factor binding protein 4	301	-1.796701074	0.10225859	No
row_37	CACNA1H	CACNA1H	calcium channel, voltage-dependent, alpha 1H subunit	311	-1.824573159	0.10139622	No
row_38	SHOX2	SHOX2	short stature homeobox 2	315	-1.830768824	0.11322682	No
row_39	AEBP1	AEBP1	AE binding protein 1	340	-1.907915235	0.08161158	No
row_40	RASGRP4	RASGRP4	RAS guanyl releasing protein 4	376	-2.075456142	0.028499093	No
row_41	NEUROG3	NEUROG3	neurogenin 3	385	-2.098732948	0.032459423	No
row_42	FGF11	FGF11	fibroblast growth factor 11	389	-2.107631922	0.04703428	No
row_43	ALDH3A2	ALDH3A2	aldehyde dehydrogenase 3 family, member A2	397	-2.13826108	0.05349167	No
row_44	MEF2C	MEF2C	MADS box transcription enhancer factor 2, polypeptide C (myocyte enhancer factor 2C)	398	-2.148215771	0.07478458	No
row_45	STAT3	STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)	401	-2.162971735	0.09201321	No
row_46	CRIM1	CRIM1	cysteine rich transmembrane BMP regulator 1 (chordin-like)	429	-2.307478189	0.05804261	No
row_47	NMUR2	NMUR2	neuromedin U receptor 2	462	-2.529047012	0.015741859	No
row_48	DMRT1	DMRT1	doublesex and mab-3 related transcription factor 1	474	-2.65857625	0.01893551	No
row_49	EVPL	EVPL	envoplakin	485	-2.816974878	0.02580446	No
row_50	GFRA3	GFRA3	GDNF family receptor alpha 3	486	-2.819857836	0.053754613	No
row_51	SNAI2	SNAI2	snail homolog 2 (Drosophila)	492	-2.860271454	0.07157903	No

Table S3 List of genes enriched in the Cell Surface Receptor Linked Signal Transduction category identified by GSEA. Related to Figure 3.

NAME	PROBE	GENE SYMBOL	GENE_TITLE	RANK IN GENE LIST	RANK METRIC SCORE	RUNNING ES	CORE ENRICHMENT
row_0	GRB10	GRB10	growth factor receptor-bound protein 10	0	4.903269768	0.07419716	Yes
row_1	CDKN1C	CDKN1C	cyclin-dependent kinase inhibitor 1C (p57, Kip2)	1	3.456902027	0.12650761	Yes
row_2	LGR5	LGR5	leucine-rich repeat-containing G protein-coupled receptor 5	11	2.555413246	0.14706793	Yes
row_3	ZIC1	ZIC1	Zic family member 1 (odd-paired homolog, Drosophila)	13	2.445233583	0.18205757	Yes
row_4	GLI1	GLI1	glioma-associated oncogene homolog 1 (zinc finger protein)	25	2.187872887	0.19303207	Yes
row_5	ADAMTS1	ADAMTS1	ADAM metallopeptidase with thrombospondin type 1 motif, 1	36	2.065867424	0.20417242	Yes
row_6	HRH3	HRH3	histamine receptor H3	37	2.059562922	0.23533809	Yes
row_7	IRS1	IRS1	insulin receptor substrate 1	53	1.913015485	0.2341051	No
row_8	OTX2	OTX2	orthodenticle homolog 2 (Drosophila)	74	1.794868827	0.22102393	No
row_9	HIPK2	HIPK2	homeodomain interacting protein kinase 2	106	1.659950376	0.18376835	No
row_10	RGS2	RGS2	regulator of G-protein signalling 2, 24kDa	115	1.601378202	0.19190411	No
row_11	FGF5	FGF5	fibroblast growth factor 5	160	1.005194783	0.11858372	No
row_12	TDGF1	TDGF1	teratocarcinoma-derived growth factor 1	166	0.946586847	0.12284728	No
row_13	STC1	STC1	stanniocalcin 1	186	-1.224491358	0.10314713	No
row_14	FLT4	FLT4	fms-related tyrosine kinase 4	191	-1.312462449	0.114959255	No
row_15	RAPGEFL1	RAPGEFL1	Rap guanine nucleotide exchange factor (GEF)-like 1	227	-1.535297155	0.06776911	No
row_16	TBXA2R	TBXA2R	thromboxane A2 receptor	256	-1.689142942	0.0369915	No
row_17	RAMP1	RAMP1	receptor (calcitonin) activity modifying protein 1	285	-1.764481425	0.007353922	No
row_18	GPR20	GPR20	G protein-coupled receptor 20	295	-1.789713264	0.016327534	No
row_19	IL13RA1	IL13RA1	interleukin 13 receptor, alpha 1	356	-1.987249494	-0.07432539	No
row_20	BDKRB2	BDKRB2	bradykinin receptor B2	360	-2.004709721	-0.050025985	No
row_21	SOCS1	SOCS1	suppressor of cytokine signaling 1	381	-2.084973097	-0.058717243	No
row_22	OPRD1	OPRD1	opioid receptor, delta 1	394	-2.121104479	-0.05076518	No
row_23	RASD1	RASD1	RAS, dexamethasone-induced 1	439	-2.354247332	-0.10367147	No
row_24	GABRA4	GABRA4	gamma-aminobutyric acid (GABA) A receptor, alpha 4	442	-2.365311384	-0.071903296	No
row_25	NMUR2	NMUR2	neuromedin U receptor 2	462	-2.529047012	-0.07186268	No
row_26	CD274	CD274	CD274 molecule	463	-2.539172888	-0.03343946	No
row_27	LAT	LAT	linker for activation of T cells	500	-2.93972373	-0.06138964	No
row_28	IL12RB1	IL12RB1	interleukin 12 receptor, beta 1	521	-3.552495718	-0.04787409	No
row_29	CLEC1A	CLEC1A	C-type lectin domain family 1, member A	524	-3.695590973	0.004024112	No

Table S4 Top ranked transcription factors or regulators reported to bind genes with decreased H3K9ac identified in *Gcn5*$^{-/-}$ EBs at day 5. Related to Figure 5.

Rank	TFs or regulators	Number of genes with H3K9ac decrease	Q-value
1	HCFC1	205	1.25E-77
2	MAX	204	4.43E-72
3	MXI1	199	4.25E-66
4	GCN5	172	2.65E-61
5	NELFE	192	2.65E-61
6	TBP	192	1.69E-59
7	CTCF	210	3.77E-59
8	SIN3A	192	9.08E-59
9	P300	188	1.14E-58
10	C-MYC	164	2.44E-57
11	E2F4	99	1.40E-48
12	ZNF	145	1.69E-43
13	POL2	218	4.62E-43
14	FLI1	103	1.48E-42
15	CHD2	143	5.48E-35

Supplemental Experimental Procedures

Chromatin immunoprecipitation

The dissociated cells from day 5 EBs were cross-linked with 1% formaldehyde (Thermo Scientific™, PI28906) for 10 minutes at room temperature then quenched with 125 mM glycine for 5 minutes. The cells were washed with ice-cold PBS containing protease inhibitor cocktail (PI, Sigma, P8340), and then resuspended in swelling buffer (5 mM PIPES pH 8.0, 85 mM KCl and 1% NP40) containing PI for 20 minutes on ice. The nuclei were pelleted and lysed in nuclei lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA and 1% SDS) and then sonicated using a Bioruptor® Plus sonication device (Diagenode B01020001). A total of 15 minutes of sonication (3 rounds of 10 cycles with 30 seconds on and 30 seconds off per cycle on the high setting) was applied to obtain chromatin fragments in the size of 150-300 bps. Sonicated samples were centrifuged to remove the insoluble debris. 30ug of chromatin fragments were diluted 1:10 in ChIP Dilution Buffer (0.01% SDS, 1% TritonX-100, 1mM EDTA, 20mM Tris-HCl, pH8.0 and 150mM NaCl) and were precleared with Dynabeads Protein A (Invitrogen™, 10002D) for 1 hour at 4°C. Precleared lysates were incubated with appropriate amount of antibodies (following manufacturer instructions) at 4°C overnight, followed by incubation with Dynabeads™ Protein A for 1 hour at 4°C. Immunoprecipitates were washed as previously described. All solutions used in the steps above were supplemented with PI. The DNA was eluted in elution buffer (50mM NaHCO3 and 1% SDS) at room temperature for 15 minutes, de-crosslinked at 65°C overnight, treated with RNaseA for 1 hour at 37°C and purified using a PCR purification kit (Qiagen, 28104).

ChIP library preparation

ChIP libraries were prepared using a Kapa Hyper Preparation kit (KAPA Biosystems, Wilmington, MA) protocol for Illumina Platforms. Briefly, for each library, 5ng of ChIP DNA was end-repaired and 3'-adenylated using a proprietary master mix, then ligated to the specific NexTflex adaptors from Bioo Scientific (Bioo Scientific, Austin, TX). The adaptor-ligated DNA was enriched using a KAPA Hyper Library Preparation kit (KAPA Biosystems, KK8502) with 5 cycles of PCR (1 cycle at 98°C for 45 seconds; 4 cycles of 98°C for 15 seconds, 60°C for 30 seconds, and 72°C for 30 seconds; 1 cycle at 72°C for 1 minute), then purified with AmpureXP beads (Beckman Coulter, A63881). The library quality was validated on a 2200 TapeStation from Agilent Technologies (Agilent, Santa Clara, CA). Concentrations of the libraries were determined using a Kapa Library Quantification Kit (KAPA Biosystems, KK4933) and loaded on cBOT (Illumina) at final concentration of 1.5nM for cluster generation, then sequenced with 50bp single-read on a HiSeq3000 sequencer (Illumina).

ChIPseq data analysis

Mapping: Sequenced DNA reads were mapped to mouse genome mm10 using bowtie (version 1.1.2) (Langmead et al., 2009) with at most 2 mismatches allowed and only the reads that were mapped to unique position were retained. 34-47 million reads were generated per sample. 90-94% reads were mapped to mouse genome, with 73-84% uniquely mapped. To avoid PCR bias, for multiple reads that were mapped to the same genomic position, only one copy was retained for further analysis. 19-29 million reads were used in the final analyses.

Peak Calling: H3K9ac peaks were detected by MACS (version 1.4.2) (Zhang et al., 2008). The window size was set as 500 bp and the p-value cutoff was 1e-5. H3K9ac peaks were initially called by comparing to the corresponding total H3. Then the peaks that overlapped ENCODE blacklisted regions (Consortium, 2012) or were not called by comparing to the corresponding total input were removed.

Differential Peak Calling: The peaks of all H3K9ac samples were merged and the numbers of reads in these merged peaks were counted for each H3K9ac sample. The count table was used to call differential H3K9ac peaks between *Gcn5*$^{fx/fx}$ and *Gcn5*$^{-/-}$ by edgeR (Robinson et al., 2010). Batch effect among the 4 replicates was corrected following edgeR user's guide. Peaks with FDR (false discovery rate) ≤ 0.05 were called as changed between *Gcn5*$^{fx/fx}$ and *Gcn5*$^{-/-}$. Genes with differential peaks in promoter (defined as -1000 bp to +500bp from TSS) were called as associated with changed H3K9ac.

Signal Landscape: Each read was extended by 150bp to its 3' end. The number of reads on each genomic position was rescaled to normalize the effective library size by edgeR to 10M and averaged over every 10bp window. The normalized values were displayed in UCSC genome browser (Kent et al., 2002).

Transcription Factor Binding: ENCODE ChIP-Seq Significance Tool (Auerbach et al., 2013) was used to identify enriched ENCODE transcription factors in the promoters of the genes associated with changed H3K9ac. The promoter was defined as -5000 bp to +2000 bp from TSS.

RNAseq data analysis

Mapping: The reads were mapped to mouse genome (mm10) by TopHat (version 2.0.10) (Kim et al., 2013) with an overall mapping rate of 84-94%. 72-91% fragments have both ends mapped to mouse genome.

Differential Expression: The number of fragments in each known gene from GENCODE Release M8 (Mudge and Harrow, 2015) was enumerated using htseq-count from HTSeq package (version 0.6.0) (Anders et al., 2015). Genes with less than 10 fragments in all the samples were removed before differential expression analysis. The differential expression between conditions was statistically assessed by R/Bioconductor package DESeq (Anders and Huber, 2010) (version 1.16.0). Genes with FDR (false discovery rate) ≤ 0.05, fold change ≥ 2 and length > 200bp were called as differentially expressed.

Principle Component Analysis (PCA): PCA was performed by R function prcomp using cpm (count of fragments in each gene per million of fragments mapped to all exons) values. The scale option was set as TRUE.

Heatmap: The normalized counts from DESeq were used to generate heatmap by Cluster 3.0 (de Hoon et al., 2004) and Java Treeview (Saldanha, 2004). The values in each gene were centered by median and rescaled so that the sum of the squares of the values is 1.0.

Gene Function and Pathway Analysis: The differential genes called by DESeq were used for Ingenuity Pathway Analysis (IPA) and Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005).

Monolayer differentiation of mESCs

Early mesoderm and endoderm lineages were generated following the protocols of Villegas et al., 2013 and (Orlova et al., 2014) with some modifications for mESCs. Briefly, mESCs were cultured on Col-IV coated plates for 2 days in differentiation media supplemented with B27, N2 and ROCK inhibitor (2.5μM). At day 2 the medium was replaced with MEDF (DMEM-High glucose (HyClone™, SH3002201) medium supplemented with 2% (v/v) FBS (Gibco™,10437-028), 0.1mM non-essential amino acids (Corning™, MT25025CI), 2mM L-glutamine (Hyclone, SH3003401), 1% (v/v) penicillin/streptomycin (Hyclone, SV30010), 0.1 mM β-mercaptoethanol (BME) (Fisher, 03446I-100), and 1mM sodium pyruvate (Gibco™, 11360070)) for 24 hours, then supplemented with Activin A (50ng/mL) for two additional days to induce differentiation towards early mesoderm and endoderm lineages.

Antibodies used in this study

Antibodies	Manufacturers	Catalog No.	Applications
Anti-phospho-ERK	Cell Signaling Technology	4370	Westerns
Anti-ERK	Cell Signaling Technology	4695	Westerns
Anti-phospho-AKT	Cell Signaling Technology	4060	Westerns
Anti-AKT	Cell Signaling Technology	4691	Westerns
Anti-phospho-p38	Cell Signaling Technology	4511	Westerns
Anti-p38	Cell Signaling Technology	9212	Westerns
Anti-phospho-c-RAF (S259)	Cell Signaling Technology	9421	Westerns
Anti-c-RAF	Cell Signaling Technology	53745	Westerns
Anti-FGFR1	Cell Signaling Technology	9740	Westerns
Anti-cMYC	Cell Signaling Technology	9402	Westerns/ChIP
Anti-H3	Abcam	ab1791	ChIP
Anti-H3K9ac	Millipore	07-352	ChIP
Anti-Rabbit IgG	Millipore	12-370	ChIP
Anti-SOX1	BD Biosciences	560749	IF/Mass Cytometry
Anti-GATA4	Abcam	ab84593	IF
Anti-Laminin	Millipore	AB2034	IF
Anti-Vimentin	Abcam	ab92547	IF
Alexa Fluor 568 Phalloidin	ThermoFisher Scienctific	A12380	IF
Donkey anti-rabbit IgG Alexa Fluor® 488	ThermoFisher Scienctific	A21206	IF
Donkey anti-rabbit IgG Alexa Fluor® 555	ThermoFisher Scienctific	A31572	IF
Donkey anti-rabbit IgG Alexa Fluor® 647	ThermoFisher Scienctific	A31573	IF
Donkey anti-mouse IgG Alexa Fluor® 488	ThermoFisher Scienctific	A21202	IF
Goat Anti-mouse IgG Alexa Fluor® 568	ThermoFisher Scienctific	A11004	IF
Donkey anti-mouse IgG Alexa Fluor® 647	ThermoFisher Scienctific	A31571	IF
Anti-NANOG	Cell Signaling Technology	3580	Mass Cytometry
Anti-OCT4	Santa Cruz	sc-5279	Mass Cytometry
Anti-SOX2	R&D Systems	MAB2018	Mass Cytometry
Anti-GATA6	R&D Systems	AF1700	Mass Cytometry
Anti-PAX3	R&D Systems	MAB2457	Mass Cytometry
Anti-PAX6	R&D Systems	AF8510	Mass Cytometry
Anti-BRACHYURY	R&D Systems	AF2085	Mass Cytometry
Anti-HAND1	R&D Systems	AF3168	Mass Cytometry
Anti-FOXA2	BD Biosciences	561580	Mass Cytometry
Anti-GATA4	BD Biosciences	560327	Mass Cytometry
Anti-SOX17	BD Biosciences	561590	Mass Cytometry

Primers used in this study

Oligo name	Sequences (5'-3')	Source
qRT-PCR (RNA analysis)		
Cdh2 fwd	CAGGGTGGACGTCATTGTAG	(Kamiya et al., 2011)
Cdh2 rev	AGGGTCTCCACCACTGATTC	
Fgf3 fwd	ACAGGCGGGAAGCATATGTA	Originally designed
Fgf3 rev	GGCCATGAACAAGAGAGGAC	
Fgf4 fwd	CGTTGTAGTTGTTGGGCAGA	Originally designed
Fgf4 rev	TTCTTCGTGGCTATGAGCAG	
Fgf5 fwd	GCGATCCACAGAACTGAAAA	(Kamiya et al., 2011)
Fgf5 rev	ACTGCTTGAACCTGGGTAGG	
Foxa2 fwd	GAGCAGCAACATCACCACAG	Originally designed
Foxa2 rev	CGTAGGCCTTGAGGTCCAT	
Gata6 fwd	CAAAAGCTTGCTCCGGTAAC	(Kamiya et al., 2011)
Gata6 rev	TGAGGTGGTCGCTTGTGTAG	
Grb10 fwd	ATCTTCCGTTTCCCATTTCC	Originally designed
Grb10 rev	CTCCTTACCTCCTCCTCCGA	
Otx2 fwd	CTTCATGAGGGAAGAGGTGG	Originally designed
Otx2 rev	GGCCTCACTTTGTTCTGACC	
Pbgd fwd	CAGGGTACAAGGCTTTCAGC	Originally designed
Pbgd rev	CGGAGTCATGTCCGGTAAC	
Prcka fwd	AACGAACTCATGGCACCTCT	Originally designed
Prcka rev	CACTGCACCGACTTCATCTG	
Sox1 fwd	CCTCGGATCTCTGGTCAAGT	(Kamiya et al., 2011)
Sox1 rev	GCAGGTACATGCTGATCATCTC	
Spry4 fwd	AGGTCCTGAACTGCACCAAG	Originally designed
Spry4 rev	GGGGATTTACACAGACGTGG	
Stat3 fwd	CTGCTCCAGGTAGCGTGTGT	Originally designed
Stat3 rev	CTCAGCCCCGGAGACAGT	
T fwd	CTGGGAGCTCAGTTCTTTCG	(Kamiya et al., 2011)
T rev	CCCCTTCATACATCGGAGAA	
ChIP-qPCR		
Bcat1 fwd	GGGTGCAAATGTGAGTCTCC	Originally designed
Bcat1 rev	GCCCAGCTCTCCATCTTCC	
Rps6ka2 fwd	CCTCACCGAGAGGAGGAAG	Originally designed
Rps6ka2 rev	CCCTGCAACTCCTTGCTTAT	
Mthfd2 fwd	AAGCGTCCGCATCTCCAC	Originally designed
Mthfd2 rev	TATCCTTCCCAAGCATCACC	
Intergenic fwd	AAGGGGCCTCTGCTTAAAAA	Originally designed
Intergenic rev	AGAGCTCCATGGCAGGTAGA	

References

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol *11*, R106.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.
Auerbach, R.K., Chen, B., and Butte, A.J. (2013). Relating genes to function: identifying enriched transcription factors using the ENCODE ChIP-Seq significance tool. Bioinformatics *29*, 1922-1924.
Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.
de Hoon, M.J., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. Bioinformatics *20*, 1453-1454.
Kamiya, D., Banno, S., Sasai, N., Ohgushi, M., Inomata, H., Watanabe, K., Kawada, M., Yakura, R., Kiyonari, H., Nakao, K.*, et al.* (2011). Intrinsic transition of embryonic stem-cell differentiation into neural progenitors. Nature *470*, 503-509.
Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res *12*, 996-1006.
Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.
Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.
Mudge, J.M., and Harrow, J. (2015). Creating reference gene annotation for the mouse C57BL6/J genome assembly. Mamm Genome *26*, 366-378.
Orlova, V.V., van den Hil, F.E., Petrus-Reurer, S., Drabsch, Y., Ten Dijke, P., and Mummery, C.L. (2014). Generation, expansion and functional analysis of endothelial cells and pericytes derived from human pluripotent stem cells. Nat Protoc *9*, 1514-1531.
Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.
Saldanha, A.J. (2004). Java Treeview--extensible visualization of microarray data. Bioinformatics *20*, 3246-3248.
Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.