Stem Cell Reports, Volume *16*

Supplemental Information

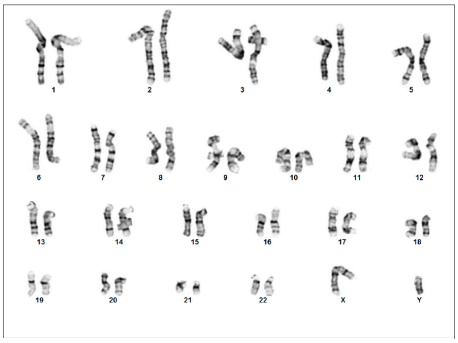
Loss of TBX3 enhances pancreatic progenitor generation from human pluripotent stem cells

Somdutta Mukherjee, Deborah L. French, and Paul Gadue

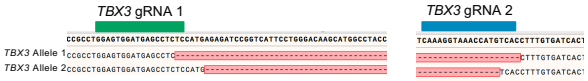
Figure S1: Characterization of *TBX3* knockout PSC lines supporting all figures

A

B



C



D

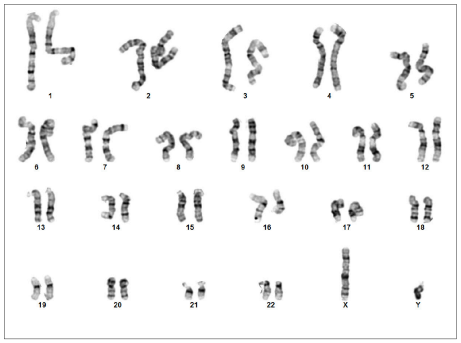


Figure S2: Loss of *TBX3* impairs hepatocyte differentiation supporting Figure 1

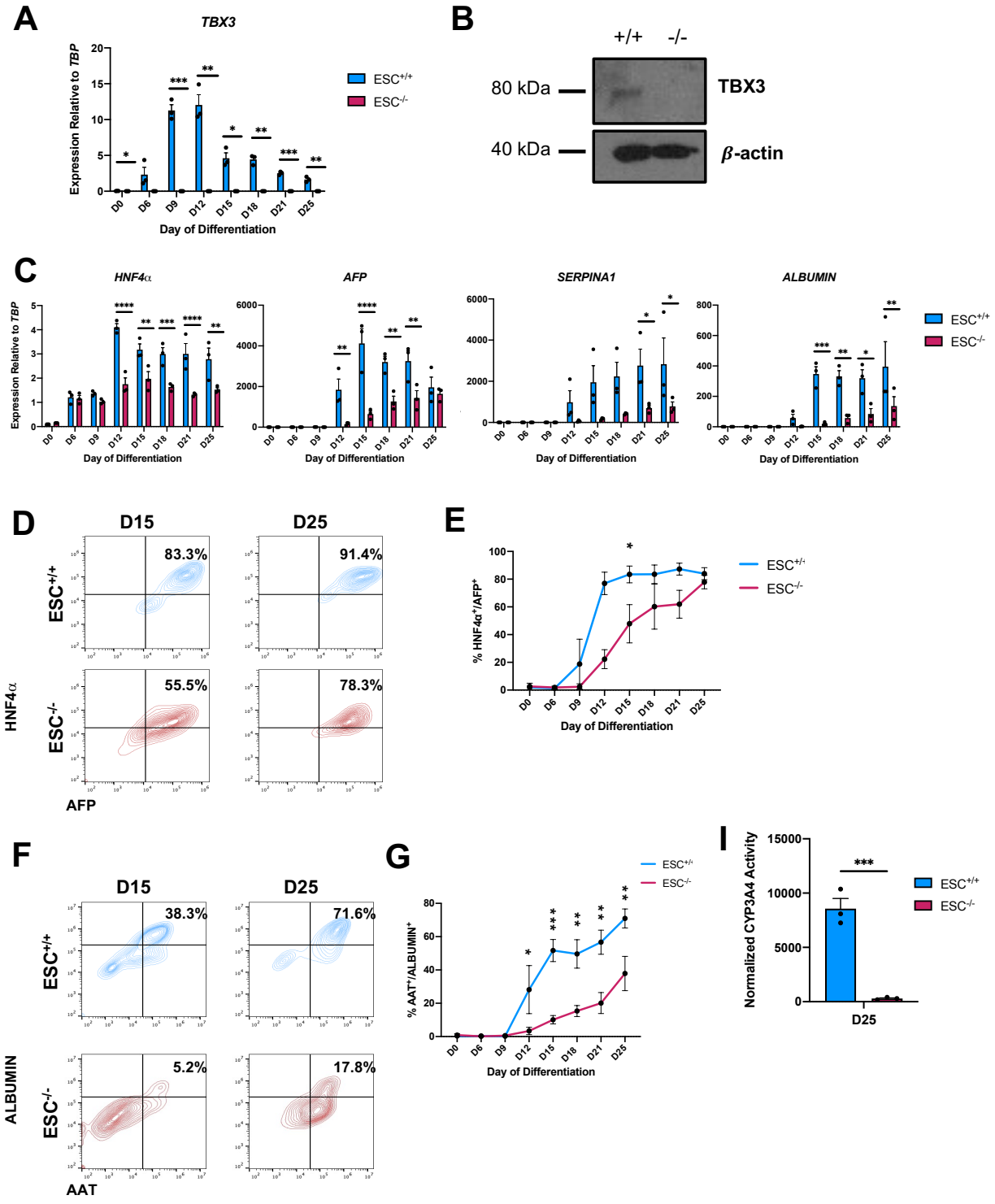


Figure S3: Loss of *TBX3* enhances pancreatic differentiation supporting Figure 2 and Figure 3

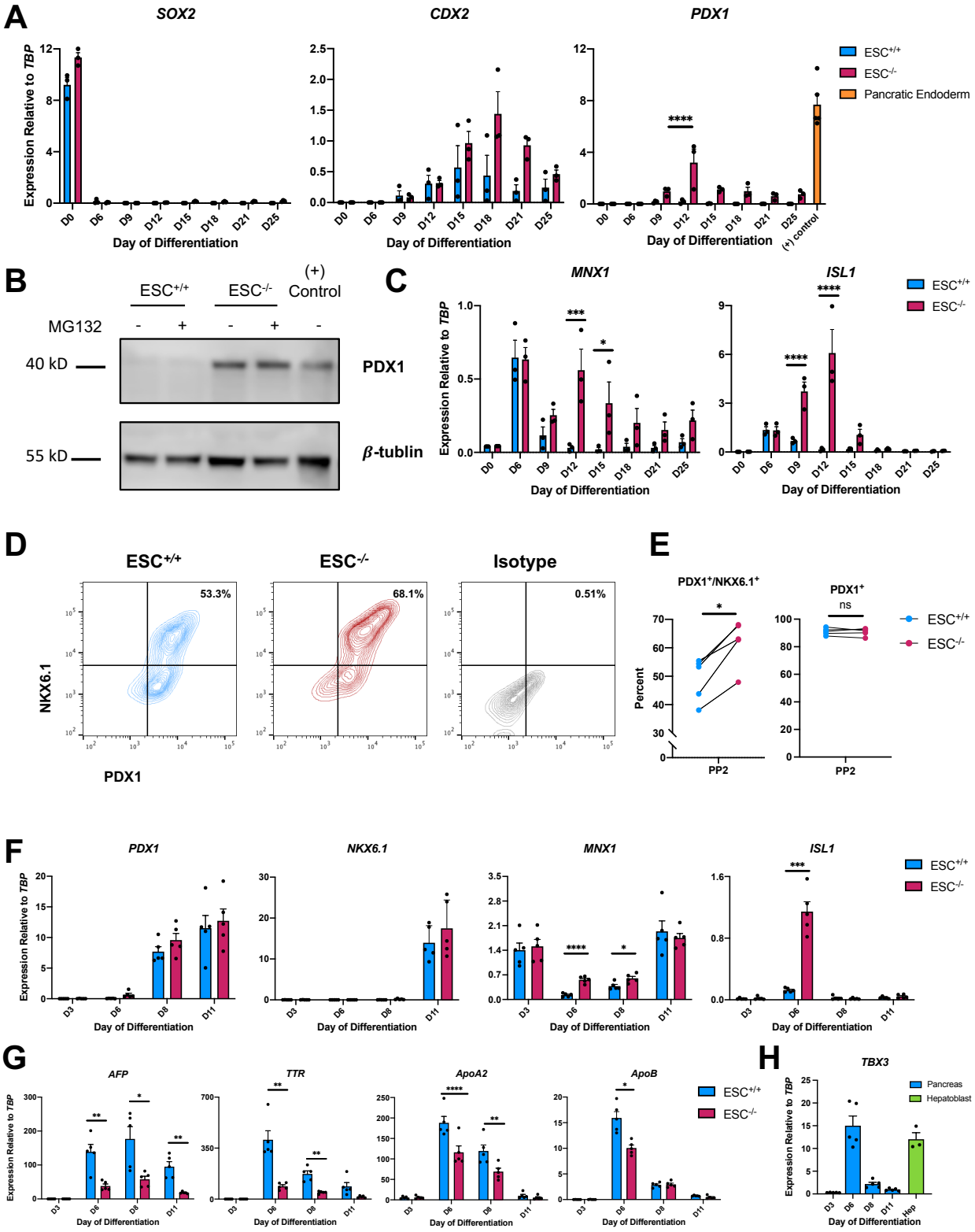


Figure S4: iPSC$^{+/+}$ cells are enriched for hepatic gene signature supporting Figure 4

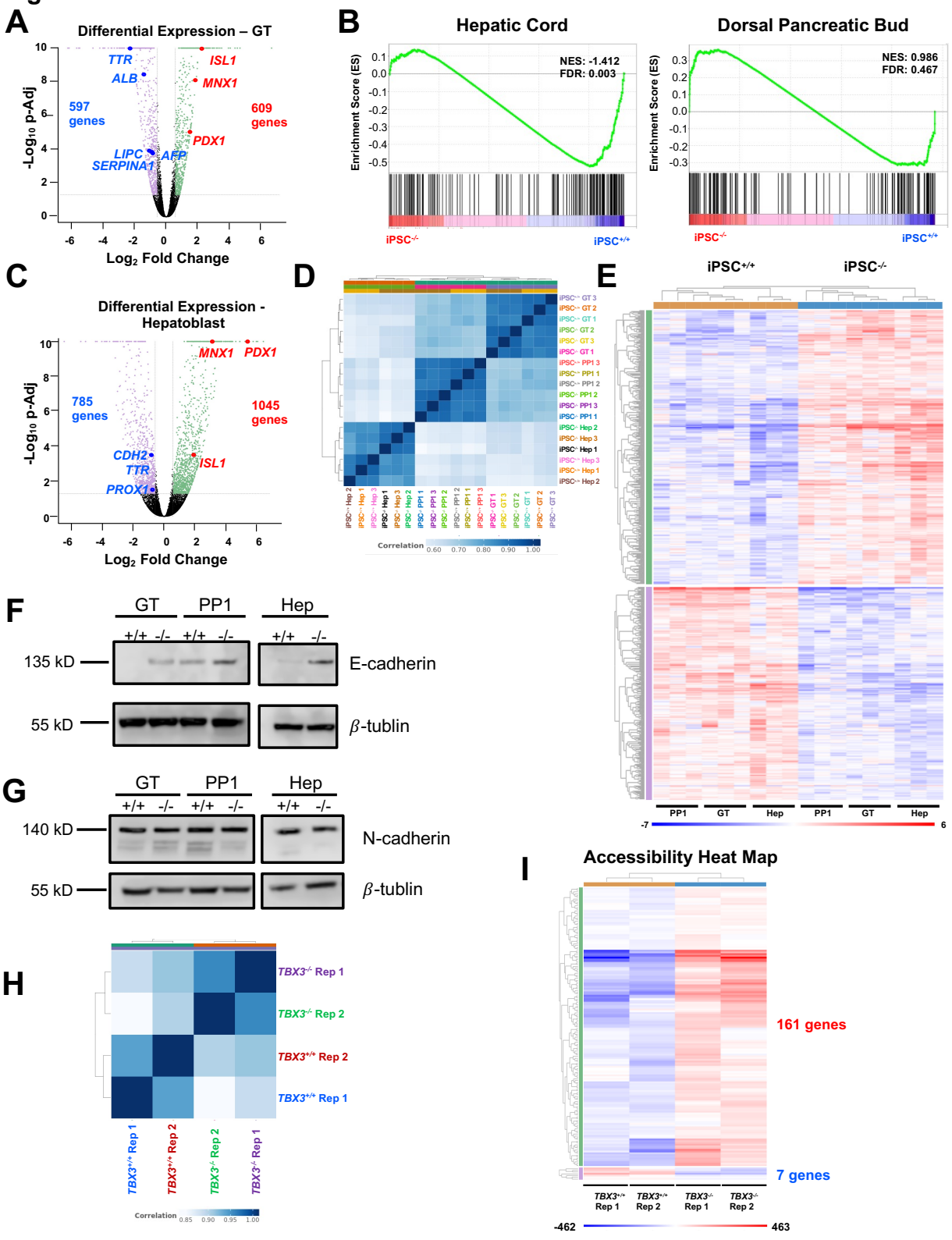


Figure S1. Characterization of *TBX3* knockout PSC lines all figures
(A) Sequencing of the region to verify deletion in both alleles of the *TBX3* locus in the iPSC$^{-/-}$ line. Bars represent locations of *TBX3* gRNA1 and *TBX3* gRNA2.
(B) Karyotype of iPSC$^{-/-}$ line.
(C) Sequencing of the region of the gRNA sequences to verify deletion in both alleles of endogenous *TBX3* locus in the ESC$^{-/-}$ line. Bars represent locations of *TBX3* gRNA1 and *TBX3* gRNA2.
(D) Karyotype of ESC$^{-/-}$ line.

Figure S2. Loss of *TBX3* impairs hepatocyte differentiation supporting Figure 1
(A) Time-course of *TBX3* expression during hepatocyte differentiation by qRT-PCR. (n =3 per time point, per cell line).
(B) Western blot of TBX3 protein in day 12 ESC$^{+/+}$ and ESC$^{-/-}$ hepatoblasts.
(C) Time-course of hepatoblast (*HNF4*α and *AFP*) and hepatocyte (*SERPINA1* and *ALBUMIN*) markers during hepatocyte differentiation by qRT-PCR (n = 3 separate experiments per time, point per cell line).
(D) Representative example of HNF4α and AFP expression at day 15 and Day 25 by flow cytometry.
(E) Time-course of percentage of HNF4α^+/AFP$^+$ cells by flow cytometry (n=3 per time point, per cell line).
(F) Representative example of AAT ad ALBUMIN expression at day 15 and Day 25 by flow cytometry.
(G) Time-course of percentage of AAT$^+$/ALBUMIN$^+$ cells by flow cytometry (n=3 per time point, per cell line).
(J) Rifampicin-induced CYP3A4 activity in ESC$^{+/+}$ and ESC$^{-/-}$ hepatocytes at day 25 (n = 3 per cell line). For all statistical analysis, *p<0.05, **p<0.01, ***p<0.001, and **** p≤ 0.0001.

Figure S3. Loss of *TBX3* enhances pancreatic differentiation supporting Figure 2 and Figure 3
(A) Time-course of anterior gut tube (*SOX2*), pancreatic endoderm (*PDX1*), and posterior gut tube (*CDX2*) markers during hepatocyte differentiation by qRT-PCR (n = 3 per time point, per cell line). PDX1 expression in ESC$^{+/+}$ differentiated to pancreatic endoderm for (+) control (n = 5).
(B) Western blot of PDX1 protein in day 15 ESC$^{+/+}$ and ESC$^{-/-}$ immature hepatocytes with or without MG132 and pancreatic endoderm for (+) control.
(C) Time-course of early pancreatic markers *ISL1* and *MNX1* during hepatocyte differentiation by qRT-PCR (n = 3 per time point, per cell line).
(D) Representative example of PDX1 and NKX6.1 expression in PP2 cells from pancreas differentiation.
(E) Quantification of the percentage of PDX1$^+$/NKX6.1$^+$ and PDX1$^+$ PP2 cells (n = 5 per cell line).
(F) Time-course of *PDX1*, *NKX6.1*, *ISL1*, and *MNX1* during pancreatic differentiation by qRT-PCR (n = 5 per time point, per cell line).
(G) Time-course of hepatoblast markers *AFP, TTR*, *ApoA2*, and *ApoB* during pancreatic differentiation by qRT-PCR (n = 5 per time point, per cell line).
(H) Time-course of *TBX3* expression during pancreatic differentiation (n = 5) by qRT-PCR and day 12 hepatoblasts (n = 3) for comparison. For all statistical analysis, *p<0.05, **p<0.01, ***p<0.001, and **** p≤ 0.0001.

Figure S4. iPSC$^{+/+}$ cells are enriched for hepatic gene signature supporting Figure 4
(A) Volcano plot of up- and downregulated genes in iPSC$^{-/-}$ versus iPSC$^{+/+}$ GT cells. p-Adj = 0.05, fold change: ≥ 1.5 and ≤ -1.5.
(B) GSEA analysis comparing normalized gene expression of samples examined in (A) to genes enriched in human fetal hepatic cords and dorsal pancreatic bud.
(C) Volcano plot of up- and downregulated genes in iPSC$^{-/-}$ versus iPSC$^{+/+}$ hepatoblast cells. p-Adj = 0.05, fold change: ≥ 1.5 and ≤ -1.5.
(D) Sample correlation of heatmap correlating RNA-seq samples of same replicate group, developmental stage, tissue type, and genotype.
(E) Heat map of genes identified by covariate analysis that are commonly up- and downregulated in iPSC$^{-/-}$ versus iPSC$^{+/+}$ cells at different developmental stages and tissues types.
(F) Western blot of E-cadherin protein in iPSC$^{+/+}$ and iPSC$^{-/-}$ cells.
(G) Western blot of N-cadherin protein in iPSC$^{+/+}$ and iPSC$^{-/-}$ cells.
(H) Sample correlation of heatmap correlating ATAC-seq samples of same replicate group and genotype.

(I) Upregulated genes from (E) were analyzed for differential accessibility by ATAC-seq. Heat map showing differentially accessible genes only (161 genes were more accessible while only 7 were less accessible).
GT = day 6 gut tube, PP1 = day 8 pancreatic progenitor 1, and Hep = day 12 hepatoblast

Table S1: *TBX3* mutant cells lines supporting all figures

Name	Genetic Background	Allele 1	Allele 2
$iPSC^{+/+}$	CHOPi004-A	Wild type	Wild type
$iPSC^{-/-}$	CHOPi004-A	6.9 Kb deletion	6.9 Kb deletion
$ESC^{+/+}$	Mel1	Wild type	Wild type
$ESC^{-/-}$	Mel1	6.9 Kb deletion	6.9 Kb deletion

Table S2: Primary and secondary antibodies supporting Figures 1-3

Antibody	Species	Dilution and Application	Source and Catalog Number
TBX3	Goat Polyclonal IgG	1:200 for Western Blot	Santa Cruz #sc-17871
β-actin	Mouse	1:2000 for Western Blot	Sigma #A1978
HNF4a	Rabbit Monoclonal IgG	1:100 for flow cytometry	Cell signaling #3113
AFP	Mouse Monoclonal IgG1	1:400 for flow cytometry	R&D Systems #MAB1368
AAT	Mouse Monoclonal IgG1	1:100 for flow cytometry	Santa Cruz #sc-59438
Albumin	Rabbit Polyclonal IgG	1:1000 for flow cytometry	DakoCytomation #A0001
PDX1-biotinylated	Goat Polyclonal IgG	1:50 for flow cytometry 1:500 for Western blot	R&D Systems #BAF2419
β-tubulin	Rabbit	1:1000 for Western blot	
NKX6.1	Mouse Monoclonal IgG1	1:250 for flow cytometry	DSHB #F55A10
E-Cadherin	Mouse	1:1000 for Western blot	
N-Cadherin	Rabbit	1:1000 for Western blot	
Mouse IgG1 - Alexa488	Goat	1:400 for flow cytometry	Jackson Immunoresearch # 115-545-205
Rabbit IgG - Alexa 647	Goat	1:400 for flow cytometry	Jackson Immunoresearch #111-605-144
Mouse IgG1 - 647	Donkey	1:400 for flow cytometry	Jackson Immunoresearch # 715-606-151
Streptavidin Pacific Blue		1:400 for flow cytometry	Invitrogen #S-11222
Goat-HRP	Rabbit	1:1000 for Western Blot	R&D Systems #HAF017
Mouse-HRP	Goat	1:5000 for Western Blot	Biorad #170-6516
Rabbit-HRP	Goat	1:5000 for Western Blot	Biorad #170-6515

Table of primary and secondary antibodies used for flow cytometry and western blot.

Table S3: Table of forward and reverse primers used for qRT-PCR supporting Figures 1-4.

Gene	Forward Primer	Reverse Primer
TBX3	5'-TGAGATGTTCTGGGCTGG-3'	5'-CTTACCAGCCACCATCCA-3'
AFP	5'-GTTTGTTCAAGAAGCCACTTAC-3'	5'-CACCCTGAAGACTGTTCATC-3'
HNF4α	5-TCCAACCCAACCTCATCCTCCTTCTT-3'	5'TCCTCTCCACTCCAAGTTCCTGTT-3'
SERPINA1	5' - AGGGCCTGAAGCTAGTGGATAGT-3'	5'-TCTGTTTCTTGGCCTCTTCGGTGT-3'
ALBUMIN	5'-GTGAAACACAAGCCCAAGGCAACA-3'	5'-TCCTCGGCAAAGCAGGTCTC-3'
SOX2	5'-CCATCACCCACAGCAAAT-3'	5'-AGTCCAGGATCTCTCTCATAA-3'
PDX1	5'-GGAGCTGGCTGTCATGTTG-3'	5'-CACTTCATGCGGCGGTTT-3'
CDX2	5'-AAGGACGTGAGCATGTACCCTAGC-3'	5'-CACGTGGTAACCGCCGTAGTC-3'
ISL1	5--CAGAAGGAGGACCGGGCTCTAAT-3'	5'-GACTGGCTACCATGCTGTTAGGTGTAT-3'
MNX1	5'-AGAAGGCGGAAACCCACAGTGTAA-3'	5'-CCCAGAGACGTAAGCATAAACCCT-3'
NKX6.1	5'-AAGAAGCACGCTGCCGAGATG-3'	5'-CCGAGTTGGGATCCAGAGGCTTATT-3'
TTR	5'-ATGGGCTCACAACTGAGGAGGAAT-3'	5'-AGATGCCAAGTGCCTTCCAGTAAGA-3'
APOA2	5'-ATGTGTGGAGAGCCTGGTTTCTCA-3'	5'-AAGCTCTGGGCTCTTGACCTTCT-3'
APOB	5'-ACTCACATCCTCCAGTGGCTGAAA-3'	5'-CGCTGATCCCTCGCCATGTT-3'
CDH1	5'-TTCCCTCGACACCCGATTCAAAGT -3'	5'-GAGTCCCAGGCGTAGACCAAGAAAT-3'
CDH2	5'-TGAGAGCAGTGAGCCTGCAGATTT-3'	5'-TGAGAGCAGTGAGCCTGCAGATTT-3'
NEUROG3	5'GGAGTCGGCGAAAGAAG-3'	5'-ATGTAGTTGTGGGCGAAG-3'
NKX2.2	5'-GAATGTTTGCGCAGCTTCGCTTCT-3'	5'-CCACTTGCTTTAGAAGACGGCTGA-3'
TBP	5'- TTGCTGAGAAGAGTGTGCTGGAGATG-3	5'-CGTAAGGTGGCAGGCTGTTGTT-3'

Supplemental Experimental Procedures

PSC lines

PSCs were cultured in an environment of 5% CO_2, 5% O_2, and 90% N_2 on 0.1% gelatin (Sigma) and irradiated mouse embryonic fibroblasts. They were grown in human embryonic stem cell (hES) medium consisting of DMEM/F12 (Corning) with 15% Knockout Serum Replacement (Gibco), 2mM L-glutamine (Corning), 1x Non-Essential Amino Acids (Gibco), 1x Penicillin/Streptomycin (Corning), 0.1 mM 2 β-mercaptoethanol (Gibco), and 10 ng/mL bFGF (R&D Systems). Medium was changed daily. PSCs were grown to 80% confluency and split at a 1:12 ratio using TryplE (Gibco) dissociation regent. Cells were replated in hES medium with 5μM ROCK inhibitor Y-27632 dihydrochloride (Tocris).

Hepatocyte Differentiation

The hepatocyte differentiation previously described (Ogawa et al., 2013) was modified as follows. PSCs were split onto 1:3 Matrigel (Corning) coated 6-well plates and cultured until they reached 80-90% confluency before starting Day 0 of the differentiation. Cells were differentiated as a monolayer in 5% O_2, 5% CO_2, and 90% N_2. Cells were cultured for 24 hours (Day 0) in RPMI medium (Corning) with 2mM L-glutamine, 50 μg/ml stabilized ascorbic acid (SAA) (Wako), 4.5×10^{-4} M MTG (Sigma), 100ng/mL Activin A (R&D Systems), and 2μM CHIR (Tocris). Cells were then cultured for 24 hours in RPMI medium with 2mM L-glutamine, 50 μg/ml SAA, 4.5×10^{-4} M MTG, 100ng/mL Activin A, and 5 ng/mL bFGF (R&D Systems). From Days 2-6, cells were cultured serum free differentiation (SFD) medium, 2mM L-glutamine, 50 μg/ml SAA, 4.5×10^{-4} M MTG, 100ng/mL Activin A, and 5 ng/mL bFGF. SFD medium consists of Iscove's DMEM (Corning) supplemented with 25% Ham's/F12 (Corning), 0.5% N2 supplement (Gibco), 1% B27 without retinoic acid (RA) supplement (Gibco), and 1% bovine serum albumin (BSA) (Sigma). At Day 6, purity of the definitive endoderm (DE) was assessed by flow cytometry with DE being 95-98% $CXCR4^+/C\text{-}kit^+$. Cells were then dissociated using TryplE for 2 minutes and replated with Y-27632 dihydrochloride onto 1:3 Matrigel coated 6-well plates at a density of 3.5×10^5 cells per well. On Day 7, the medium was replaced with hepatoblast specification consisting of H16 medium supplemented with supplemented 2mM L-glutamine, 50 μg/ml SAA, 4.5×10^{-4} M MTG, 40 ng/mL bFGF, and 50 ng/mL BMP4 (R&D Systems). H16 base medium consists of DMEM Low glucose (Gibco) supplemented with 25% Ham's/F12, 1% B27 with RA (Gibco), and 1% BSA. Medium was changed every other day. On Day 13 medium was replaced with hepatocyte maturation medium A consisting of H16 medium supplemented with 2mM L-glutamine, 50 μg/ml SAA, 4.5×10^{-4} M MTG, 20 ng/mL Hepatocyte Growth Factor (HGF) (R&D Systems), 20 ng/mL Oncostatin-M (OSM) (R&D Systems), and 40 ng/ml dexamethasone (Dex) (Sigma). Medium was changed every other day. On Day 21, cells were transferred to an environment of 20% O_2 and 5% CO_2, and 90% N_2. Medium was replaced with hepatoblast maturation medium B consisting of H21 medium supplemented with 2mM L-glutamine, 50 μg/ml SAA, 4.5×10^{-4} M MTG, 20 ng/mL Hepatocyte Growth Factor (HGF) (R&D Systems), 20 ng/mL Oncostatin-M (OSM) (R&D Systems), and 40 ng/ml dexamethasone (Dex) (Sigma). H21 medium is the same as H16 except using DMEM high glucose (Gibco) as the base medium. Medium was changed every other day until Day 25. Cells were harvested on Day 0, Day 6, Day 9, Day 12, Day 15, Day 18, Day 21 and Day 25 for analysis.

Cytochrome P450 activity assay

The cytochrome p450 activity assay was performed using P450-GloTM CYP3A4 Luciferin-IPA kit (Promega, V9002). On Day 22 of the hepatocyte differentiation, cells were cultured with 25 μM Rifampicin (Sigma) or with DMSO (Tocris) as a vehicle control. To confirm that the induced activity was specific for CYP3A4 enzyme, inhibition controls included the selective inhibitor ketoconazole at 1 μM (Sigma) in the presence of 25 μM Rifampicin. Net signal was calculated by subtracting background luminescence values (no-cell control) from Rifampicin or DMSO values. Luminescence values were normalized for $5x10^5$ cells to account for differences in cell numbers between wells.

Pancreatic β-cell Differentiation

The pancreatic β-cell differentiation previously described (Rezania et al., 2014) was modified as follows. PSCs were split onto 1:30 Matrigel coated 6-well plates and cultured until they reached 80-90% confluency before starting Day 0 of the differentiation. Cells were grown in 20% O_2 and 5% CO2, and 90% N_2, and differentiated as a monolayer. Cells were cultured for 24 hours in RPMI medium with 100ng/mL Activin A, and 3μM CHIR (Day 0). Cells were cultured for another 24 hours (Day 1) in RPMI medium with 100 ng/mL

Activin A, 0.3 μM CHIR, and 5 ng/mL bFGF. On Day 2, cells were differentiated to definitive endoderm (DE) with serum free differentiation (SFD) medium containing 100 ng/mL Activin A. At Day 3, purity of the DE was assessed as described above. From Day 3-5, cells were cultured in DMEM/F12 medium with 1% fetal bovine serum (FBS) (Gibco), 50μg/ml stabilized ascorbic acid (SAA), 1.25 mM IWP2 (Tocris), and 50 ng/ml FGF7 (R&D Systems). Medium was changed every day. On Day 6 and 7, cells were cultured in DMEM high glucose medium with 1% B27 without RA, 1X Glutamax (Gibco), 50 μg/ml SAA, 0.5% ITS-X (Gibco), 50 ng/mL FGF7, 0.5μM SANT1 (Sigma), 1μM RA (Sigma), 100nM LDN (Tocris), and 500nM Phorbol 12-myristate 13-acetate (Tocris). Medium was changed daily. From Day 8-10, cells were cultured in DMEM high glucose medium with 1% B27 without RA, 1X Glutamax, 50 μg/ml SAA, 0.5% ITS-X, 2 ng/mL FGF7, 0.5μM SANT1, 0.1μM RA, 100nM LDN, and 250nM Phorbol 12-myristate 13-acetate. Medium was changed daily. Cells were harvested on Day 0, Day 3, Day 6, Day 8, and Day 11 for analysis.

RNA-seq Genomic Analysis

Cells were harvested on day 12 of the hepatocyte differentiation by dissociation with 0.25% Trypsin/EDTA (Gibco) for 4 minutes. RNA was extracted as described below and sent to Genewiz (South Plainfield, NJ) for library preparation and sequencing. Cells were harvested on day 6 and day 8 of the pancreas differentiation as described above. RNA was extracted and sent to the Center for Applied Genomics Biorepository Core at the Children's Hospital of Philadelphia for library preparation and sequencing. Data was analyzed by ROSALIND (https://rosalind.onramp.bio/), with a HyperScale architecture developed by ROSALIND, Inc. (San Diego, CA). Reads were trimmed using cutadapt (Martin, 2011). Quality scores were assessed using FastQC[2]. Reads were aligned to the Homo sapiens genome build hg19 using STAR (Dobin et al., 2013). Individual sample reads were quantified using HTseq (Anders et al., 2015) and normalized via Relative Log Expression (RLE) using DESeq2 R library (Love et al., 2014). Read Distribution percentages, violin plots, identity heatmaps, and sample MDS plots were generated as part of the QC step using RSeQC (Wang et al., 2012). Deseq2 was also used to calculate fold changes and p-values and perform optional covariate correction. Clustering of genes for the final heatmap of differentially expressed genes was done using the PAM (Partitioning Around Medoids) method using the fpc R library. Hypergeometric distribution was used to analyze the enrichment of pathways, gene ontology, domain structure, and other ontologies. The topGO R library was used to determine local similarities and dependencies between GO terms in order to perform Elim pruning corrections. Several database sources were referenced for enrichment analysis, including Interpro (Mitchell et al., 2019), NCBI (Geer et al., 2009), KEGG (Kanehisa and Goto, 2000; Kanehisa et al., 2017, 2019), MSigDB (Liberzon et al., 2011; Subramanian et al., 2005), REACTOME (Fabregat et al., 2018), WikiPathways (Slenter et al., 2018). Enrichment was calculated relative to a set of background genes relevant for the experiment. Functional enrichment analysis of pathways, gene ontology, domain structure and other ontologies was performed using HOMER (Heinz et al., 2010). Gene set enrichment analysis was performed (Mootha et al., 2003; Subramanian et al., 2005) using a gene list comparing gene expression in dorsal pancreatic bud and hepatic cord tissues dissected from human embryos using laser capture technology (Jennings et al., 2017). The gene list was sorted by a p-value of <0.05, then by fold change. The top 200 upregulated genes were the most enriched in the dorsal pancreatic bud, and the top 200 downregulated genes were most enriched in the hepatic cord. GSEA was performed on these subsets using the "gene_set" permutation and an FDR cutoff of 5%.

ATAC-seq Genomic Analysis

Cells were harvested on day 12 of the hepatocyte differentiation by dissociation with 0.25% Trypsin/EDTA (Gibco) for 4 minutes. $1x10^6$ cells were frozen in 90% FBS (Gibco) and 10% DMSO (Tocris), and sent to Genewiz (South Plainfield, NJ) for tagmentation, library preparation, and sequencing. Data was analyzed by ROSALIND (https://rosalind.onramp.bio/), with a HyperScale architecture developed by ROSALIND, Inc. (San Diego, CA). Reads were trimmed using cutadapt (Martin, 2011). Quality scores were assessed using FastQC. Reads were aligned to the Homo sapiens genome build hg19 using bowtie2 (Langmead and Salzberg, 2012). Per-sample quality assessment plots were generated with HOMER (Heinz et al., 2010) and Mosaics (Kuan et al., 2011). Peaks were called using MACS2 (Zhang et al., 2008). Peak overlaps and differential accessibility were calculated using the DiffBind R library (Ross-Innes et al., 2012). Differential accessibility was calculated at gene promoter sites. Read distribution percentages, identity heatmaps, and FRiP plots were generated as part of the QC step using ChIPQC R library (Carroll and Stark, 2014) and

HOMER. HOMER was also used to generate known and de novo motifs and perform functional enrichment analysis of pathways, gene ontology, domain structure and other ontologies.

RNA extraction, cDNA synthesis, and qRT-PCR

Cells were harvested by dissociation with 0.25% Trypsin/EDTA for 4 minutes. Cellular RNA was isolated using the PureLink RNA Micro Scale Kit (Invitrogen) following the manufacturer's protocol. Random hexamers (Invitrogen) were used with the SuperScript III Reverse Transcriptase System (Invitrogen) to synthesize cDNA from 500 ng of extracted RNA. qRT-PCR reactions were done in triplicate on a Roche LightCycler 480 II using SYBR Select Master Mix (Applied Biosystems). Serial dilutions of H9 embryonic stem cell genomic DNA were used to generate a standard curve, and *TBP* (Veazey and Golding, 2011) was used as a house keeping gene to determine relative gene expression levels. The primers that were used in this study can be found in Table S3.

Western Blot

Cells were harvested with 1.5x Laemmli buffer (75 mM Tris-HCl3, 15% glycerol, 3% SDS, 3.75 mM EDTA, and 200 mM NaF)**.** Cell lysates were boiled at 95^{o}C for 20 minutes and protein was quantified using Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Samples were aliquoted and diluted with 4x Laemmli buffer. 20 µg of protein from each sample were run on a 4%-12% Bis-Tris SDS-polyacrylamide gel (Invitrogen) and transferred to a 0.45 µm pore size nitrocellulose membrane (Thermo Fisher Scientific). The membrane was stained with Ponceau S (Sigma-Aldrich) to ensure successful transfer. The membrane was washed with 1X TBS (Bio-Rad) with 0.1% Tween-20 (Sigma) (TBST), then blocked in 5% nonfat dry milk in TBST for 1 hour at room temperature. The membrane was probed with primary antibody diluted in 5% nonfat dry milk in TBST overnight at 4°C. The membrane was washed with TBST and placed in a horseradish peroxidase conjugated secondary antibody diluted in 5% nonfat dry milk in 1X TBST for 1 hour at room temperature. Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific) was added to the membrane and exposed onto HyBlot CL autoradiography film (Denville Scientific) for visualization the TBX3 blots. Super Signal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific) was added to the membrane and visualized using a C-DiGit Blot Scanner and ImageStudioDigits Software Version 5.2.5 (LiCor) for the PDX1, E-Cadherin, and N-Cadherin blots. A list of antibodies used can be found in Table S2.

Flow Cytometry

Cells were harvested by dissociation with 0.25% Trypsin/EDTA for 3 to 5 minutes. For intracellular staining, cells were fixed with 1.6% paraformaldehyde (Electron Microscopy Science) for 30 minutes at 37°C. Cells were washed in 1X PBS (Corning), then permeabilized and stained in 1X saponin buffer (Biolegend). Primary and secondary antibodies were diluted to the appropriate concentrations in saponin buffer and cells were stained for thirty minutes each at room temperature. Following staining, cells were washed in saponin, and resuspended in FACS buffer (1X PBS (Corning) with 0.1% BSA (Sigma) and 0.1% sodium azide (Sigma)). For extracellular staining, conjugated primary antibodies were diluted to the appropriate concentration in FACS buffer and cells were stained for fifteen minutes at room temperate. Following the staining, cells were washed and resuspended in FACS buffer. All samples were analyzed on a CytoFLEX V2-Br-R2 flow cytometer (Beckman Coulter Life Sciences) and FlowJo Version 10.6.2 (Beckton Dickenson) software program. Cell sorting was carried out on a single cell suspension using BD FACSAria II (Becton Dickinson). A list of antibodies used can be found in Table S2.

Supplemental References

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq-A Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Carroll, T.S., and Stark, R. (2014). Assessing ChIP-seq sample quality with ChIPQC. 1–21.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: Ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., Haw, R., Jassal, B.,

Korninger, F., May, B., et al. (2018). The Reactome Pathway Knowledgebase. Nucleic Acids Res. *46*, D649–D655.
Geer, L.Y., Marchler-Bauer, A., Geer, R.C., Han, L., He, J., He, S., Liu, C., Shi, W., and Bryant, S.H. (2009). The NCBI BioSystems database. Nucleic Acids Res. *38*, 492–496.
Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol. Cell *38*, 576–589.
Jennings, R.E., Berry, A.A., Gerrard, D.T., Wearne, S.J., Strutt, J., Withey, S., Chhatriwala, M., Piper Hanley, K., Vallier, L., Bobola, N., et al. (2017). Laser Capture and Deep Sequencing Reveals the Transcriptomic Programmes Regulating the Onset of Pancreas and Liver Differentiation in Human Embryos. Stem Cell Reports *9*, 1387–1394.
Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. *28*, 27–30.
Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. *45*, D353–D361.
Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., and Tanabe, M. (2019). New approach for understanding genome variations in KEGG. Nucleic Acids Res. *47*, D590–D595.
Kuan, P.F., Chung, D., Pan, G., Thomson, J.A., Stewart, R., and Keleş, S. (2011). A statistical framework for the analysis of ChIP-Seq data. J. Am. Stat. Assoc. *106*, 891–903.
Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.
Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739–1740.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 1–21.
Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal *17*, 10–12.
Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.Y., El-Gebali, S., Fraser, M.I., et al. (2019). InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res. *47*, D351–D360.
Mootha, V.K., Lindgren, C.M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., et al. (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. *34*, 267–273.
Ogawa, S., Surapisitchat, J., Virtanen, C., Ogawa, M., Niapour, M., Sugamori, K.S., Wang, S., Tamblyn, L., Guillemette, C., Hoffmann, E., et al. (2013). Three-dimensional culture and cAMP signaling promote the maturation of human pluripotent stem cell-derived hepatocytes. Development *140*, 3285–3296.
Rezania, A., Bruin, J.E., Arora, P., Rubin, A., Batushansky, I., Asadi, A., O'Dwyer, S., Quiskamp, N., Mojibian, M., Albrecht, T., et al. (2014). Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. Nat. Biotechnol. *32*, 1121–1133.
Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389–393.
Slenter, D.N., Kutmon, M., Hanspers, K., Riutta, A., Windsor, J., Nunes, N., Mélius, J., Cirillo, E., Coort, S.L., DIgles, D., et al. (2018). WikiPathways: A multifaceted pathway database bridging metabolomics to other omics research. Nucleic Acids Res. *46*, D661–D667.
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.
Veazey, K.J., and Golding, M.C. (2011). Selection of stable reference genes for quantitative RT-PCR comparisons of mouse embryonic and Extra-Embryonic stem cells. PLoS One *6*.
Wang, L., Wang, S., and Li, W. (2012). RSeQC: Quality control of RNA-seq experiments. Bioinformatics *28*, 2184–2185.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.