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

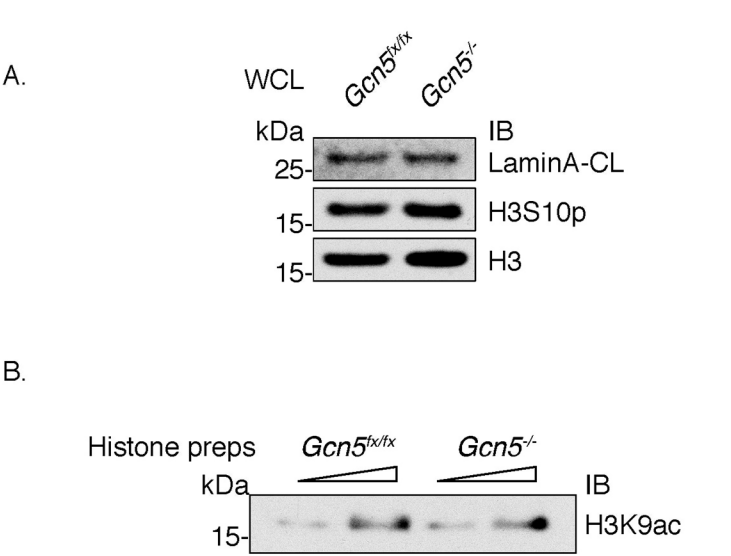


Figure S1 Loss of *Gcn5* does not cause overt defect in proliferation or apoptosis. Related to Figure 1

(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.

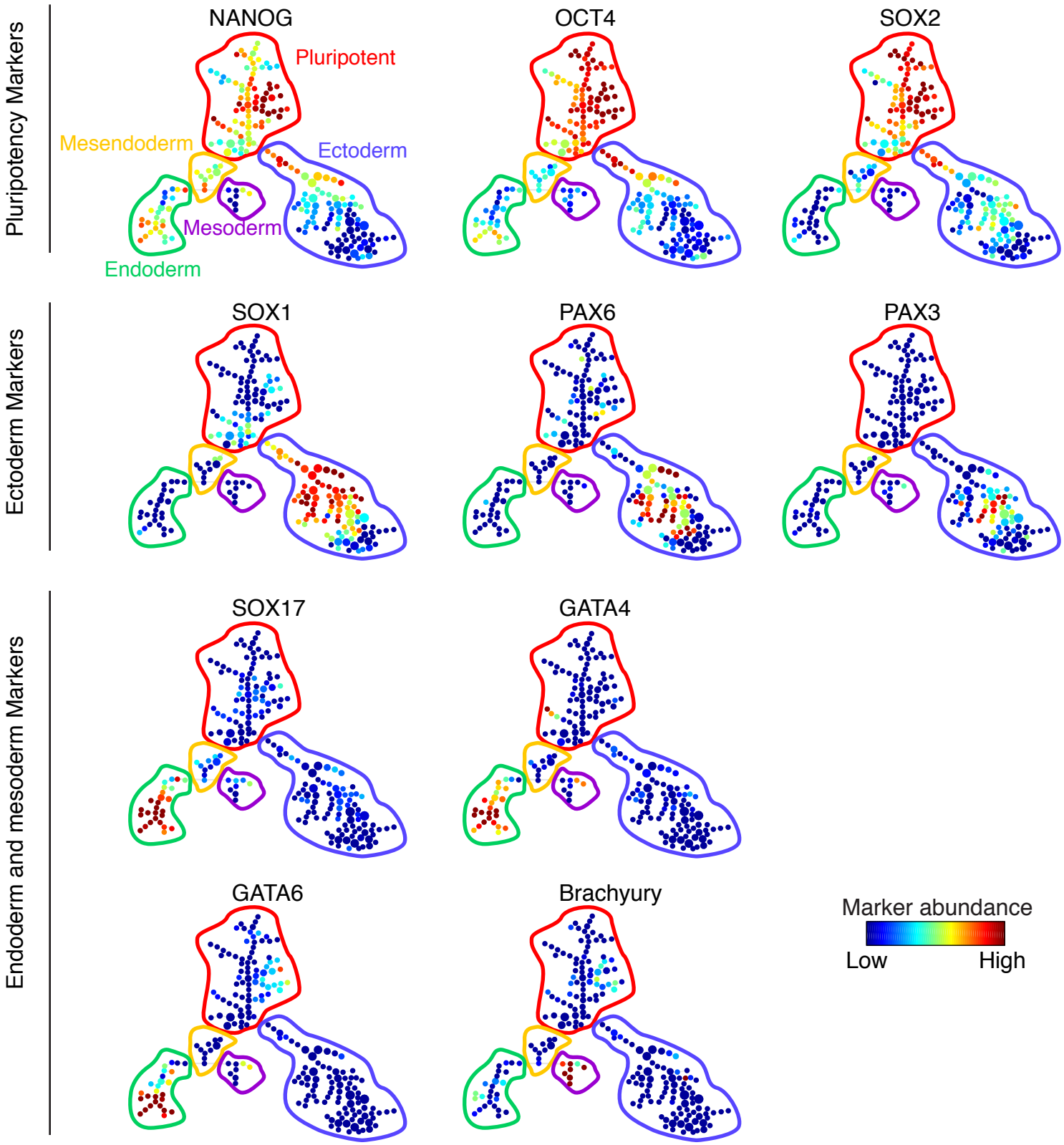


Figure S2 Mass cytometry to delineate heterogeneous cell populations during ESC differentiation. Related to Figure 2.

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).

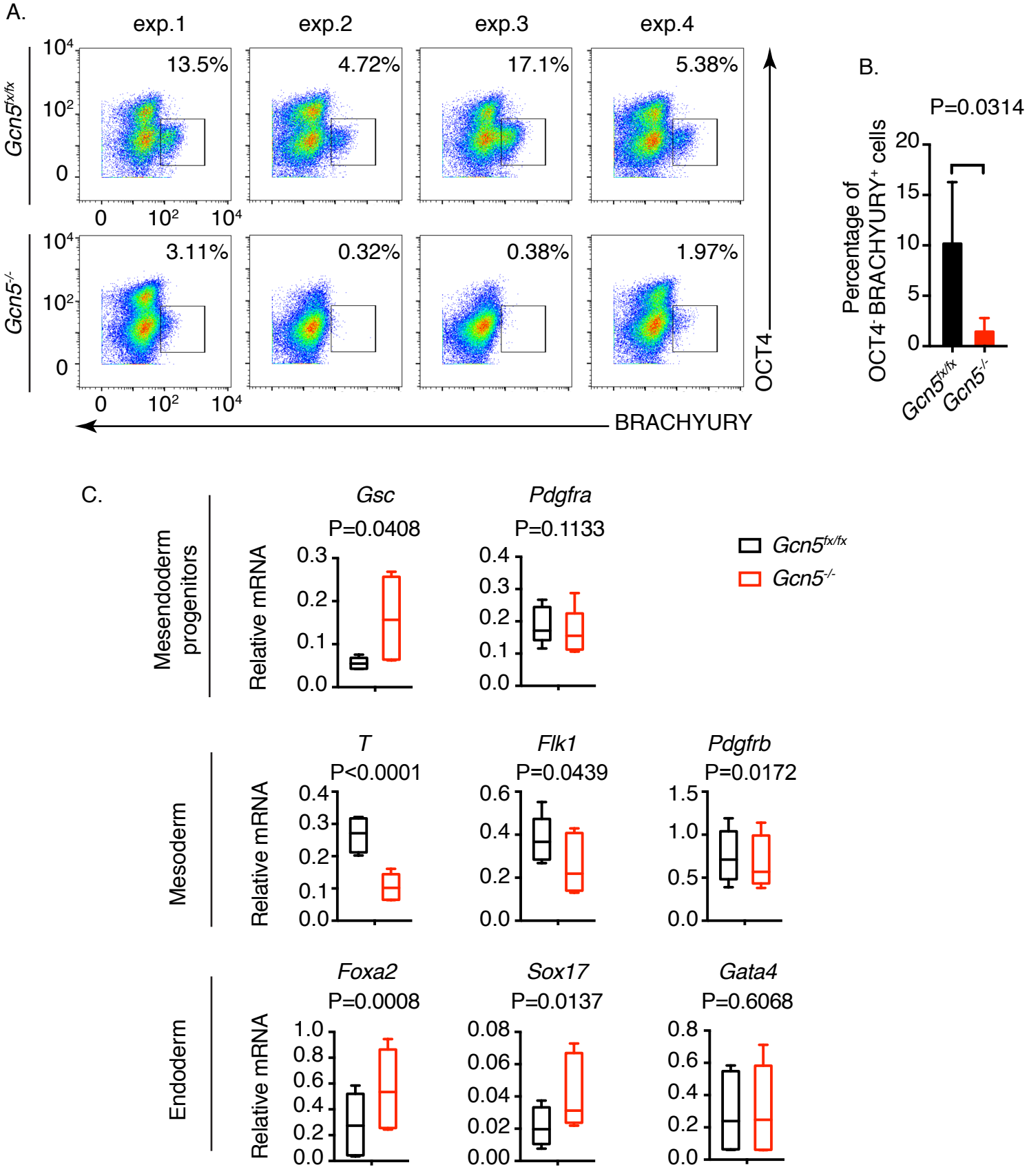
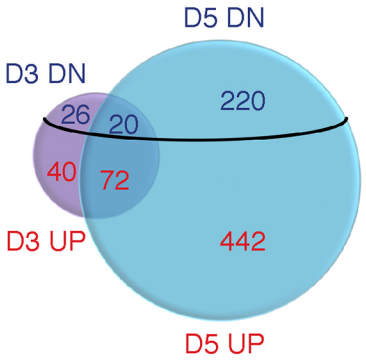
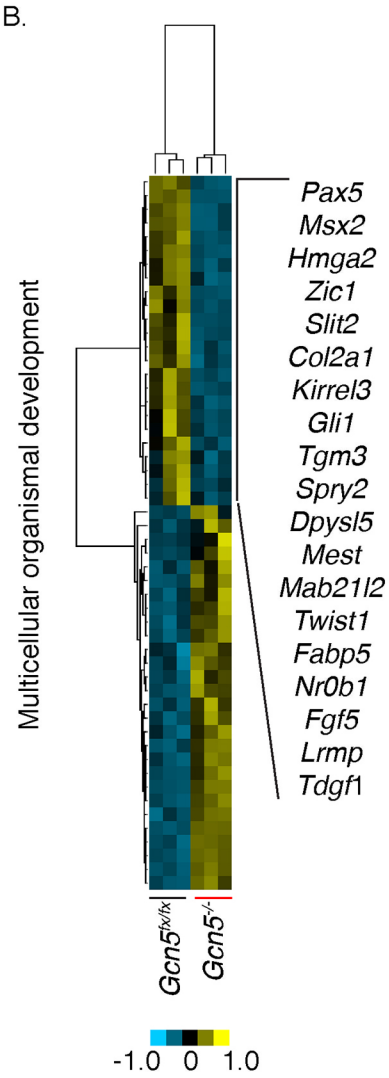


Figure S3 Loss of *Gcn5* impedes mesoderm differentiation. Related to Figure 2
(A) Gated mesoderm population of day 5 EBs from 4 independent mass cytometry experiments.
(B) Quantification of (A).
(C) qRT-PCR analysis of marker genes for indicated populations derived from monolayer mesoderm/endoderm differentiation 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





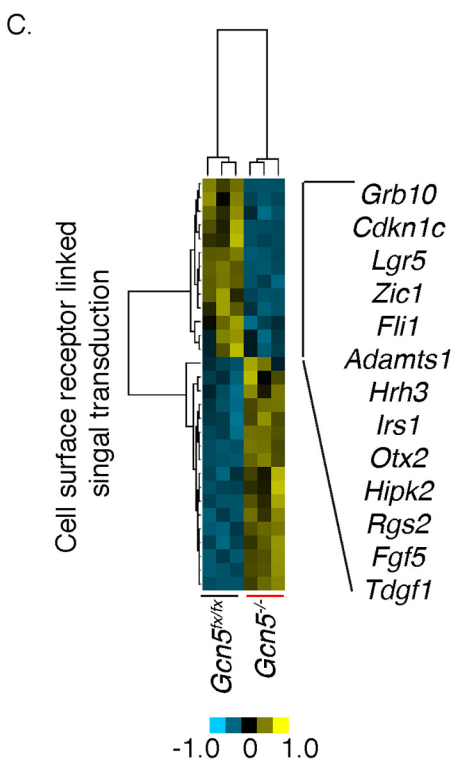


Figure S4 Loss of *Gcn5* impacted genes critical for development and signaling. Related to Figure 3

(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.