

Supplementary Materials for

Unbiased transcription factor CRISPR screen identifies ZNF800 as master repressor of enteroendocrine differentiation

Lin Lin *et al.*

Corresponding authors: Lin Lin, l.lin@hubrecht.eu; Hans Clevers, h.clevers@hubrecht.eu

Science **382**, 451 (2023)
DOI: 10.1126/science.adl2246

The PDF file includes:

Materials and Methods
Figs. S1 to S22
References

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S8
MDAR Reproducibility Checklist

Materials and Methods

Experimental model and subject details

One human small intestine organoid line (male) (18), and two human colon organoid lines (both donors are female) (58) were recruited in this study. Tissues from human ileum and colon were obtained from the UMC Utrecht and The Diakonessen Hospital Utrecht with informed consent of each patient. The study was approved by the ethical committee and was in accordance with the Declaration of Helsinki and according to Dutch law. This study is compliant with all relevant ethical regulations regarding research involving human participants.

Organoid culture

Human small intestinal cells and colon cells were isolated, processed and cultured as described previously (14). Organoids were expanded in the expansion medium and subsequently differentiated as described previously (19). In brief, ileum and colon organoids were split in expansion medium on average every 10 days and in optimal maturation medium on average every 12 days. The composition of optimal maturation medium is: R-Spondin conditioned medium (10%), Noggin conditioned medium (2%, U-Protein Express), B-27 Supplement (1%, ThermoFisher), n-Acetyl Cysteine (1,25 mM, Sigma), Nicotinamide (10 mM, Sigma), Wnt surrogate (0.015nM, U-Protein Express), human EGF (50 ng/mL, PeproTech), human IL-22(2 ng/ml, PeproTech), A83-01 (500 nM, Tocris), Rho kinase inhibitor Y-27632 dihydrochloride (10 mM, Abmole Bioscience) and Primocin (100 mg/mL, InvivoGen) supplemented in basic culture medium consisting of Advanced Dulbecco's modified Eagle's medium (DMEM)/F12 with B27, Glutamax, HEPES, penicillin/streptomycin (Thermo Fisher). For passaging the organoids, organoids were removed from the Cultrex Basement Membrane Extract (BME, Growth Factor Reduced, Type 2, R&D Systems), washed with AdDMEM/F12 (Thermo Fisher) and mechanically dissociated into small fragments using 1mL pipette tips. Cell clumps were washed once to remove dead cells and replated in fresh BME (R&D Systems). Organoid clones carrying the overexpression construct (FLAG-ZNF800-P2A-tagBFP; FLAG-ARX-P2A-tagBFP; FLAG-PAX4-P2A-tagBFP) underwent expansion, patterning and maturation stages without doxycycline treatment. Once organoids were cultured and stably maintained in maturation medium, 1 μ g/mL doxycycline (Sigma-Aldrich) was added to the culture right after splitting for further maturation and examination. The parallel condition without adding doxycycline was always used as a control.

CRISPR engineering in human intestinal and colon organoids

For fluorescent reporter construction, CHGA-IRES-iRFP670 homology-directed repair (HDR) knock-in construct was established as previously described (19). For gene knockout assay, sgRNAs were designed to target the respective genes in the exon, and cloned into a backbone construct p2TU6sg2XBbsIFE (Addgene #71485) (59) using NEBuilder HiFi DNA Assembly (NEB). The mixture of sgRNA plasmids, spCas9 plasmids together with HDR constructs were introduced into the cells using NEPA electroporation system (NEPAGENE) (60). Organoid clones derived from single cell expansion were picked manually and expanded for genotyping. All organoid lines with knock-in reporters used in this study are clonal lines with homozygous knock-in reporters. All knockout organoid lines produced in this study are clonal organoids with

homozygous knockout. The sgRNA sequences and genotyping primers are listed in table S8. For overexpression of ZNF800, a plasmid containing full-length cDNA of ZNF800 (pDONR223-ZNF800) was obtained from Addgene (#88573) (61). The cDNA of ZNF800 was amplified by primers flanking start codon and last codon sequence (table S8).

For overexpression of ARX and PAX4, first-strand cDNA was synthesized from total RNA using ProtoScript® II First Strand cDNA Synthesis Kit (NEB). Primers flanking start codon and last codon sequence of ARX and PAX4 were designed to amplify the full-length cDNA (table S8).

A plasmid with tol2 transposon system and doxycycline-inducible expression cassette driving ZNF800/ARX/PAX4 was established by assembling digested p2T fragment, tight TRE promoter, cDNA of ZNF800/ARX/PAX4 and hPGK-PuroR-rTetR by NEBuilder HiFi DNA Assembly (NEB). The p2T fragment was obtained by digesting p2T-CAG-SpCas9-BlastR (Addgene #107190) (62) with Spel and SphI restriction enzymes (NEB). Tight TRE promoter and hPGK-PuroR-rTetR fragment is by PCR amplification from pCW-Cas9 (Addgene #50661) (63). The mixture of overexpression plasmids and tol2 transposase plasmids were introduced into the cells using NEPA electroporation system (NEPAGENE) (60). pDONR223-ZNF800 was a gift from Marc Vidal (Addgene plasmid # 88573 ; <http://n2t.net/addgene:88573> ; RRID:Addgene_88573). pCW-Cas9 was a gift from Eric Lander & David Sabatini (Addgene plasmid # 50661 ; <http://n2t.net/addgene:50661> ; RRID:Addgene_50661). sgBbsI (p2Tol-U6-2xBbsI-sgRNA-HygR) was a gift from Richard Sherwood (Addgene plasmid # 71485 ; <http://n2t.net/addgene:71485> ; RRID:Addgene_71485). p2T-CAG-SpCas9-BlastR was a gift from Richard Sherwood (Addgene plasmid # 107190 ; <http://n2t.net/addgene:107190> ; RRID:Addgene_107190).

For sgRNA library of CRISPR screen, sgRNA sequences were designed and purchased as pooled oligos from CustomArray, Inc (GenScript, WA) as previously described (22) (table S8). Oligos were cloned into the lentiCRISPR-v2-FE backbone (A gift from Prof. Richard Sherwood) by using NEBuilder HiFi DNA Assembly (NEB). Assembly mixture was subsequently transformed into NEB Stable Competent *E.coli* (High Efficiency) (NEB) according to the manufacturer's directions. Following recovery, a small dilution series were plated to assess transformation efficiency and the remainder was grown in liquid culture in LB medium overnight at 30 °C, followed by maxiprep with NucleoBond Xtra Maxi kit (Macherey-Nagel).

Lentivirus production and quantification

HEK293T cells (RRID:CVCL_0063) were cultured in Dulbecco's Modified Eagle Medium (DMEM; Thermo Fisher) supplemented with 10% v/v FBS (Sigma-Aldrich). HEK293T cells were plated at a confluence of 1×10^5 /cm² in 150mm cell culture dishes (Falcon) before the day of transfection. Transfection mixture was prepared with TransIT-Lenti Transfection Reagent (Mirus) according to the manufacturer's directions. In briefly, two tubes of reagent mixture were set up separately: Tube A contains 4mL Opti-MEM (Thermo Fisher), 9.7 µg of pMDLg/pRRE, 6.5 µg pRSV/REV, 3.3 µg pVSV-G and 13 µg of sgRNA library; Tube B contains 4mL Opti-MEM (Thermo Fisher) and 98 µL of TransIT-Lenti Transfection Reagent (Mirus). Two tubes were mixed by adding Tube B to Tube A and incubated 20 min at room temperature. Cells were transfected by adding the mixture dropwise. Cells were then incubated at 37 °C for 8 hours and refreshed with full medium. After 24 hours, supernatants of transfected HEK293T cells were

collected and stored at 4 °C. Transfected HEK293T cells were refreshed with full medium. Same procedures were repeated after 48 hours of transfection, and supernatants of transfected HEK293T cells were collected at both 48 hours and 72 hours after transfection. All batches of supernatants containing lentiviral particles were first centrifuged at 300 g at 4 °C, and filtered through a 45 µm pore size filter. Following filtration, lentiviral particles were concentrated by using Lenti-X Concentrator (Takara Bio) according to the manufacturer's directions. Lentiviral particles were resuspended in organoid expansion medium and immediately aliquoted for subsequent quantification and storage at -80 °C.

For virus titration, genomic RNA from lentiviral particles was isolated from concentrated aliquots using Nucleospin RNA Virus Kit (Macherey-Nagel) and treated with DNaseI (Thermo Fisher Scientific) to remove residual plasmid. Quantitation was performed in a qRT-PCR reaction using Lenti-X qRT-PCR Titration Kit (Takara Bio) according to the manufacturer's directions, yielding an approximate concentration of 2×10^5 virus particles per µL of concentrated virus mixture.

TFome-wide CRISPR knockout screen

Human SI organoids were mechanically dispersed by pipetting several times, followed by TrypLE (Thermo Fisher Scientific) digestion and filtration to obtain single cells. Single cells were washed and resuspended in full expansion medium containing Rho kinase inhibitor (10 µM) and Polybrene (8 µg/mL) with a concentration of 100,000 cells per well of 48-well suspension plate. Concentrated virus mixture was added to the cells in suspension. Spin infection was performed at 500 g for 1 hour at 32 °C, followed by recovery for 6 hours at 37 °C. Cells were then collected by spinning at 3000 rpm for 5 min, plated with fresh BME (R&D Systems) and cultured in full expansion medium. Puromycin was added to select infected cells 2 days after infection at a final concentration of 0.5 µg/mL.

To achieve an optimal infection titration with a multiplicity of infection (MOI) smaller than 0.2, a panel of test infection experiments were performed with a series dilution of the concentrated virus mixture based on qRT-PCR quantification. The dilution value resulted in less than 20% survival of the input cells was chosen for the final library infection. 2×10^7 cells were used for library infection per biological replicate ($n=2$). Organoids with stable lentiviral integration were expanded in full expansion medium and passaged in a ratio of 1:10. A split of the organoids from each biological replicate were collected as control bulk populations at expansion stage. After passaging, organoids were replated for two-step differentiation as previously described (19). After patterning and maturation, organoids were passaged once in a ratio of 1:10 by mechanically breaking using 1mL pipette tips. Organoids were further cultured for another 12 days. A split of the differentiated organoids from each biological replicate were collected as control bulk populations at differentiation stage. The other differentiated organoids were removed from BME (R&D Systems), washed with ice-cold AdDMEM/F12 (Thermo Fisher) and digested with TrypLE (Thermo Fisher) at 37 °C. Single cells were collected by passing the dissociated organoids through a cell strainer (Falcon). Prior to FACS, cells were stained with DAPI (Sigma-Aldrich) to identify live cells. Cells were sorted based on their reporter intensity and collected directly into cell lysis buffer (gating strategy: Fig. S1A). In average, 6×10^6 cells per biological replicate were collected as triple-reporter negative controls; 3×10^6 MUC2+ cells per biological replicate were collected; 7×10^5 CHGA+ cells per biological replicate were collected.

CRISPR screen analysis

CRISPR screen sequencing libraries were constructed as previously described (22) and sequenced on a Nextseq2000. Raw sequencing data and mapped library read counts can be found at NCBI GEO DataSets: GSE229584. In brief, genomic DNA was isolated from different cell populations using DNeasy Blood & Tissue kit (QIAGEN) following the manufacturer's protocol. Two sequential PCR reactions were performed to amplify the gRNA fragments integrated in the genomic DNA, using primers flanking the gRNA protospacers containing staggered barcodes (PCR1) and primers with full-length Illumina sequencing adapters (PCR2). All primer sequences are listed in table S8.

Reads were demultiplexed requiring a perfect match to a designed 8-nt index (Illumina Nextera DNA adapter series) and up to 1 mismatch to a designed 3-nt to 9-nt staggered barcode residing at the beginning of the read. gRNAs were counted by requiring a perfect match to a designed 19-nt, 20-nt, or 21-nt gRNA in the library using *mageck count* from MAGeCK CRISPR screen analysis toolkit (0.5.9.5) (64). MAGeCK was run on the mapped sgRNA read counts with default parameters. In brief, *mageck mle* was performed between the bulk population at expansion and differentiation stages by using plasmid pool as baseline with the setting of *--permutation-round 5*; *mageck mle* was performed between the CHGA+ and MUC2+ cell population by using triple-reporter negative population as baseline with the setting of *--permutation-round 5*. Embedded in *mageck mle* function, β scores for each gene are estimated by maximizing the joint log-likelihood of observing all sgRNA read counts on all different samples (table S1 and S2). Statistical analysis (wald test p-value) was performed to examine the reproducibility of all 4 sgRNAs for each gene in all replicates (65). Positive hits were prioritized based on the β score distribution of each gene in these two types of comparisons (CHGA+ EECs vs Triple-reporter negative population; MUC2+ goblet cells vs Triple-reporter negative population) with a slope cutoff = 1, intercept cutoff = 0.6. The selection of genes for further validation was based on their gene expression values in human SI organoids (19). In addition, *mageck test* was performed between the CHGA+ cells and triple-reporter negative cell population to calculate the Log₂FoldChange of individual sgRNAs (table S3). For all MAGeCK analysis, read counts from 100 non-targeting control sgRNAs embedded in the library were used as a control for normalization.

Confocal Imaging

Live organoids with fluorescent reporters were imaged on a Leica SP8 Confocal microscope. Fluorescent images were processed for max projection of all z-stacks using ImageJ software.

Flow cytometry

Flowcytometry sorting of organoids for CRISPR screening was performed on BD Influx cell sorter and BD FACSJazz cell sorter (BD Bioscience). For cells subjected to gDNA isolation, Buffer AL from DNeasy Blood & Tissue kit (QIAGEN) was used as collection buffer. For cells subjected to RNA isolation, Buffer RA1 from NucleoSpin RNA Kit (Macherey-Nagel) was used as collection buffer. For cells subjected to analysis, we performed flowcytometry analysis with analysis buffer that is comprised of 5mM EDTA, 25mM HEPES and 1% BSA in DPBS (Thermo

Fisher). Organoid cells were stained with DAPI (Sigma-Aldrich) or propidium iodide (Thermo Fisher) to identify live cells. Flowcytometry analysis was performed with CytoFLEX benchtop flow cytometer (Beckman Coulter) and analyzed with FlowJo software.

RNA extraction and real-time qPCR

Organoids or cells collected by FACS were subjected to RNA isolation using NucleoSpin RNA Kit (Macherey-Nagel) following the manufacturer's protocol. 2000~4000 CHGA+ or MUC2+ cells were sorted per replicate for each reverse transcription reaction. Equivalent cell numbers from different conditions were used for each experiment. Reverse transcription reactions were performed with DNase-treated RNA using GoScript Reverse Transcriptase kit (Promega). Afterwards, cDNA was subjected to qPCR analysis using iQ SYBR Green supermix (BioRad) in CFX Connect Real-Time PCR machine (BioRad). For gene expression profiling, qPCR was performed with gene-specific qPCR primers. Ct readouts of each gene were first normalized to the housekeeping gene GAPDH (ΔCt), and the relative fold changes between experimental groups and control groups were calculated with the $2^{-\Delta\Delta Ct}$ method. All qPCR primers used in study are listed in table S8.

Transmission electron microscopy

The protocol used for processing organoids for transmission electron microscopy (TEM) was as previously described (19). In brief, organoids were fixed for 3 hours at room temperature with 1.5% glutaraldehyde in 0.067 M cacodylate and 1 % sucrose (pH 7.4), followed by incubation in 1% osmium tetroxide and 1.5% K₄Fe(CN)₆ in 0.1 M sodium cacodylate (pH 7.4) for 1 hour at 4 °C. Organoids were then rinsed with MilliQ water and dehydrated at room temperature in a graded ethanol series (70, 90, up to 100%) and embedded in Epon polymerized for 48 hours at 60 °C. Ultrathin sections of 60 nm were cut using a diamond knife (Diatome) on a Leica UC7 ultramicrotome, and transferred onto 50 Mesh copper grids covered with a Formvar and carbon film. Sections were post-stained with uranyl acetate and lead citrate. All TEM data were collected autonomously as virtual nanoscopy slide on Tecnai T12 microscopes (Thermo Fisher) at 120 kV using an Eagle camera. Data were stitched, uploaded, shared and annotated using Omero and PathViewer. The final pictures were directly acquired at the microscope in a manual standard way using the Eagle camera at 4kx4k.

Western blot

Organoids were collected from BME and washed with DPBS (Thermo Fisher). Organoids were lysed with RIPA Buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1% SDS, 0.5% Na-Deoxycholate, 1% NP-40, 1X Complete protease inhibitors (Roche)) by incubation on ice for 15 min, followed by sonication and centrifugation. Total protein extract from the supernatant was measured using a BCA assay kit (Thermo Fisher). For immunoprecipitation (IP) experiment, target proteins bound on the magnetic beads were directly eluted in reducing SDS-PAGE sample buffer (Biorad). Samples were loaded on a 4-20% Mini-PROTEAN TGX Precast Gel (Biorad) and subsequently transferred onto a PVDF membrane (Millipore). Membrane was blocked with 5% milk in PBST for 1 hour, and incubated with primary antibody overnight at 4 °C. Primary antibodies used in this study are: polyclonal Anti-ZNF800 antibody (Atlas Antibodies Cat#

HPA023090, RRID:AB_1859457), anti-ARX antibody (R and D Systems Cat# AF7068, RRID:AB_10973178), anti-PAX4 antibody (Thermo Fisher Scientific Cat# PA1-108, RRID:AB_2539882), anti-FLAG antibody (Sigma-Aldrich Cat# F1804, RRID:AB_262044), β -Actin (13E5) rabbit antibody (HRP-conjugated) (Cell Signaling Technology Cat#5125). Secondary antibody used is goat anti-rabbit-HRP (Thermo Fisher Cat#31460, RRID:AB_228341). Membranes were developed using Pierce ECL Western Blotting Substrate (Thermo Fisher) and imaged using ImageQuant LAS4000 western blot imaging systems (GE).

ELISA

ELISA was performed on the supernatant of organoids differentiated from different genotypes. Organoids were passaged and differentiated for 12 days. Medium was removed from the organoid cultures. Organoids still embedded in the BME were gently washed three times with DPBS (Thermo Fisher). 500 μ L of DPBS was added into the organoid cultures (per well of 12-well culture plate) and incubated overnight at 37 °C. Next day, supernatants from each well were centrifuged at 300 g for 3 min. Supernatants were collected and subjected to ELISA assay according to the respective manufacturer's protocols: SEROTONIN high sensitive ELISA (LDN), Human GLP1 (7-36) ELISA Kit (Abcam) and Human Ghrelin ELISA Kit (Ghrelin-28) (Abcam).

Immunohistochemistry staining

Organoids were removed from BME (R&D Systems), washed with DPBS (Thermo Fisher) and fixed with 4% paraformaldehyde (PFA) for 1 hour at room temperature. Organoids were then washed, dehydrated using a graded ethanol series and washed in xylene before embedding in paraffin. Sections cut from paraffin block were stained with antibodies according to manufacturer's instructions. Slides were imaged using a SLIDEVIEW VS200 Slide Scanner (Olympus). Antibodies used in this study are: polyclonal Anti-ZNF800 Antibody (Atlas Antibodies Cat# HPA023090, RRID:AB_1859457), mouse anti-Mucin2 (clone CCP58) (Monosan Cat#MON9810), goat anti-Chromogranin A (C-20) (Santa Cruz Biotechnology Cat# sc-1488), mouse anti-Glucagon (C-11) (Santa Cruz Biotechnology Cat#sc-514592, RRID:AB_2629431), rabbit anti-Motilin (Atlas Antibodies Cat# HPA069392, RRID:AB_2686136), goat anti-Somatostatin (D-20) (Santa Cruz Biotechnology Cat# sc-7819, RRID:AB_2302603), rabbit anti-Neurotensin (FL-170) (Santa Cruz Biotechnology Cat# sc-20806, RRID:AB_2155562), rabbit anti-Cholecystokinin (Thermo Fisher Scientific Cat# PA5-84814, RRID:AB_2791963), goat anti-Serotonin (Abcam Cat# ab66047, RRID:AB_1142794), sheep anti-ARX antibody (R and D Systems Cat# AF7068, RRID:AB_10973178), rabbit anti-PAX4 antibody (Thermo Fisher Scientific Cat# PA1-108, RRID:AB_2539882), rabbit anti-SOX4 (Thermo Fisher Scientific Cat# PA5-40681, RRID:AB_2608078), rabbit anti-NFIC antibody (Proteintech Cat# 16399-1-AP, RRID:AB_2878252), rabbit anti-TEF antibody (Thermo Fisher Scientific Cat# PA5-36187, RRID:AB_2553406), rabbit anti-ZHX2 antibody (Proteintech Cat# 20136-1-AP, RRID:AB_10666438), mouse anti-NEUROD1 antibody (BD Biosciences Cat# 563566, RRID:AB_2738279).

scRNA-seq organoid preparation

WT and ZNF800^{-/-} organoids were subjected to FACS on BD Influx cell sorter (BD Bioscience). DAPI staining was used to isolate live cells. 5000 live CHGA+ cells and 20,000 live CHGA- cells were separated and sorted into DPBS (Thermo Fisher) supplemented with 0.04% BSA. Cells were then pooled and counted. 4,000 live cells from each genotype were loaded to droplet-based scRNA-seq using the 10x Genomics platform. Libraries were prepared using the 10x Genomics Chromium 3' Gene Expression solution v3.1 and sequenced on a NovaSeq6000 (Illumina).

scRNA-seq analysis

Raw base call files were converted to FASTQ format using the *mkfastq* function from the Cell Ranger toolkit (10X genomics, version 7.0.1). Raw sequencing data can be found at NCBI GEO DataSets: GSE229585. Reads were then aligned to a custom human reference genome (based on GRCh38) in which the endogenous *MUC2*, *DEFA5* and *CHGA* genes are replaced by the tagged transgene constructs (MUC2-mNeonGreen, DEFA5-IRES-DsRed and CHGA-IRES-iRFP60) and transcript count tables were compiled using the Cell Ranger *count* function (including introns).

The resulting count tables were loaded into R (version 4.1.0) using the *Read10X* function from the Seurat package (version 4.1.0.9004). Cells with < 5000 or >100000 unique transcripts, or if $\geq 15\%$ of the total unique transcripts came from mitochondrial genes were excluded from further analysis. We captured averagely 27704 unique transcripts per single cell and 5637 genes per single cell. Normalized counts and scaled gene expression values were obtained using the SCTtransform (v2) method (66). Cells were projected in uniform manifold approximation and projection (UMAP) space using the first 20 principal components (PCs) which were calculated on the scaled gene expression values from the top 3000 most variable features. Clustering was performed on the shared nearest-neighbors graph calculated from the first 20 PCs using the Leiden algorithm (resolution 1.2) and clusters were assigned cell type labels through examination of the expression of known marker genes. Paneth cells were annotated as previously described (19) by calculating a module score (using the *AddModuleScore* function in Seurat) for Paneth cell-specific genes *DEFA6*, *PLA2G2A*, *PRSS2*, *REG3A*, *ITLN2* and *DEFA5*, and setting a cut-off based on the score distribution (>1 in this dataset). Differential gene expression between conditions (WT vs ZNF800^{-/-}) or clusters was determined by Wilcoxon Rank Sum test (implemented by the *FindMarkers* Seurat function) (table S7). The differential expression of ZNF800 targets between the WT and ZNF800^{-/-} conditions were performed on the intersection of genes associated with peaks in both the anti-ZNF800 and anti-FLAG ChIP data which were detected in at least 5% of single cells in either of the two conditions (table S5).

The assessment of transcription factor activity was performed as previously described using the pySCENIC implementation (v0.12.1) of the SCENIC pipeline (67). We performed 10 SCENIC runs using the unnormalized count data as input, along with the auxiliary input databases “*motifs-v9-nr.hgnc-m0.001-o0.0*”, “*hs_hgnc_tfs*” and “*hg38_refseq-r80_10kb_up_and_down_tss.mc9nr*” (accessed from <https://resources.aertslab.org/cistarget/>). Transcription factors were retained only if they were identified in ≥ 8 independent runs and the per cell transcription factor activity was calculated as the mean AUCell value across runs. Differential transcription factor activity (AUCell) was assessed by Wilcoxon Rank Sum test (table S4).

Chromatin immunoprecipitation in human small intestinal organoids

The protocol of chromatin immunoprecipitation (ChIP) performed in human SI organoids was adapted from previous study (68). In brief, organoids were dissociated with TrypLE (Thermo Fisher), fixed with formaldehyde (final concentration: 1%) and quenched by Glycine (final concentration: 0.125M). Chromatins were extracted by two-step lysis protocol in the presence of protease inhibitors: leupeptin hemisulfate, pepstatin A and phenylmethylsulfonyl fluoride. Chromatins were further sheared by sonication using Bioruptor Pico (Diagenode) with 7 cycles of 30s “ON” and 30s “OFF”. The size of the sheared chromatins was checked by taking an aliquot for analysis on a Bioanalyzer high-sensitivity DNA chip (Agilent).

For WT organoids, chromatin extracts were subjected to pre-clearing by using 10 µg of rabbit IgG (Sigma-Aldrich) and 60 µL of Dynabeads Protein G (Thermo Fisher). 10% of the pre-cleared chromatin was collected as input control. The rest of the pre-cleared chromatin extracts were then incubated with 4 µg of anti-ZNF800 antibody (Atals Antibodies Cat# HPA023090, RRID:AB_1859457) per replicate ($n=3$) overnight at 4 °C. The antibody-bound chromatins were then pulled down by adding 60 µL of Dynabeads Protein G (Thermo Fisher) per replicate. For ZNF800^{-/-} organoids with or without doxycycline-induced overexpression of *ZNF800*, chromatin extracts were incubated with 100 µL of Anti-FLAG M2 Magnetic Beads (Thermo Fisher) per replicate ($n=3$) overnight at 4 °C.

The magnetic beads containing pull-down chromatins were washed and eluted in 100 mM NaHCO3 and 0.1% SDS. Eluted chromatins were treated with RNaseA/T1 (Thermo Fisher) and Proteinase K (Promega). DNA purification was performed using phenol:chloroform:isoamyl alcohol isolation, followed by ethanol precipitation. ChIP-qPCR was first performed to confirm ZNF800 is indeed bound to target regions of interest. Sequencing libraries for ChIP-seq were constructed by using NEBNext Ultra II DNA Library Prep kit for Illumina (NEB) and NEBNext Multiplex Oligos for Illumina (Dual Index Primers Set 1) (NEB).

ChIP-seq analysis

ChIP-seq libraries were sequenced on a NextSeq2000 (Illumina). Libraries were demultiplexed based on the dual indexes, requiring a perfect match to each index. Read quality report was generated using Fastqc (0.11.9) (69) and checked manually. Read mapping to the human genome (hg38) was done with Bowtie2 (2.4.1) (70) using parameters: -q -N 0 -L 22 -X 500 --no-unal. Duplicate reads were marked and removed using SAMtools (version 1.10) (71). Unmapped reads, low mapping quality reads ($Q<30$) were discarded. Peak calling was performed by using MACS2's *callpeak* function (2.2.6) with a cutoff of $q<0.01$ (72). The multiBamSummary and plotCorrelation commands from deepTools (3.5.1) (73) were used to calculate Pearson's correlation matrix between replicates. Visualization of transcription factor binding and histone marks were done by computeMatrix and plotHeatmap commands from deepTools (3.5.1) (73) and Integrative Genomics Viewer (2.10.3) (74). The consensus peaks from three biological replicates were isolated by BedTools (2.30.0) (75) *intersect* function. ChIP-seq peaks were mapped to human genome and annotated to TxDb.Hsapiens.UCSC.hg38.refGene by using ChIPseeker (1.32.1) package (76, 77). Gene-ontology analysis was performed by *enrichGO* function (table S6). Gene set enrichment plot was done with *cnetplot* function. Raw sequencing data and processed ChIP-seq bigwig files can be found at NCBI GEO DataSets: GSE229583.

ChIP-seq qPCR

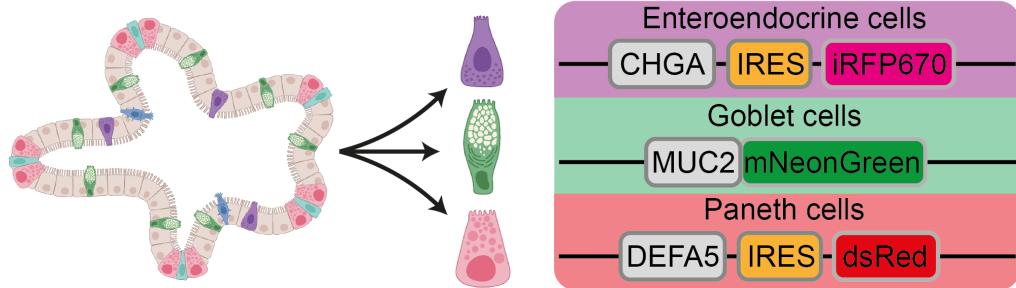
The eluted pull-down chromatins were first purified using phenol:chloroform:isoamyl alcohol isolation, followed by ethanol precipitation. A standard curve of each locus-targeting primer sets was determined by a series of sequentially diluted input DNA samples. Afterwards, eluted pull-down chromatin samples along with 0.8% input samples were subjected to qPCR analysis using iQ SYBR Green supermix (BioRad) in CFX Connect Real-Time PCR machine (BioRad). Ct readouts of each locus-targeting primer sets were mapped to the standard curve constructed by input DNAs. The relative fold changes between experimental groups and chromatin inputs were calculated. All ChIP-qPCR primers used in study are listed in table S8.

Statistical analysis

Data are presented as means with standard error of the mean (SEM) to indicate the variation within each experiment. Sample sizes (n) presented in this study are all biological replicates. Statistics analysis was performed in Prism 8 and R 4.2.0. Multiple t -tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli were used for comparison between different conditions versus control group. Wilcoxon test was used for comparison between two groups. Annotation for P values in figure legends regardless of statistical test type are: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$.

Fig. S1

A



B

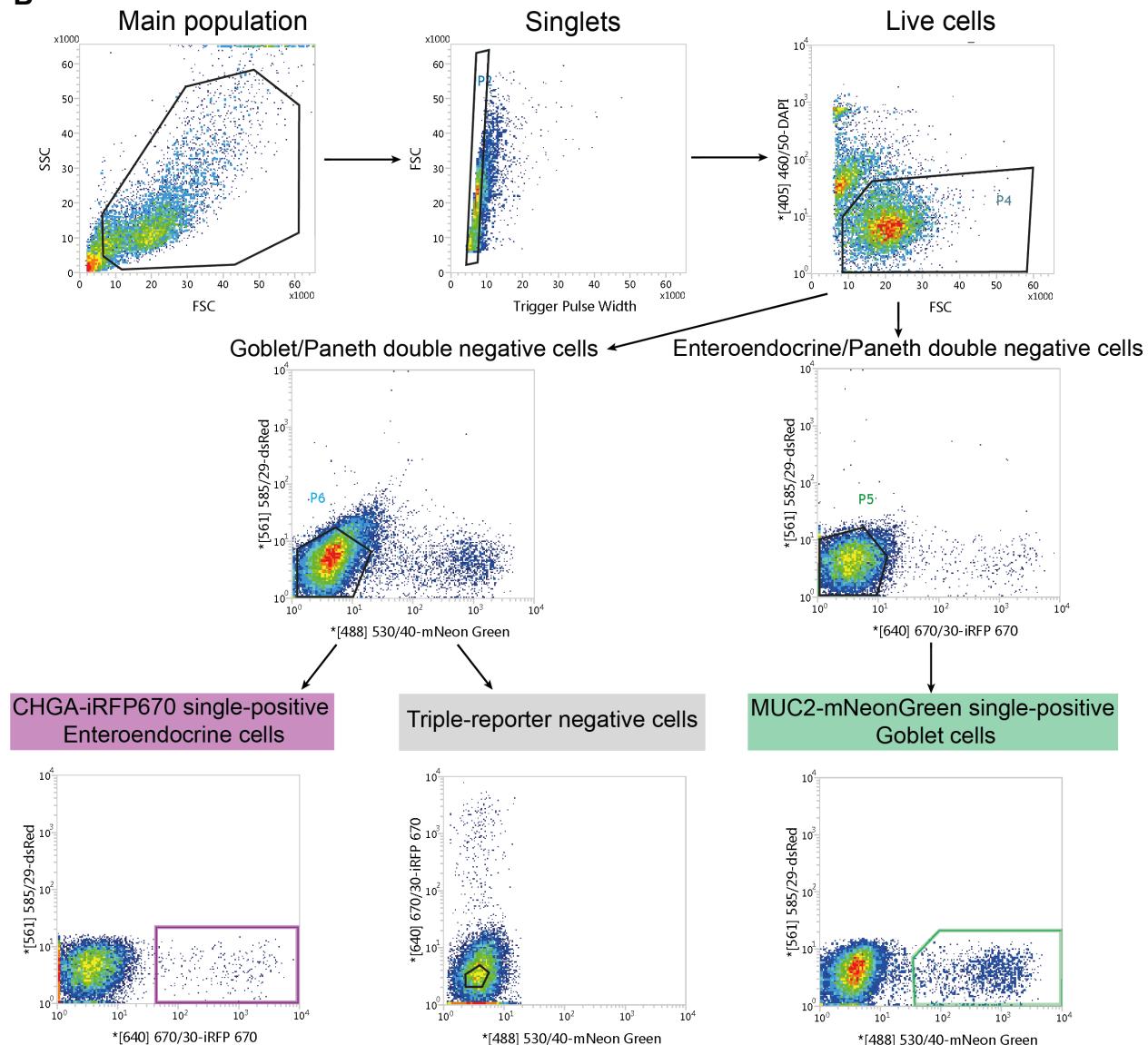


Fig. S1. Flow cytometry gating strategy used to sort differentiated secretory lineages by endogenous reporters

(A) Schematic of triple-reporter in human SI organoids. (B) Gating strategy used for sorting cells with single-positive of different reporters with representative flow cytometry plots.

Fig. S2

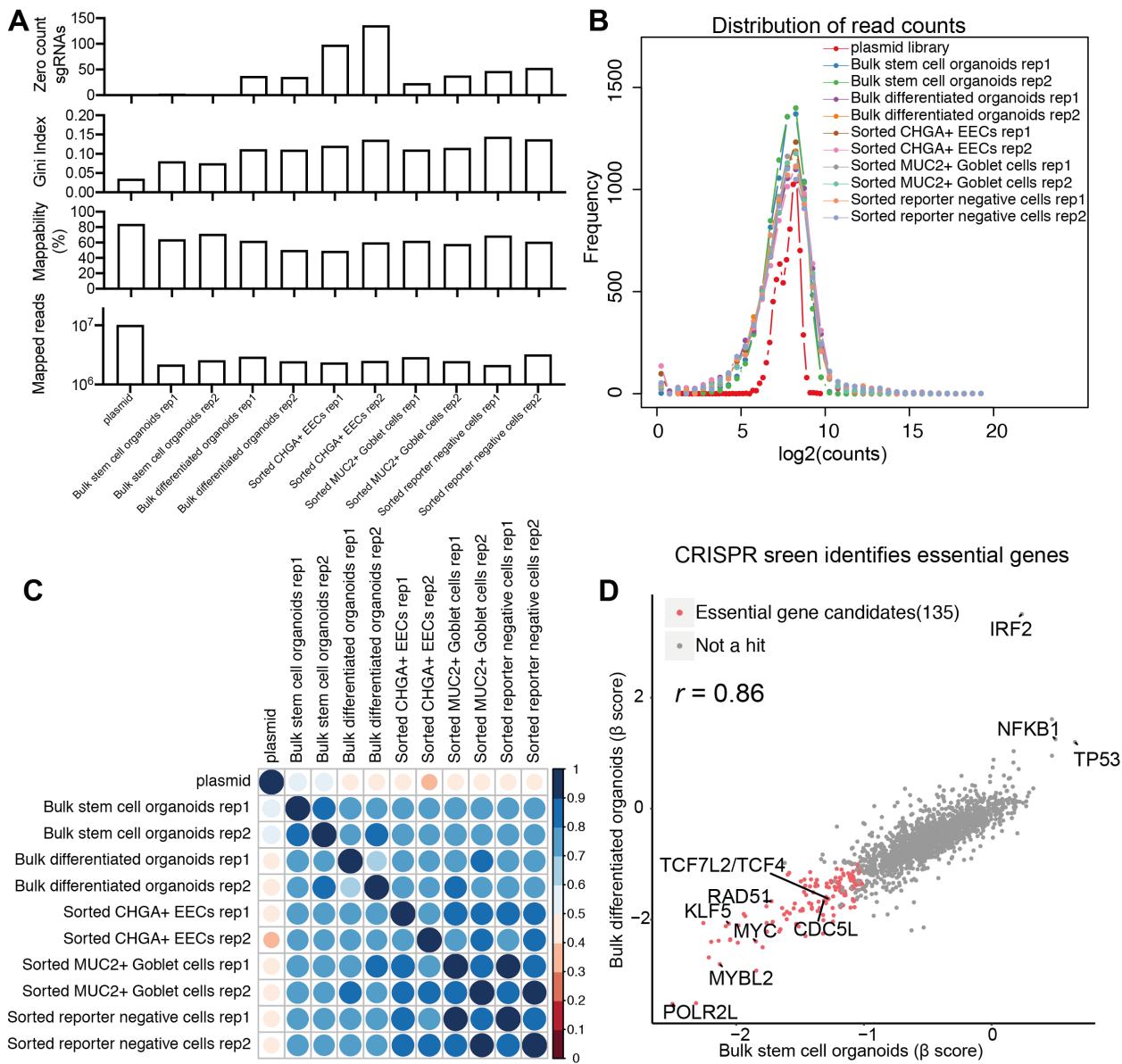


Fig. S2. Quality check of TFome-wide CRISPR screen

(A) Bar plots summarizing the number of zero-count sgRNAs, gini index, sequencing library mappability and mapped reads in all CRISPR screen sequencing libraries. (B) Frequency distribution of sgRNAs from TFome-wide screening library shown in \log_2 -normalized read counts from each class of biological replicates ($n=2$). (C) Heatmap presenting the correlation of sgRNA read counts from all CRISPR screen sequencing libraries. Spearman's correlation coefficient is calculated for all possible pairs of comparisons. (D) Scatter plot of enrichment β score of each TF gene, comparing stem cell bulk population and differentiated bulk population versus the plasmid library control. Spearman's correlation coefficient between two sets of β score is calculated.

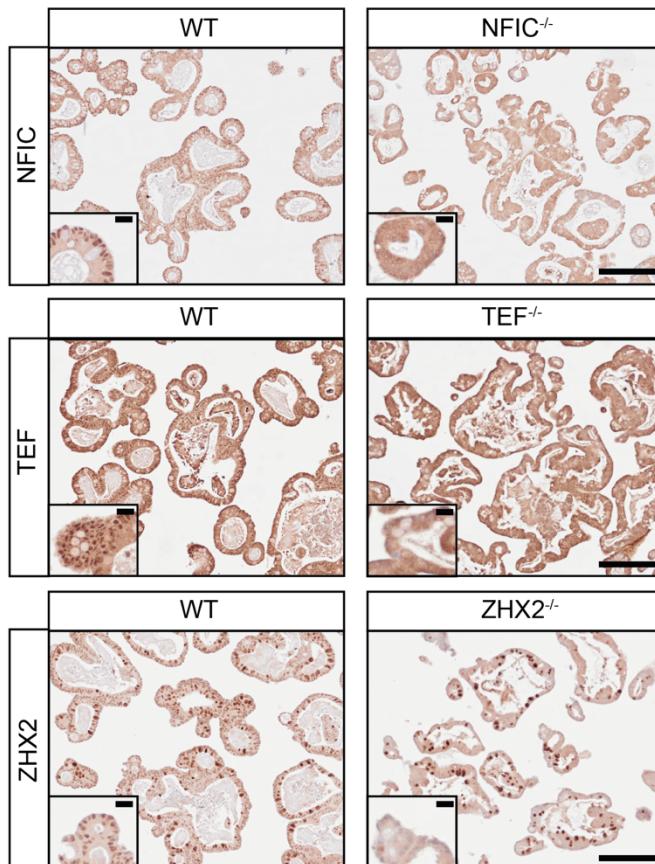
Fig. S3

A

CRISPR gRNA targeting locus



B



C

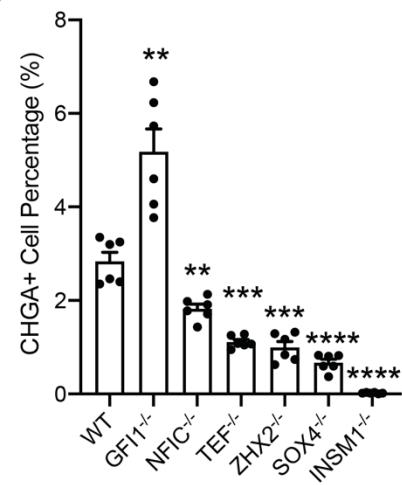


Fig. S3. Validation of individual hits

(A) Sequencing results of genotyping different knockout organoid lines. **(B)** Representative IHC staining of NFIC, TEF and ZHX2 in the respective knockout organoids versus WT organoids. Scale bars, 200 μ m; Zoom-in scale bars, 20 μ m. **(C)** Proportion of EECs as determined by FACS analysis of the CHGA reporter. Data are shown as mean \pm SEM. Statistics analysis is conducted between WT organoids and organoids with the respective gene knockout. ** P <0.01, *** P <0.001, **** P <0.0001 by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, n =6.

Fig. S4

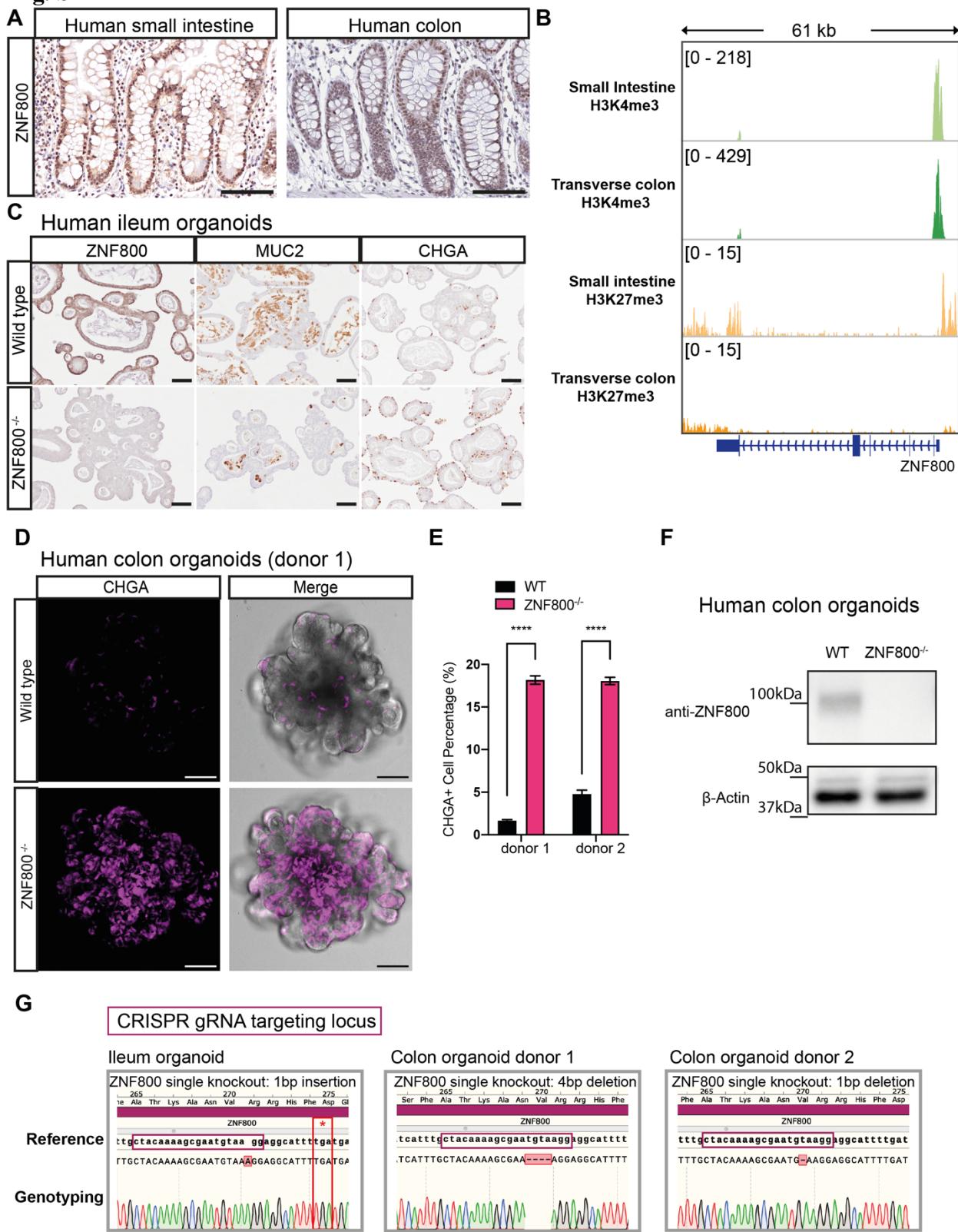
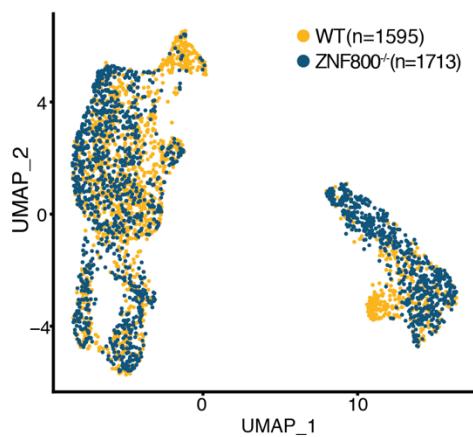


Fig. S4. Characterization of ZNF800 expression and ZNF800^{-/-} organoids

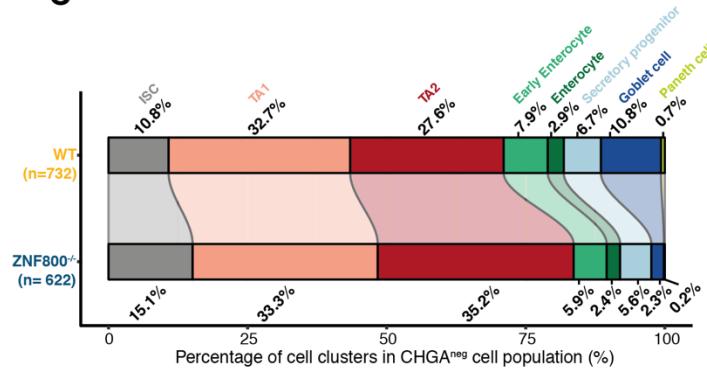
(A) IHC staining of ZNF800 in human small intestine tissue and human colon tissue (HPA023090, Human Protein Atlas). Scale bars, 100 μ m. **(B)** ChIP-seq tracks of chromatin marks at the *ZNF800* gene locus. **(C)** Representative IHC staining images of ZNF800, MUC2 and CHGA in WT and ZNF800^{-/-} organoids. Scale bars, 100 μ m. **(D)** Representative confocal images of WT and ZNF800^{-/-} human colon organoids. Enteroendocrine cells are highlighted by CHGA fluorescent reporter. Scale bars, 100 μ m. **(E)** Proportion of EECs as determined by FACS analysis of the CHGA reporter. Data are shown as mean \pm SEM. *** P <0.0001 by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, n =3. **(F)** Western blot profiling ZNF800 expression in whole organoid lysate of WT and ZNF800^{-/-} human colon organoids. β -actin served as loading control. **(G)** Sequencing results of genotyping ZNF800^{-/-} organoid lines.

Fig. S5

A



C



B

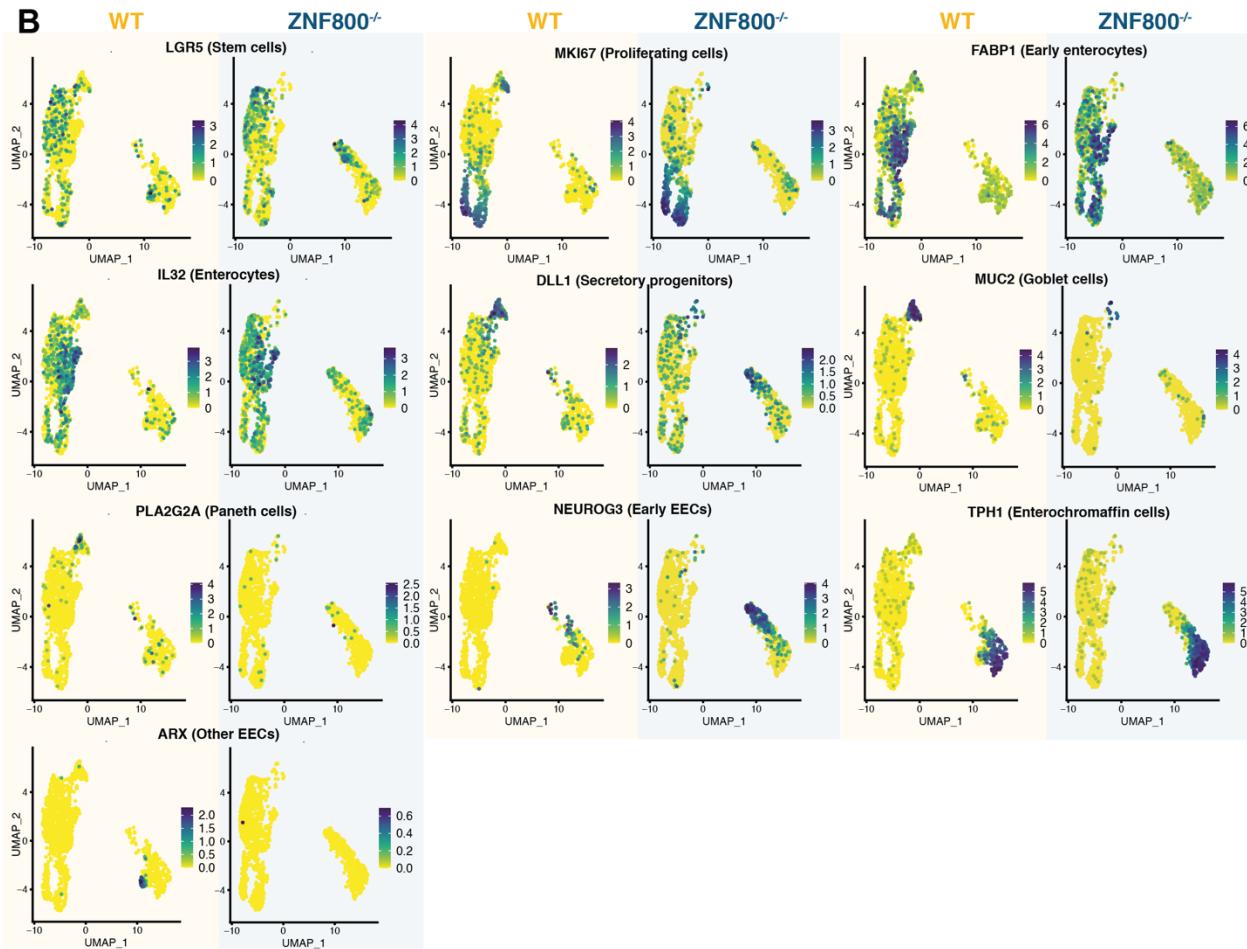


Fig. S5. scRNA-seq analysis of WT and ZNF800^{-/-} organoids

(A) UMAP plot showing the scRNA-seq datasets from WT (n=1595 cells) and ZNF800^{-/-} organoids (n=1713 cells). (B) UMAP plots showing various cell marker expression levels split by conditions from scRNA-seq. Cells are colored by normalized expression of the indicated genes. (C) Stacked bar plot showing the comparison of the cell cluster proportions identified in the CHGA- population of the scRNA-seq datasets from WT and ZNF800^{-/-} organoids.

Fig. S6

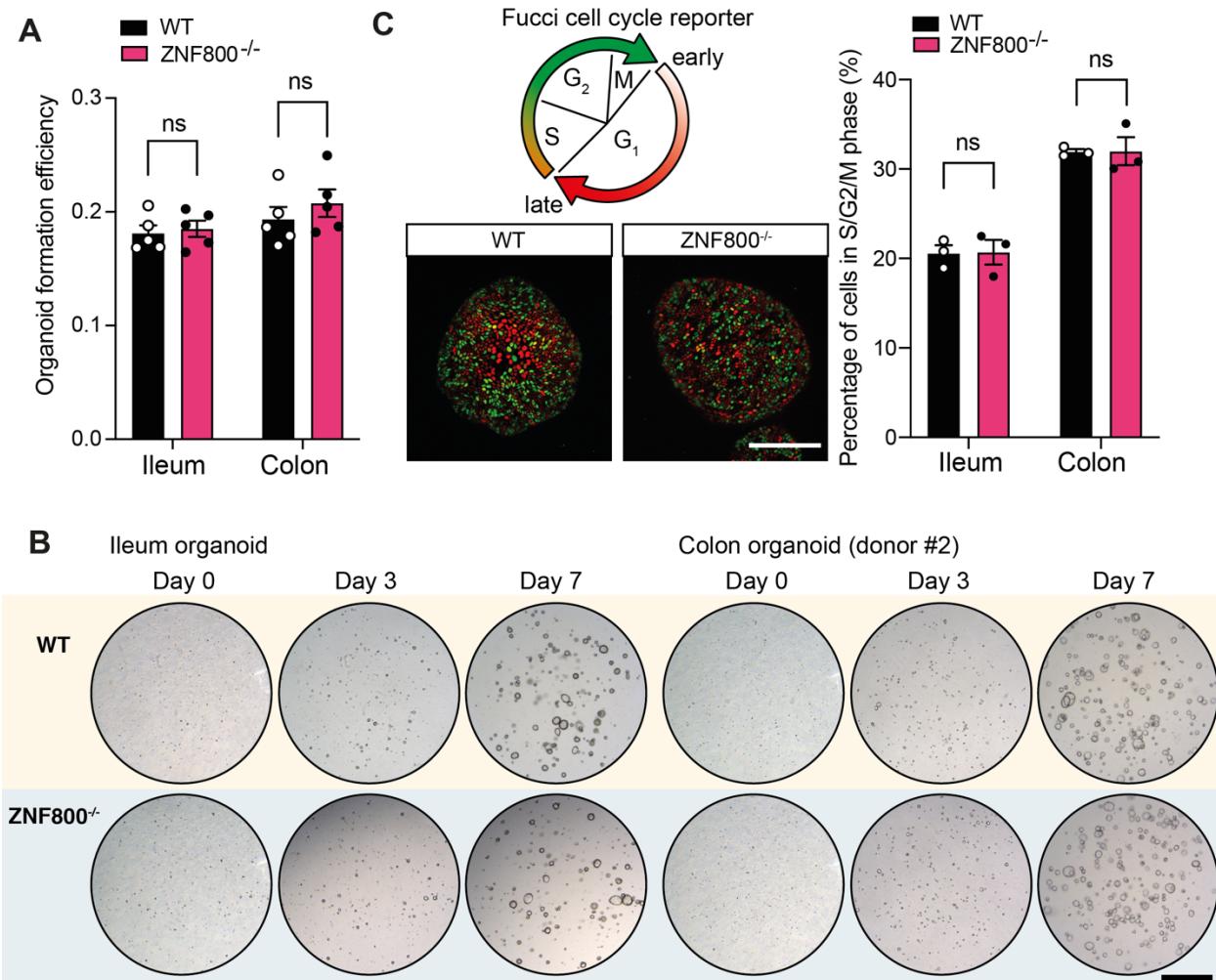


Fig. S6. ZNF800^{-/-} does not affect ISC homeostasis

(A) Quantification of clonal formation efficiency of WT and ZNF800^{-/-} organoids from human ileum and colon tissues. Data are shown as mean \pm SEM. ns, not significant by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, *n*=5. (B) Representative images of the progressive clonal formation from single cells of WT and ZNF800^{-/-} organoids from human ileum and colon tissues. Scale bars, 500 μ m. (C) Left top: Schematic of the presented fluorescent colors in different cell phases by Fucci cell cycle reporter construct, Left bottom: Representative confocal images of Fucci cell reporter in WT and ZNF800^{-/-} organoids from human ileum tissue. Scale bars, 200 μ m. Right: Quantification of cell cycle progression (S/G2/M phase) in WT and ZNF800^{-/-} organoids from human ileum and colon tissues. Data are shown as mean \pm SEM. ns, not significant by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, *n*=3.

Fig. S7

A

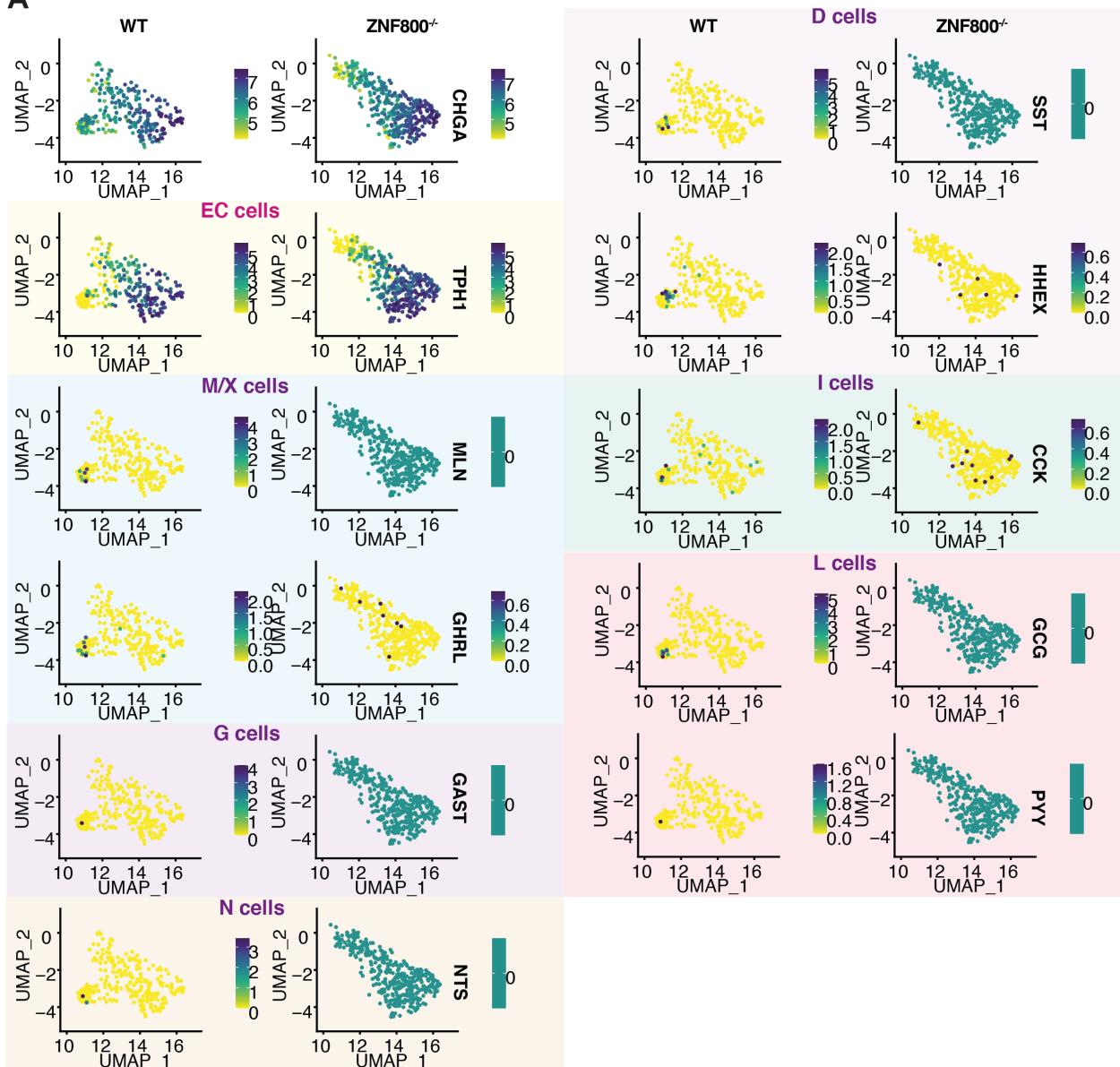


Fig. S7. Single-cell RNA expression of EEC subtype markers in CHGA+ cells of WT and ZNF800^{-/-} organoids

(A) UMAP plots showing various EEC subtype cell marker expression levels in CHGA+ population of the scRNA-seq datasets from WT and ZNF800^{-/-} organoids. Cells are colored by normalized expression of the indicated genes.

Fig. S8

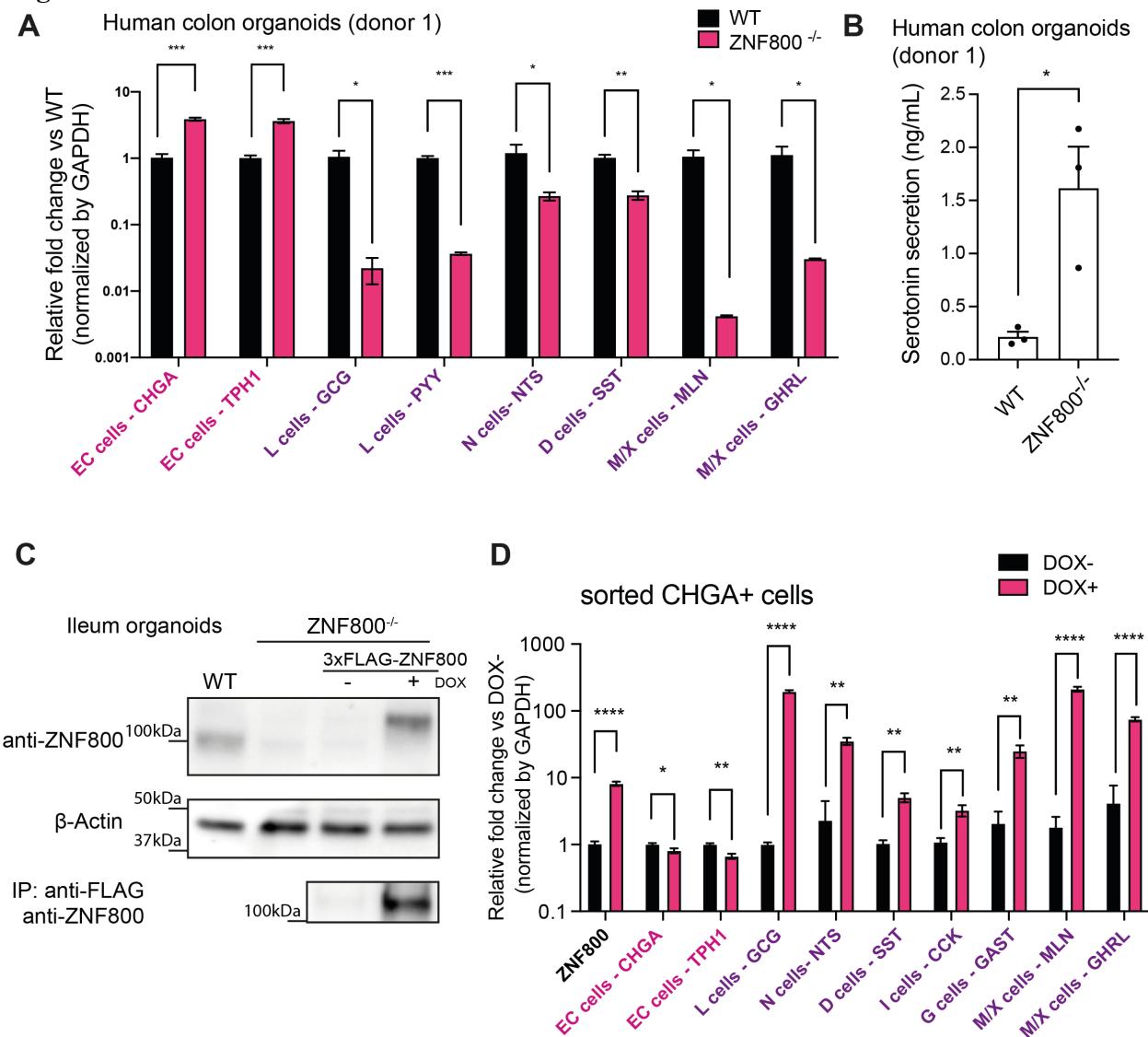


Fig. S8. Characterization of EC-biased phenotypes in ZNF800^{-/-} organoids

(A) RT-qPCR quantification of various EEC markers in WT and ZNF800^{-/-} human colon organoids. Data are shown as mean \pm SEM. * $P<0.05$, ** $P<0.01$, *** $P<0.001$ by multiple t -tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$.

(B) ELISA quantification of serotonin secretion of WT and ZNF800^{-/-} human colon organoids. Data are shown as mean \pm SEM. * $P<0.05$ by multiple t -tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$.

(C) Top panel: western blot profiling ZNF800 expression in whole organoid lysate of WT and ZNF800^{-/-} organoids with or without overexpression of FLAG-tagged ZNF800. Middle panel: β -actin served as loading control. Bottom panel: immunoprecipitation (IP) of ZNF800 by anti-FLAG antibody in ZNF800^{-/-} organoids with or without overexpression of FLAG-tagged ZNF800.

(D) RT-qPCR quantification of ZNF800 and various EEC markers in ZNF800^{-/-} human small intestinal organoids with or without dox-induced overexpression of ZNF800. Data are shown as mean \pm SEM. * $P<0.05$, ** $P<0.01$, **** $P<0.0001$ by multiple t -tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$.

Fig. S9

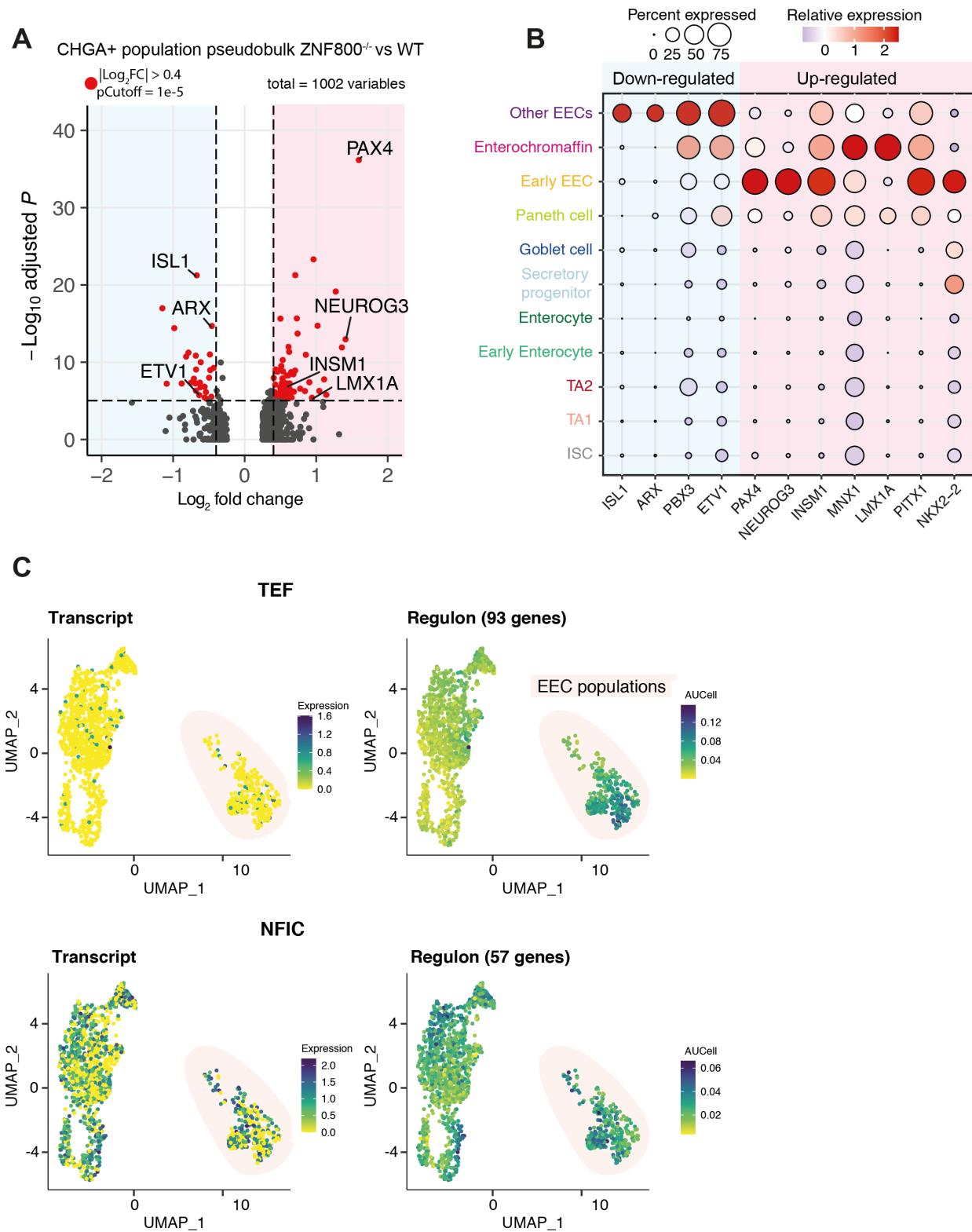


Fig. S9. Differential gene expression and regulon activities from scRNA-seq of WT and $ZNF800^{-/-}$ organoids

(A) Volcano plot showing the differentially expressed genes in ZNF800^{-/-} versus WT organoids. Selective TFs are highlighted. The full list of differentially expressed genes can be found in table S7. **(B)** Dot plot showing the relative expression and the percentage of cells expressing selected TFs across scRNA-seq clusters. **(C)** UMAP projection of TEF and NFIC gene expressions and regulon activities predicted by SCENIC. EEC populations are highlighted in the shaded areas.

Fig. S10

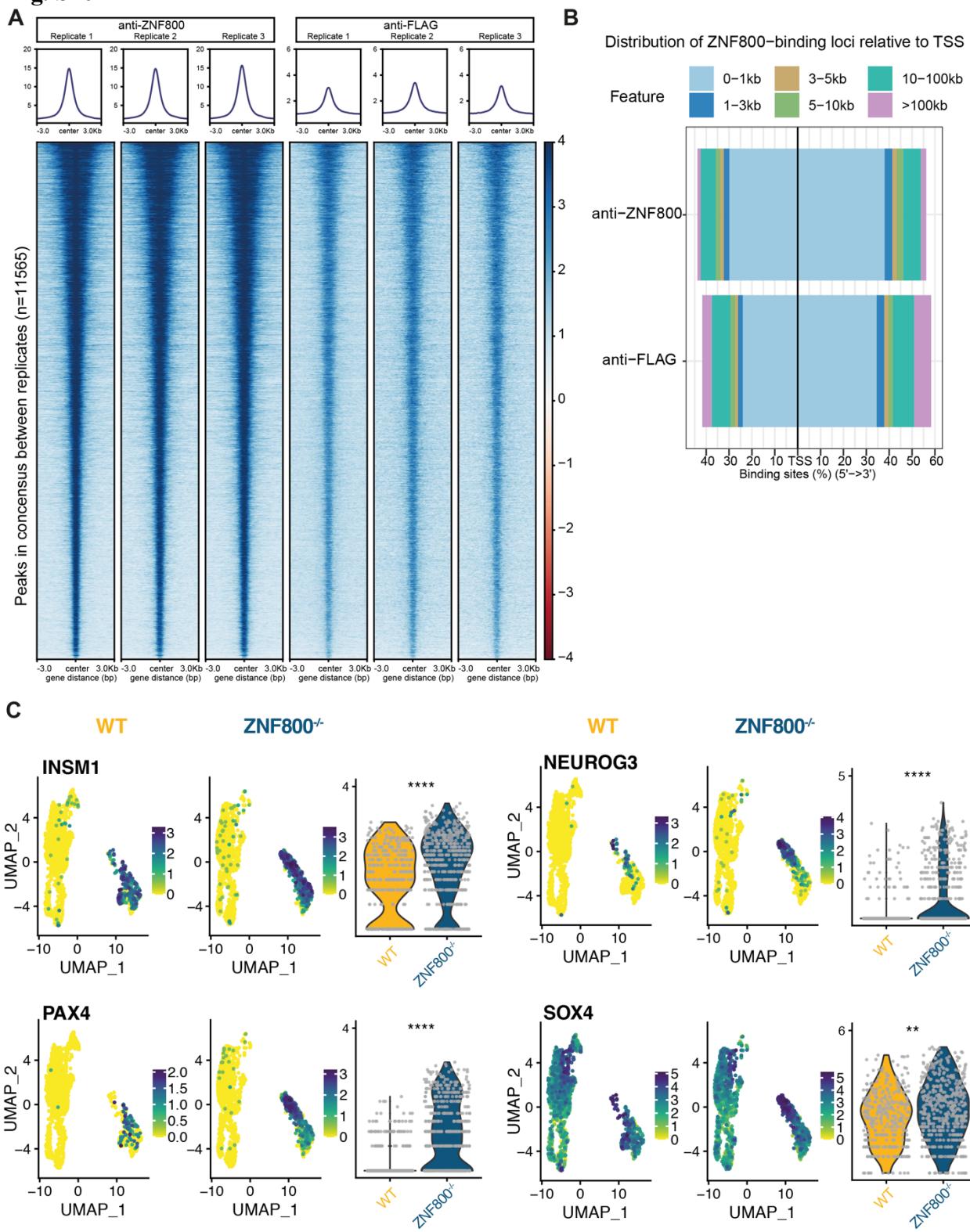
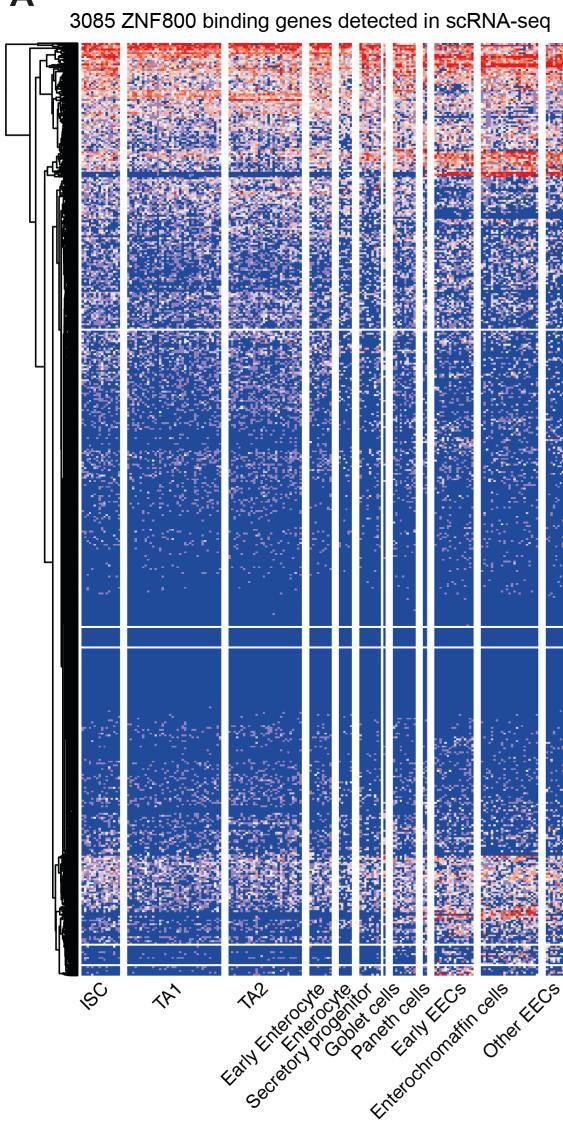


Fig. S10. ZNF800 ChIP-seq analysis revealed direct downstream targets in line with scRNA-seq

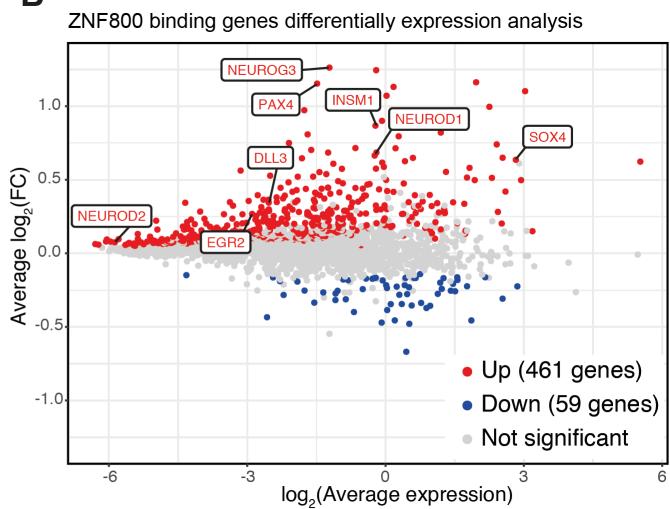
(A) Heatmaps showing ChIP-seq coverage at 11565 sites with consensus ZNF800 occupancy in different ChIP-seq datasets and replicates. **(B)** Annotation of ZNF800 ChIP-seq peaks in human genome relative to TSS. **(C)** UMAP plots showing gene expression levels of *INSM1*, *NEUROG3*, *PAX4* and *SOX4* in WT and ZNF800^{-/-} organoids by scRNA-seq. Cells are colored by normalized expression of the indicated genes. Violin plots showing the comparisons of gene expressions in CHGA+ EEC populations between WT and ZNF800^{-/-} conditions. Statistics analysis is conducted between WT and ZNF800^{-/-} organoids. ** $P<0.01$, *** $P<0.0001$ by Wilcoxon Test.

Fig. S11

A



B



C

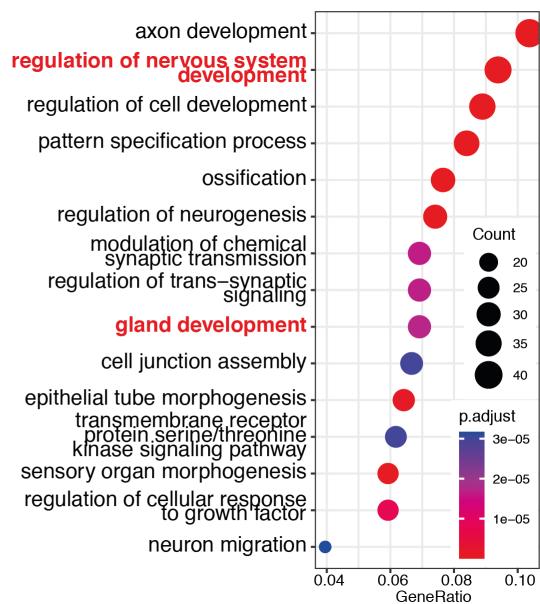


Fig. S11. ZNF800 functions as a repressor on target genes

(A) Heatmap showing the normalized expression of 3085 ZNF800 binding genes detected in scRNA-seq. (B) MA plot of the differential gene expression of ZNF800 binding genes in scRNA-seq. Genes with statistically significance (adjusted $p < 0.05$) were highlighted with colors. Full list can be found in table S5. (C) Gene ontology of 870 genes nearest to top 1000 sites with ZNF800 occupancy in WT organoids. Top fifteen selected terms of 261 significant terms are shown ($p.\text{adjust} < 0.01$). Full list can be found in table S6.

Fig. S12

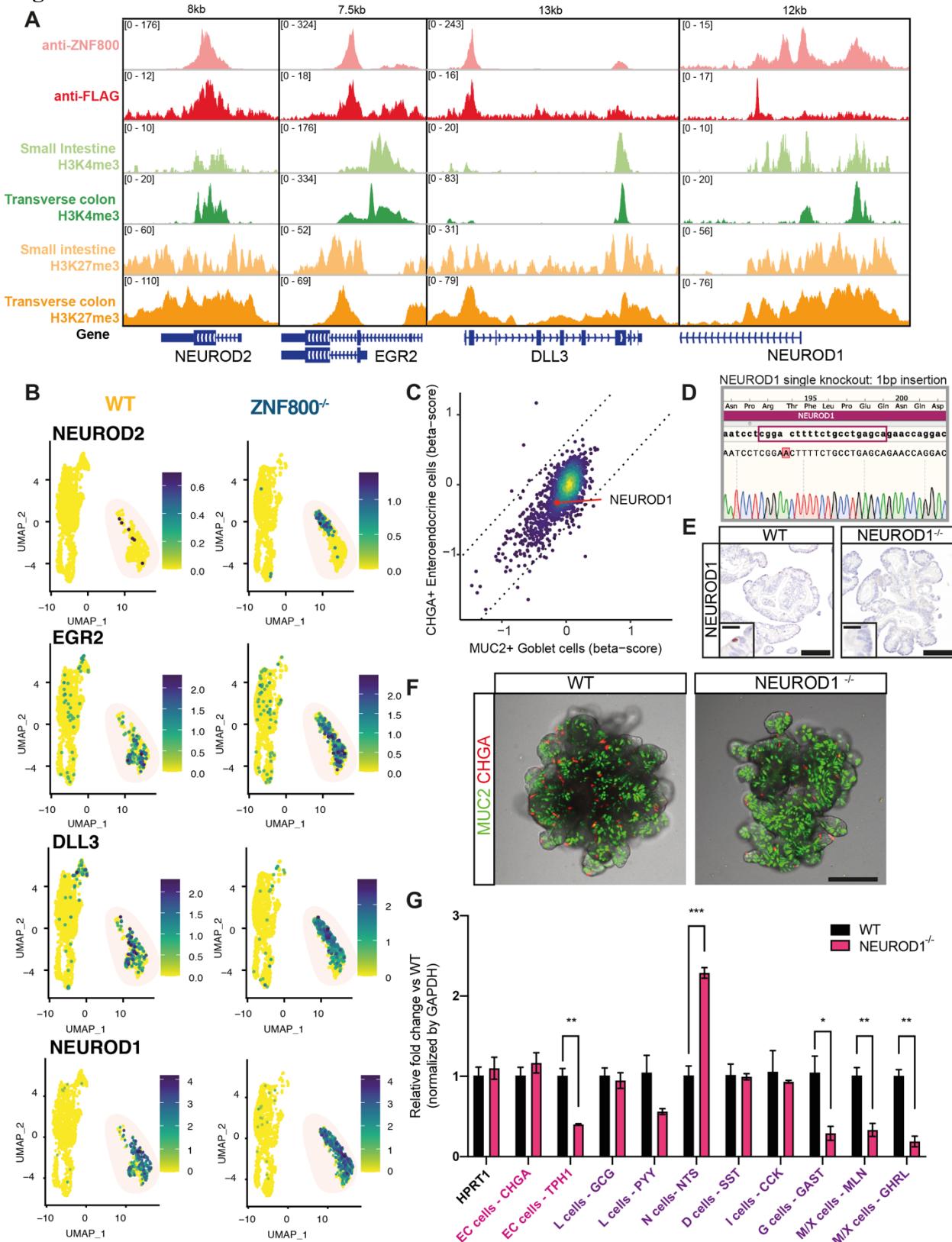


Fig. S12. ZNF800 ChIP-seq analysis revealed other EEC specific downstream targets

(A) ChIP-seq tracks at the *NEUROD2*, *EGR2*, *DLL3* and *NEUROD1* loci. **(B)** UMAP plots showing gene expression levels of *NEUROD2*, *EGR2*, *DLL3* and *NEUROD1* in WT and ZNF800^{-/-} organoids by scRNA-seq. Cells are colored by normalized expression of the indicated genes. EEC populations are highlighted in the shaded area. **(C)** Scatter plot of enrichment β score of TFome CRISPR screen with *NEUROD1* highlighted. **(D)** Sequencing results of genotyping *NEUROD1*^{-/-} organoids. **(E)** Representative IHC staining of *NEUROD1* in WT and *NEUROD1*^{-/-} organoids. Scale bars, 100 μ m; Zoom-in scale bars, 20 μ m. **(F)** Representative confocal images of WT and *NEUROD1*^{-/-} human small intestinal organoids. Representative marker genes for enteroendocrine (CHGA, red) and goblet (MUC2, green) cells are highlighted by fluorescent reporters. Scale bars, 200 μ m. **(G)** RT-qPCR quantification of various EEC markers in CHGA+ cell populations of WT and *NEUROD1*^{-/-} organoids. Data are shown as mean \pm SEM. * $P<0.05$, ** $P<0.01$, *** $P<0.001$ by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$.

Fig. S13

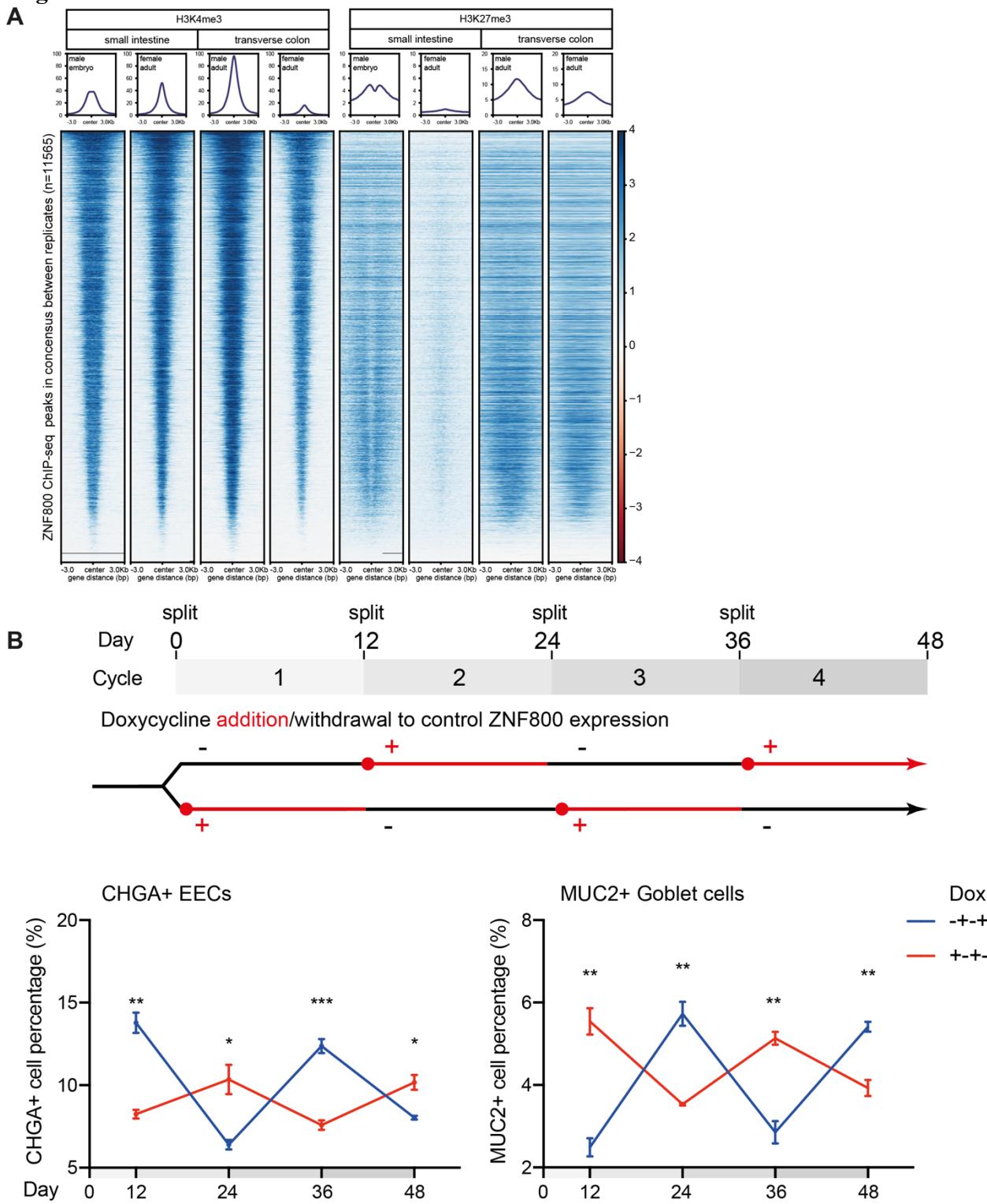


Fig. S13. ZNF800 ChIP-seq analysis revealed co-occupancy with bivalent chromatin state

(A) Heatmaps showing ChIP-seq coverage of H3K4me3 and H3K27me3 in human small intestine and colon tissues at 11565 sites with consensus ZNF800 occupancy. (B) Top: Schematic illustration of the experimental setup of the ZNF800 dox on-and-off assay; Bottom:

Proportion of EECs and goblet cells as determined by FACS analysis of the respective reporters in ZNF800^{-/-} organoids with or without dox-induced ZNF800 expression at different time points. Data are shown as mean \pm SEM. * $P<0.05$, ** $P<0.01$, *** $P<0.001$ by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$.

Fig. S14

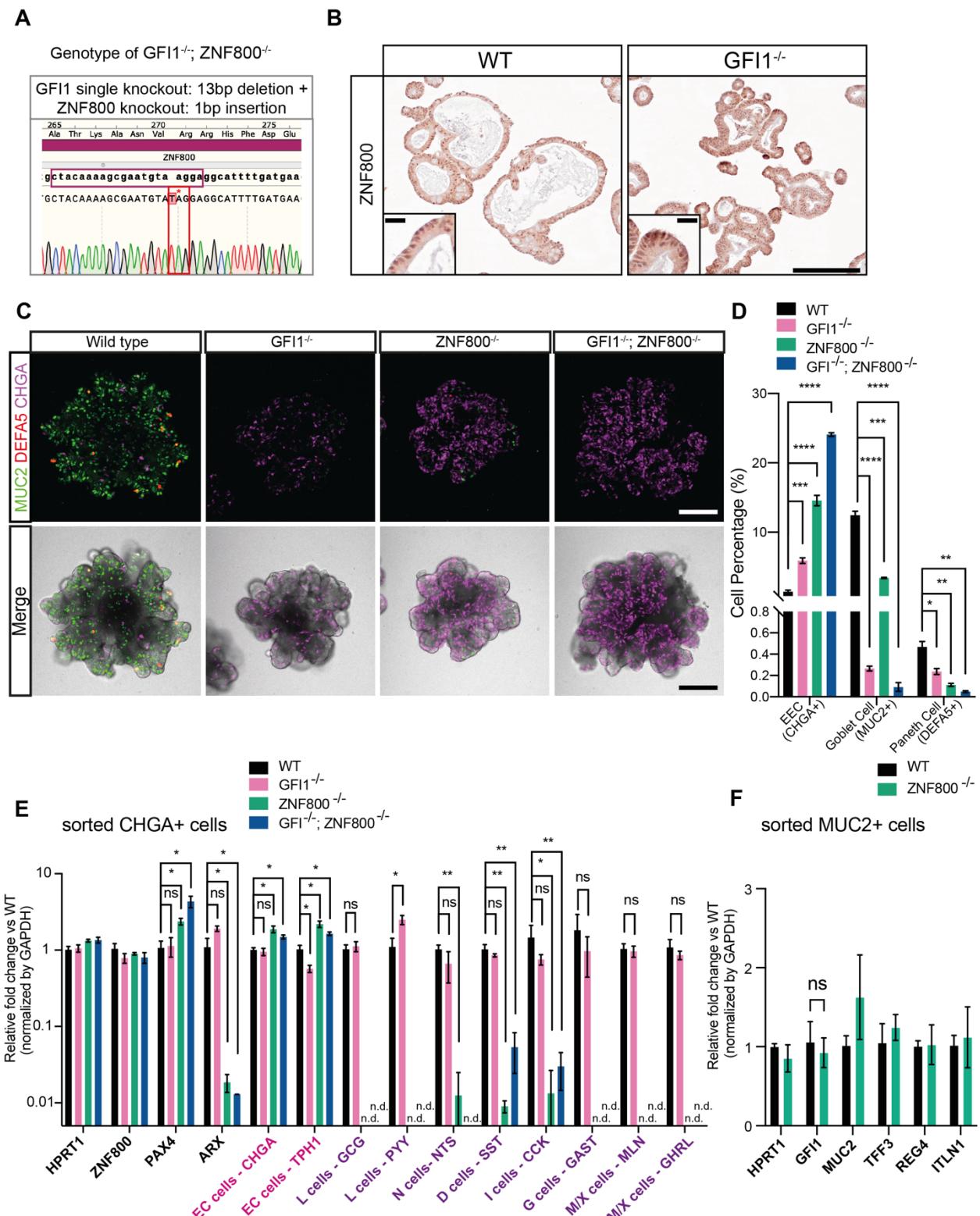


Fig. S14. GFI1 and ZNF800 function independently in repressing EEC differentiation

(A) Sequencing results of genotyping GFI1^{-/-};ZNF800^{-/-} organoids. **(B)** Representative IHC staining of ZNF800 in WT and GFI1^{-/-} organoids. Scale bars, 200 μ m; Zoom-in scale bars, 20 μ m. **(C)** Representative confocal images of WT, GFI1^{-/-} and GFI1^{-/-};ZNF800^{-/-} organoids. Representative marker genes for enteroendocrine (CHGA, magenta), goblet (MUC2, green) and Paneth (DEFA5, red) cells are highlighted by fluorescent reporters. Scale bars, 200 μ m. **(D)** Proportion of EECs, goblet and Paneth cells as determined by FACS analysis of the respective reporters in WT, GFI1^{-/-} and GFI1^{-/-};ZNF800^{-/-} organoids. Data are shown as mean \pm SEM. * $P<0.05$, ** $P<0.01$, *** $P<0.001$, **** $P<0.0001$ by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$. **(E)** RT-qPCR quantification of various EEC TFs and markers in CHGA+ cell populations of WT, GFI1^{-/-} and GFI1^{-/-};ZNF800^{-/-} organoids. Data are shown as mean \pm SEM, n.d. not detected. * $P<0.05$, ** $P<0.01$ by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$. **(F)** RT-qPCR quantification of GFI1 and various goblet cell markers in MUC2+ cell populations of WT and ZNF800^{-/-} organoids. Data are shown as mean \pm SEM. ns, not significant by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$.

Fig. S15

A

Top differentially expressed genes ($ZNF800^{-/-}$ vs WT) in different cell clusters

| ISC | TA1 | TA2 | Secretory progenitor | Early enterocyte | Goblet cell | Early EEC | Enterochromaffin | Other EECs |
|-----------------|-----------------|-----------------|----------------------|------------------|-----------------|-----------------|------------------|---------------|
| <u>COL4A2</u> | <u>COL4A2</u> | <u>COL4A2</u> | <u>TMEM178B</u> | <u>COL4A2</u> | <u>COL4A2</u> | <u>PAX4</u> | <u>RNF217</u> | <u>CADM2</u> |
| <u>TMEM178B</u> | <u>MATK</u> | <u>IGFBP2</u> | <u>COL4A2</u> | <u>BTBD11</u> | <u>TMEM178B</u> | <u>FAM241B</u> | <u>TMEM178B</u> | <u>RNF217</u> |
| <u>BTBD11</u> | <u>BTBD11</u> | <u>PITX1</u> | <u>RNF217</u> | <u>RNF217</u> | <u>NEUROG3</u> | <u>TMEM178B</u> | <u>BTBD11</u> | <u>IGFBP2</u> |
| <u>RNF217</u> | <u>TMEM178B</u> | <u>MATK</u> | <u>BTBD11</u> | <u>IGFBP2</u> | <u>HEATR3</u> | <u>COL4A2</u> | <u>KCNQ3</u> | |
| <u>BICC1</u> | <u>IGFBP2</u> | <u>TMEM178B</u> | <u>BICC1</u> | <u>TMEM178B</u> | <u>NTN1</u> | <u>ARF5</u> | <u>DPY19L2</u> | |
| <u>ST6GAL1</u> | <u>BICC1</u> | <u>BTBD11</u> | <u>SPON1</u> | <u>BICC1</u> | <u>SHC3</u> | <u>NEUROG3</u> | <u>MMP16</u> | |
| <u>IL7</u> | <u>HTRA1</u> | <u>BICC1</u> | <u>RIN3</u> | <u>PID1</u> | <u>BTBD11</u> | <u>CADPS</u> | <u>PAX4</u> | |
| <u>IGFBP2</u> | <u>REG1A</u> | <u>HTRA1</u> | <u>ST6GAL1</u> | <u>CELF2</u> | <u>FAM241B</u> | <u>LRRN2</u> | <u>PITX1</u> | |
| <u>THBS2</u> | <u>TGFB1</u> | <u>REG1A</u> | | <u>IL7</u> | <u>PCSK1N</u> | <u>HMGB3</u> | <u>PGM5</u> | |
| <u>CACNB2</u> | <u>ASIC2</u> | <u>REG1A</u> | | | | <u>MS4A8</u> | <u>TAC1</u> | |
| <u>ASIC2</u> | <u>SULT1C2</u> | <u>FCGBP</u> | | | | <u>STXBP5L</u> | <u>CMC1</u> | |
| <u>UTRN</u> | <u>CCND2</u> | <u>REG1B</u> | | | | <u>BRINP3</u> | <u>PRSS2</u> | |
| <u>REG1A</u> | <u>HMGCS1</u> | | | | | | | |
| <u>TSHZ2</u> | <u>SCD</u> | | | | | | | |
| <u>SHC3</u> | <u>MSMO1</u> | | | | | | | |
| <u>VAV3</u> | <u>ECM1</u> | | | | | | | |
| <u>PDZRN3</u> | <u>IDI1</u> | | | | | | | |
| | <u>FDPS</u> | | | | | | | |

Up-regulated: shown in more than one cell cluster

Down-regulated

p.adj < 0.05, $|\text{Log}_2 \text{FC}| > 0.5$, ranked by p.adj

B

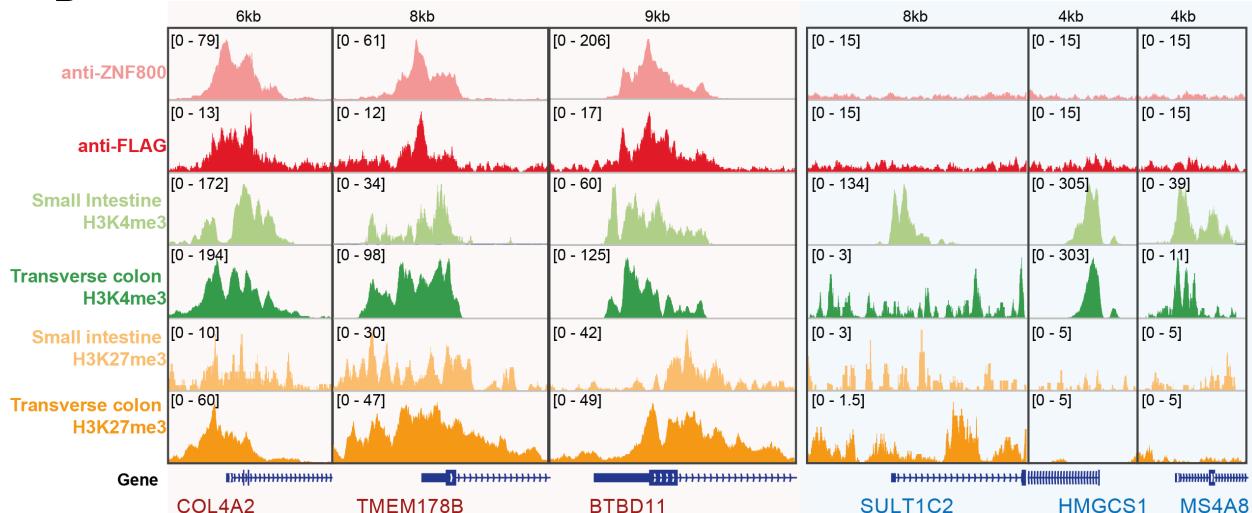


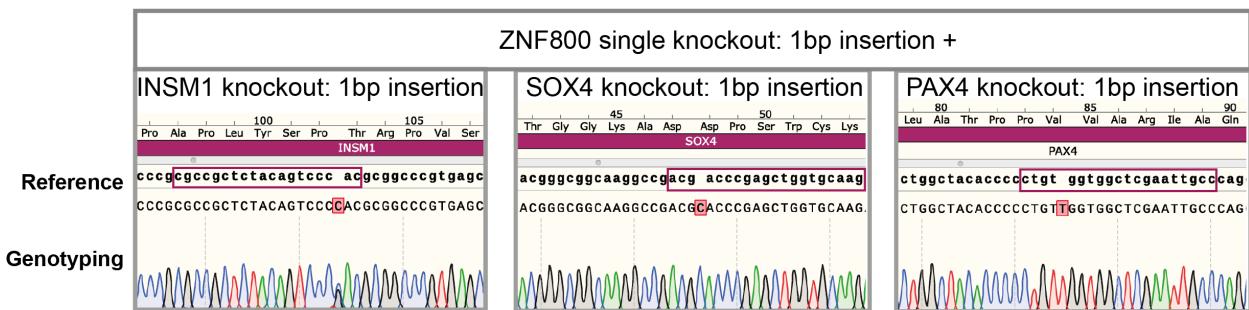
Fig. S15. $ZNF800^{-/-}$ causes differential gene expression in different cell clusters

(A) List of top differentially expressed genes in $ZNF800^{-/-}$ versus WT organoids in different cell clusters. Full list can be found in table S7. (B) ChIP-seq tracks at the loci of up-regulated genes *COL4A2*, *TMEM178B*, *BTBD11* and down-regulated genes *SULT1C2*, *HMGCS1*, *MS4A8*.

Fig. S16

A

Genotyping of double knockout: CRISPR gRNA targeting locus



B

PAX4 single knockout: 1bp deletion

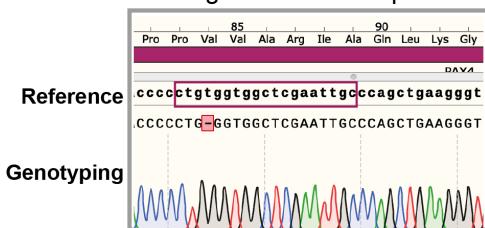


Fig. S16. Sequencing genotyping of different knockout organoid lines

(A) Sequencing results of genotyping double knockout organoids. (B) Sequencing results of genotyping PAX4^{-/-} knockout organoids.

Fig. S17

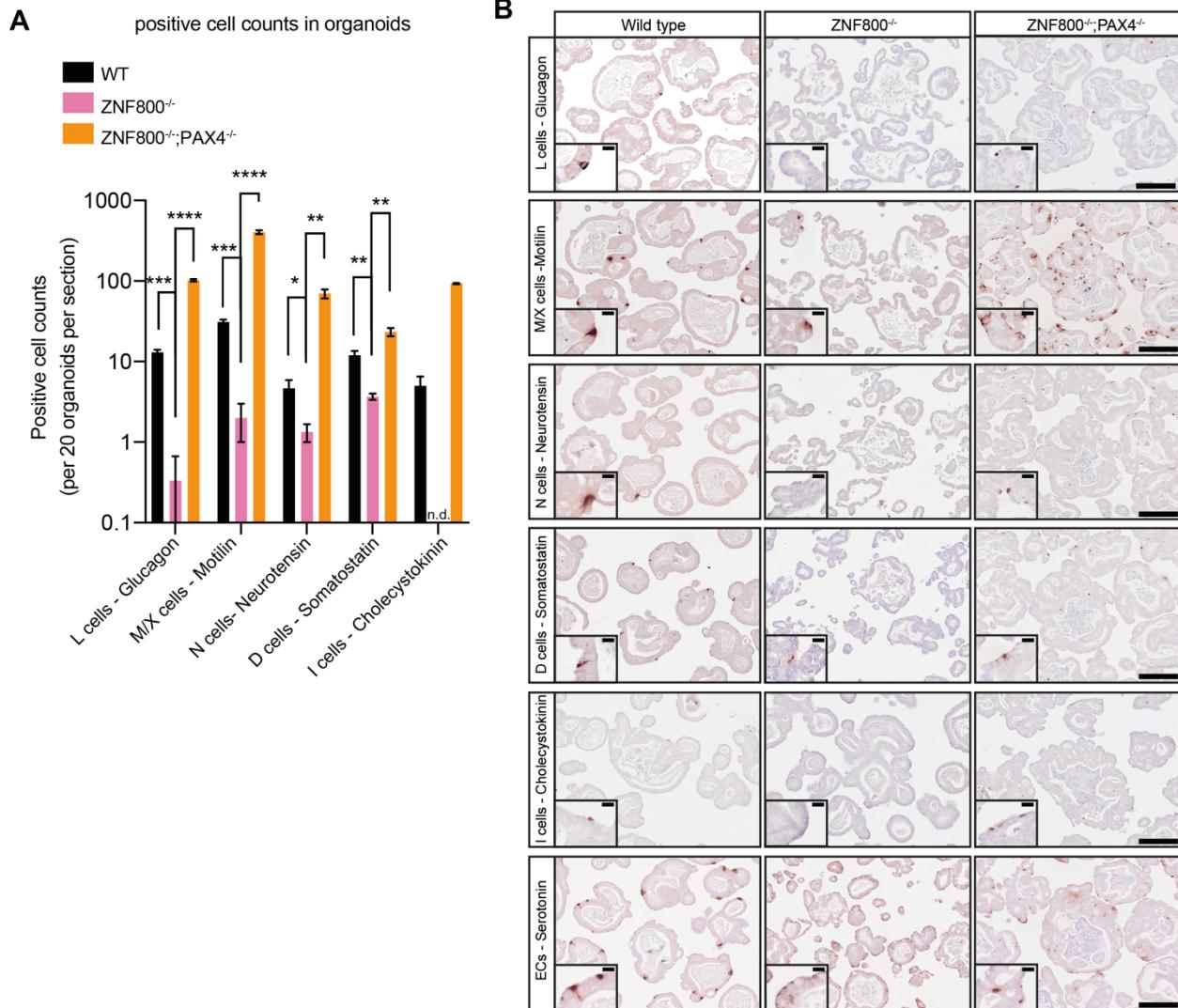


Fig. S17. IHC staining of EEC subtype hormones in organoids from different genotypes

(A) Quantification of cells positive for staining of hormone proteins. Data are shown as mean \pm SEM. * $P<0.05$, ** $P<0.01$, *** $P<0.001$, **** $P<0.0001$ by multiple t -tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with $Q=5\%$, $n=3$. **(B)** Representative IHC staining images of glucagon, motilin, neurotensin, somatostatin, cholecystokinin and serotonin in WT, ZNF800^{-/-} and ZNF800^{-/-}; PAX4^{-/-} organoids. Scale bars, 200 μ m; Zoom-in scale bars, 20 μ m.

Fig. S18

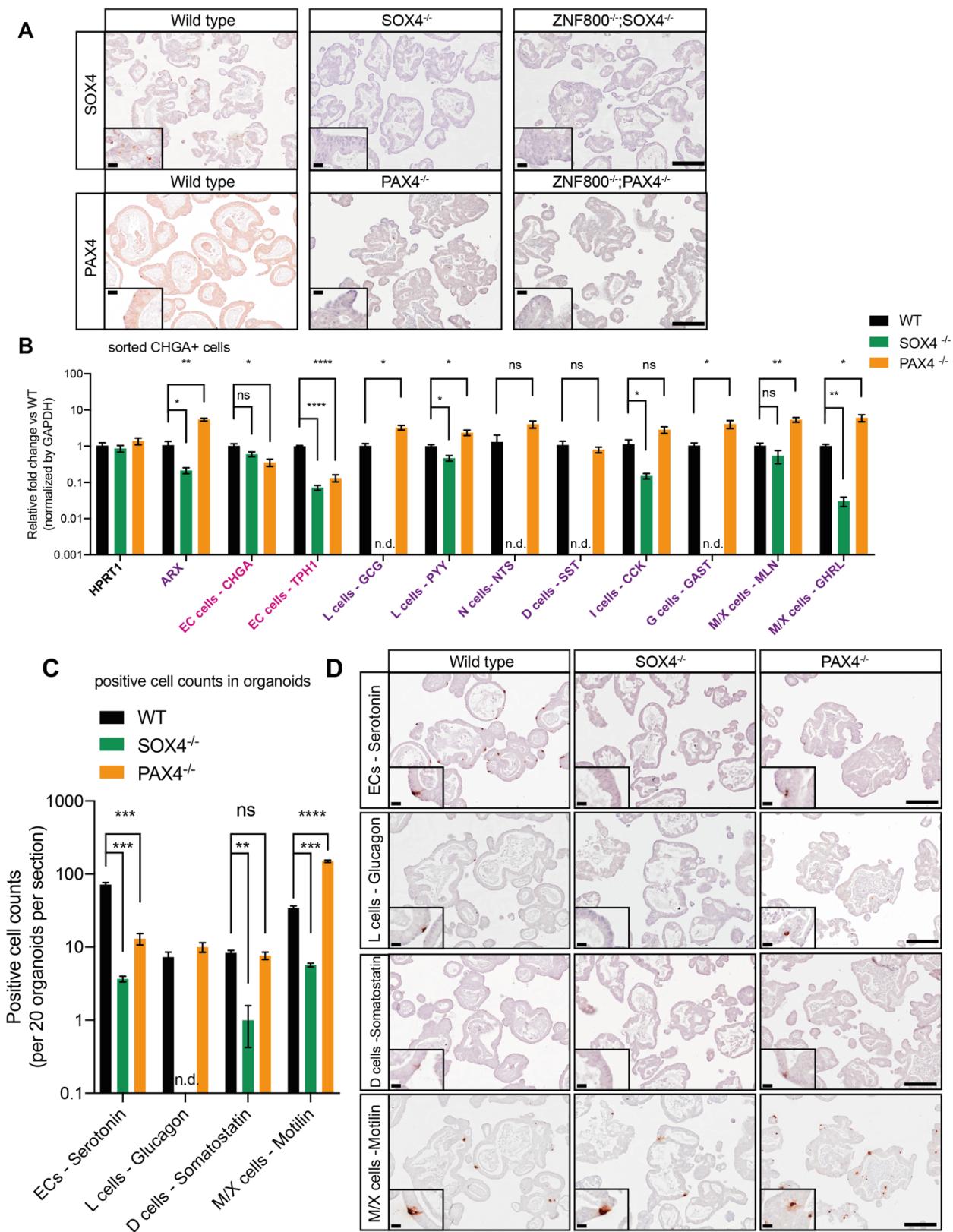


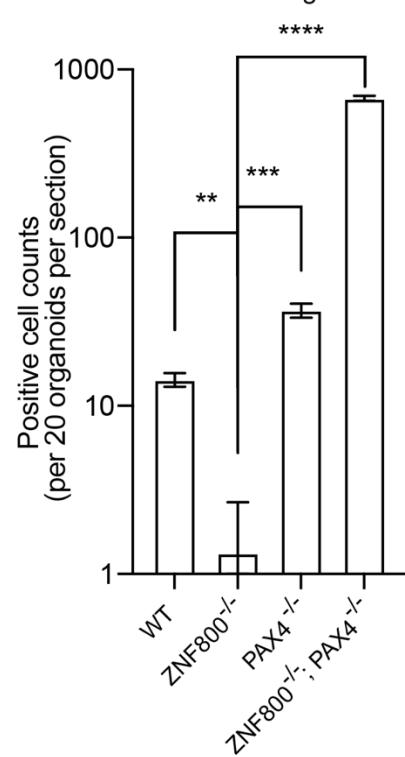
Fig. S18. Characterization of SOX4 and PAX4 single knockout

(A) Top: Representative IHC staining images of SOX4 in WT, SOX4^{-/-} and ZNF800^{-/-}; SOX4^{-/-} organoids. Bottom: Representative IHC staining images of PAX4 in WT, PAX4^{-/-} and ZNF800^{-/-}; PAX4^{-/-} organoids. Scale bars, 200 μ m; Zoom-in scale bars, 20 μ m. **(B)** RT-qPCR quantification of various EEC markers in CHGA+ cell populations of WT, SOX4^{-/-} and PAX4^{-/-} organoids. Data are shown as mean \pm SEM, n.d. not detected. * $P<0.05$, ** $P<0.01$, **** $P<0.0001$ by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$. **(C)** Quantification of cells positive for staining of hormone proteins. Data are shown as mean \pm SEM, n.d. not detected. ns, not significant, ** $P<0.01$, *** $P<0.001$, **** $P<0.0001$ by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$. **(D)** Representative IHC staining images of serotonin, glucagon, somatostatin and motilin in serial sections of WT, SOX4^{-/-} and PAX4^{-/-} organoids. Scale bars, 200 μ m; Zoom-in scale bars, 20 μ m.

Fig. S19

A

ARX+ cell counts in organoids



B

ARX staining

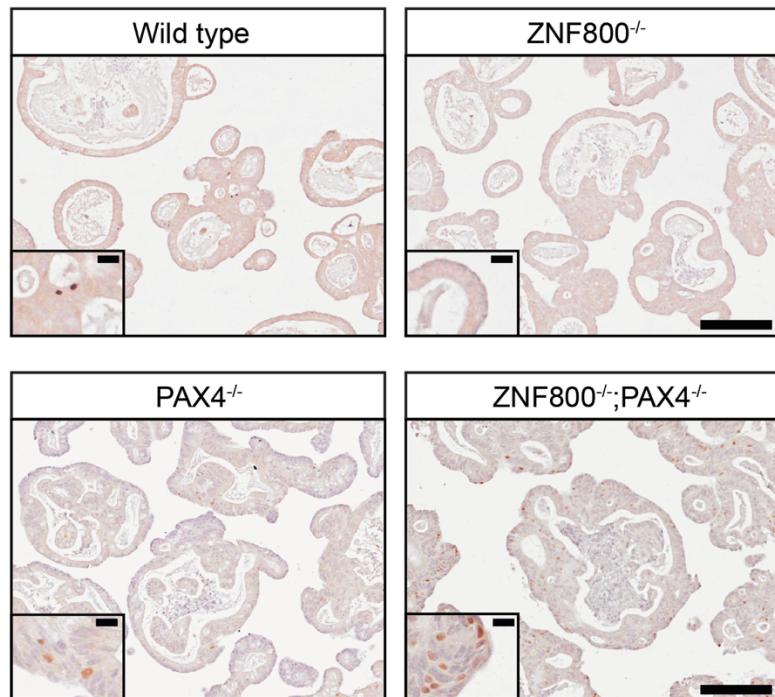


Fig. S19. ARX is re-expressed upon PAX4 knockout

(A) Quantification of ARX positive cells by IHC staining in organoid from different genotypes. Data are shown as mean \pm SEM. ** $P<0.01$, *** $P<0.001$, **** $P<0.0001$ by multiple t -tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with $Q=5\%$, $n=3$. **(B)** Representative IHC staining images of ARX in WT, ZNF800^{-/-}, PAX4^{-/-} and ZNF800^{-/-};PAX4^{-/-} organoids. Scale bars, 200 μ m; Zoom-in scale bars, 20 μ m.

Fig. S20

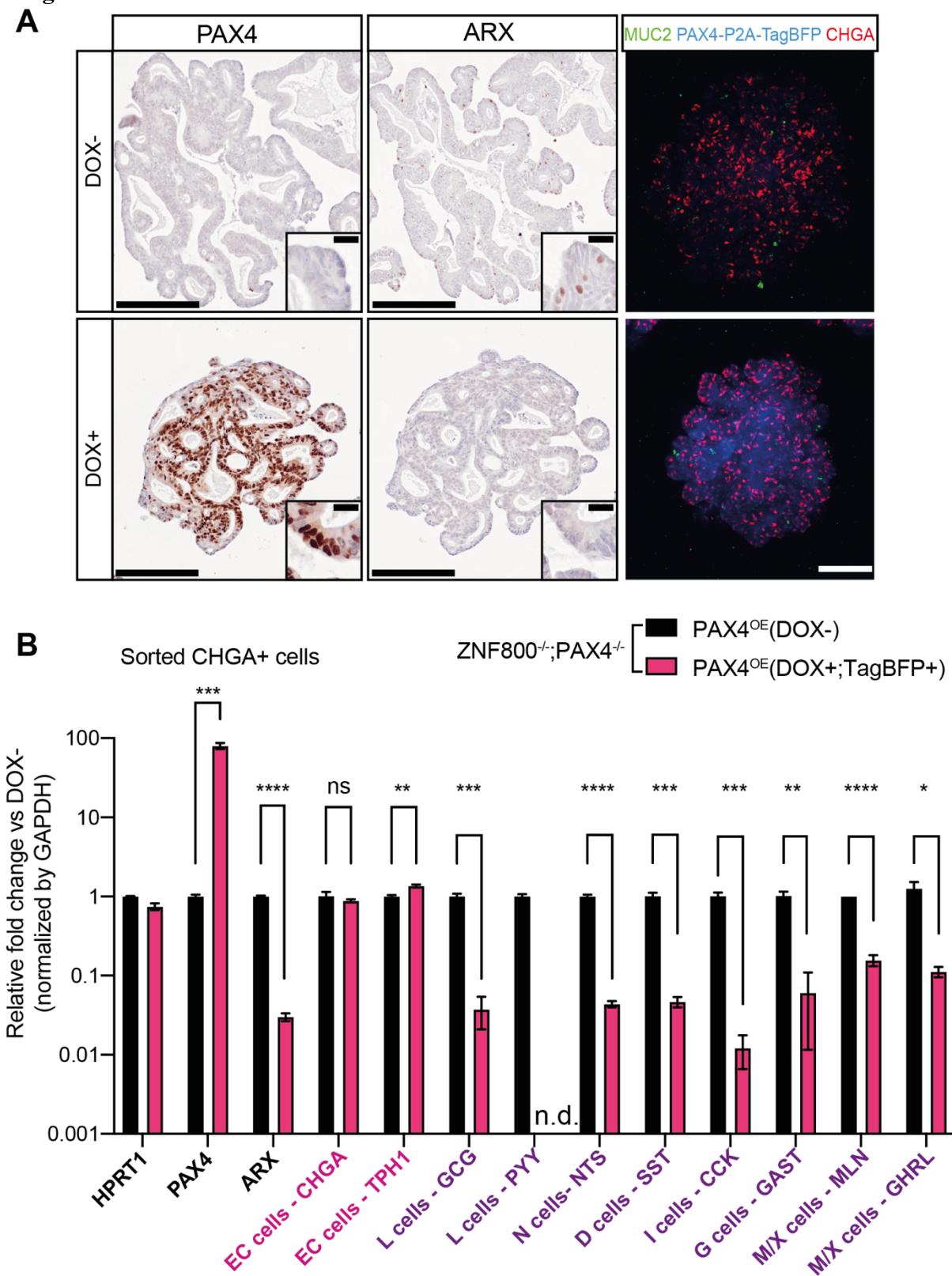


Fig. S20. Re-expression of PAX4 represses ARX expression and reverses EC-biased EEC differentiation

(A) Left, representative IHC staining images of PAX4 and ARX in serial sections of ZNF800^{-/-};PAX4^{-/-} organoids with or without dox-induced PAX4 expression. Scale bars, 200 μ m; Zoom-in scale bars, 20 μ m; Right, representative confocal images of organoids with fluorescent reporters. Scale bars, 200 μ m. **(B)** RT-qPCR quantification of various EEC markers in CHGA+ cell populations of ZNF800^{-/-};PAX4^{-/-} organoids with or without dox-induced PAX4 expression. Data are shown as mean \pm SEM, n.d. not detected. ns, not significant, * $P<0.05$, ** $P<0.01$, *** $P<0.001$, **** $P<0.0001$ by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, *n*=3.

Fig. S21

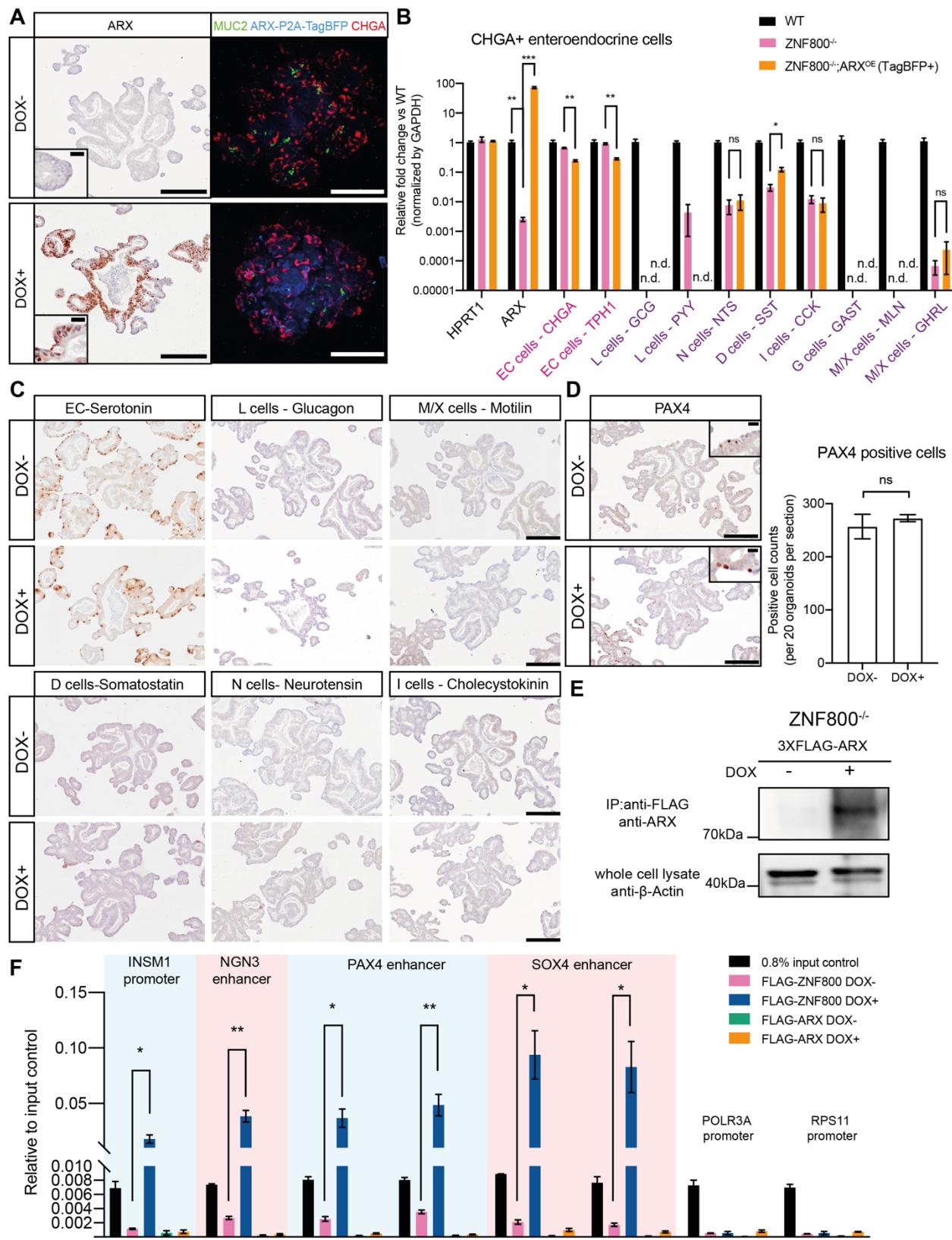


Fig. S21. Re-expression of ARX does not reverse PAX4 upregulation induced by ZNF800^{-/-}

(A) Left, representative IHC staining images of ARX in ZNF800^{-/-} organoids with or without dox-induced ARX expression. Scale bars, 200 μ m; Zoom-in scale bars, 20 μ m; Right, representative confocal images of organoids with fluorescent reporters. Scale bars, 200 μ m. **(B)** RT-qPCR quantification of various EEC markers in CHGA+ cell populations of WT and ZNF800^{-/-} organoids with or without dox-induced ARX expression. Data are shown as mean \pm SEM, n.d. not detected. ns, not significant, * $P<0.05$, ** $P<0.01$, *** $P<0.001$, by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$. **(C)** Representative IHC staining images of serotonin, glucagon, motilin, somatostatin, neuropeptides and cholecystokinin in serial sections of ZNF800^{-/-} organoids with or without dox-induced ARX expression. Scale bars, 200 μ m. **(D)** Left: Representative IHC staining images of PAX4 in ZNF800^{-/-} organoids with or without dox-induced ARX expression. Scale bars, 200 μ m; Zoom-in scale bars, 20 μ m. Right: Quantification of PAX4 positive cells. Data are shown as mean \pm SEM. ns, not significant by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$. **(E)** Top panel: Western blot profiling of the anti-FLAG IP of ARX in ZNF800^{-/-} organoids with or without overexpression of FLAG-tagged ARX. Bottom panel: β -actin served as IP input control. **(F)** ChIP-qPCR profiling of ZNF800 and ARX binding activity in various chromatin loci by anti-FLAG IP in ZNF800^{-/-} organoids. Data are shown as mean \pm SEM, * $P<0.05$, ** $P<0.01$ by multiple *t*-tests using two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q=5%, $n=3$.

Fig. S22

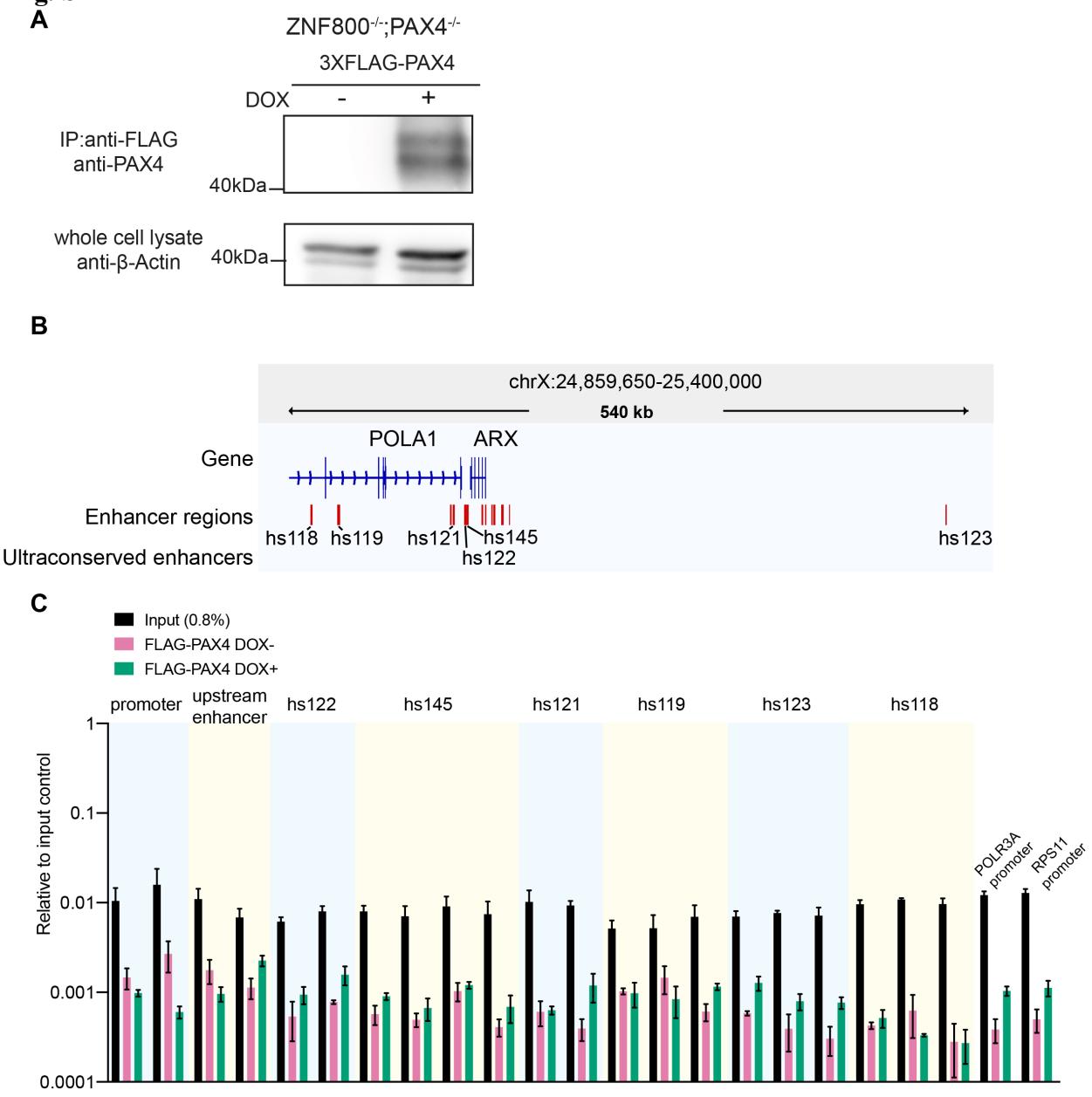


Fig. S22. PAX4 does not bind to ARX enhancer loci

(A) Top panel: Western blot profiling of the anti-FLAG IP of PAX4 in ZNF800^{-/-};PAX4^{-/-} organoids with or without overexpression of FLAG-tagged PAX4. Bottom panel: β -actin served as IP input control. (B) Schematic of promoter, intronic enhancers, upstream enhancers and distal ultraconserved enhancers at ARX loci. (C) ChIP-qPCR profiling of PAX4 binding activity at the ARX loci by anti-FLAG IP in ZNF800^{-/-};PAX4^{-/-} organoids. Data are shown as mean \pm SEM, $n=3$.

Table S1. CRISPR screen MAGeCK analysis Bulk populations vs plasmid pool (separate file).

Table S2. CRISPR screen MAGeCK analysis EEC/Goblet vs triple-reporter negative population (separate file).

Table S3. CRISPR screen MAGeCK analysis EEC vs triple-reporter negative population sgRNA enrichment (separate file).

Table S4. SCENIC Regulon AUC by cell cluster (separate file).

Table S5. Differential gene expression analysis of ZNF800 ChIP binding genes (separate file).

Table S6. Gene ontology analysis of top1000 peaks of ZNF800 binding (separate file).

Table S7. scRNA-seq differential gene expression by cell cluster (separate file).

Table S8. Oligo and primer collection. This table includes all the primers and oligos used in this study (separate file).

References and Notes

1. N. Barker, J. H. van Es, J. Kuipers, P. Kujala, M. van den Born, M. Cozijnsen, A. Haegebarth, J. Korving, H. Begthel, P. J. Peters, H. Clevers, Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003–1007 (2007). [doi:10.1038/nature06196](https://doi.org/10.1038/nature06196) [Medline](#)
2. Y. Sei, X. Lu, A. Liou, X. Zhao, S. A. Wank, A stem cell marker-expressing subset of enteroendocrine cells resides at the crypt base in the small intestine. *Am. J. Physiol. Gastrointest. Liver Physiol.* **300**, G345–G356 (2011). [doi:10.1152/ajpgi.00278.2010](https://doi.org/10.1152/ajpgi.00278.2010) [Medline](#)
3. S. J. A. Buczacki, H. I. Zecchini, A. M. Nicholson, R. Russell, L. Vermeulen, R. Kemp, D. J. Winton, Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* **495**, 65–69 (2013). [doi:10.1038/nature11965](https://doi.org/10.1038/nature11965) [Medline](#)
4. O. Basak, M. van de Born, J. Korving, J. Beumer, S. van der Elst, J. H. van Es, H. Clevers, Mapping early fate determination in Lgr5+ crypt stem cells using a novel Ki67-RFP allele. *EMBO J.* **33**, 2057–2068 (2014). [doi:10.1525/embj.201488017](https://doi.org/10.1525/embj.201488017) [Medline](#)
5. O. Basak, J. Beumer, K. Wiebrands, H. Seno, A. van Oudenaarden, H. Clevers, Induced quiescence of Lgr5+ stem cells in intestinal organoids enables differentiation of hormone-producing enteroendocrine cells. *Cell Stem Cell* **20**, 177–190.e4 (2017). [doi:10.1016/j.stem.2016.11.001](https://doi.org/10.1016/j.stem.2016.11.001) [Medline](#)
6. H. Mutoh, B. P. Fung, F. J. Naya, M.-J. Tsai, J. Nishitani, A. B. Leiter, The basic helix-loop-helix transcription factor BETA2/NeuroD is expressed in mammalian enteroendocrine cells and activates secretin gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3560–3564 (1997). [doi:10.1073/pnas.94.8.3560](https://doi.org/10.1073/pnas.94.8.3560) [Medline](#)
7. F. J. Naya, H.-P. Huang, Y. Qiu, H. Mutoh, F. J. DeMayo, A. B. Leiter, M.-J. Tsai, Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neuroD-deficient mice. *Genes Dev.* **11**, 2323–2334 (1997). [doi:10.1101/gad.11.18.2323](https://doi.org/10.1101/gad.11.18.2323) [Medline](#)
8. M. S. Gierl, N. Karoulias, H. Wende, M. Strehle, C. Birchmeier, The zinc-finger factor Insm1 (IA-1) is essential for the development of pancreatic β cells and intestinal endocrine cells. *Genes Dev.* **20**, 2465–2478 (2006). [doi:10.1101/gad.381806](https://doi.org/10.1101/gad.381806) [Medline](#)
9. S. Desai, Z. Loomis, A. Pugh-Bernard, J. Schrunk, M. J. Doyle, A. Minic, E. McCoy, L. Sussel, Nkx2.2 regulates cell fate choice in the enteroendocrine cell lineages of the intestine. *Dev. Biol.* **313**, 58–66 (2008). [doi:10.1016/j.ydbio.2007.09.047](https://doi.org/10.1016/j.ydbio.2007.09.047) [Medline](#)
10. A. Du, K. W. McCracken, E. R. Walp, N. A. Terry, T. J. Klein, A. Han, J. M. Wells, C. L. May, Arx is required for normal enteroendocrine cell development in mice and humans. *Dev. Biol.* **365**, 175–188 (2012). [doi:10.1016/j.ydbio.2012.02.024](https://doi.org/10.1016/j.ydbio.2012.02.024) [Medline](#)
11. S. Gross, D. C. Garofalo, D. A. Balderes, T. L. Mastracci, J. M. Dias, T. Perlmann, J. Ericson, L. Sussel, Lmx1a functions in intestinal serotonin-producing enterochromaffin cells downstream of Nkx2.2. *Development* **130**, 682 (2013). [doi:10.1242/dev.130682](https://doi.org/10.1242/dev.130682)
12. H. Gehart, J. H. van Es, K. Hamer, J. Beumer, K. Kretzschmar, J. F. Dekkers, A. Rios, H. Clevers, Identification of enteroendocrine regulators by real-time single-cell

differentiation mapping. *Cell* **176**, 1158–1173.e16 (2019). [doi:10.1016/j.cell.2018.12.029](https://doi.org/10.1016/j.cell.2018.12.029) [Medline](#)

13. T. Sato, R. G. Vries, H. J. Snippert, M. van de Wetering, N. Barker, D. E. Stange, J. H. van Es, A. Abo, P. Kujala, P. J. Peters, H. Clevers, Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262–265 (2009). [doi:10.1038/nature07935](https://doi.org/10.1038/nature07935) [Medline](#)
14. T. Sato, D. E. Stange, M. Ferrante, R. G. J. Vries, J. H. Van Es, S. Van den Brink, W. J. Van Houdt, A. Pronk, J. Van Gorp, P. D. Siersema, H. Clevers, Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **141**, 1762–1772 (2011). [doi:10.1053/j.gastro.2011.07.050](https://doi.org/10.1053/j.gastro.2011.07.050) [Medline](#)
15. G. Gradwohl, A. Dierich, M. LeMeur, F. Guillemot, Neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1607–1611 (2000). [doi:10.1073/pnas.97.4.1607](https://doi.org/10.1073/pnas.97.4.1607) [Medline](#)
16. M. Jenny, C. Uhl, C. Roche, I. Duluc, V. Guillermin, F. Guillemot, J. Jensen, M. Kedinger, G. Gradwohl, Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *EMBO J.* **21**, 6338–6347 (2002). [doi:10.1093/emboj/cdf649](https://doi.org/10.1093/emboj/cdf649) [Medline](#)
17. J. Lee, T. Sugiyama, Y. Liu, J. Wang, X. Gu, J. Lei, J. F. Markmann, S. Miyazaki, J. Miyazaki, G. L. Szot, R. Bottino, S. K. Kim, Expansion and conversion of human pancreatic ductal cells into insulin-secreting endocrine cells. *eLife* **2**, e00940 (2013). [doi:10.7554/eLife.00940](https://doi.org/10.7554/eLife.00940) [Medline](#)
18. J. Beumer, J. Puschhof, J. Bauzá-Martinez, A. Martínez-Silgado, R. Elmentaite, K. R. James, A. Ross, D. Hendriks, B. Artegiani, G. A. Busslinger, B. Ponsioen, A. Andersson-Rolf, A. Saftien, C. Boot, K. Kretzschmar, M. H. Geurts, Y. E. Bar-Ephraim, C. Pleguezuelos-Manzano, Y. Post, H. Begthel, F. van der Linden, C. Lopez-Iglesias, W. J. van de Wetering, R. van der Linden, P. J. Peters, A. J. R. Heck, J. Goedhart, H. Snippert, M. Zilbauer, S. A. Teichmann, W. Wu, H. Clevers, High-resolution mRNA and secretome atlas of human enteroendocrine cells. *Cell* **181**, 1291–1306.e19 (2020). [doi:10.1016/j.cell.2020.04.036](https://doi.org/10.1016/j.cell.2020.04.036) [Medline](#)
19. G.-W. He, L. Lin, J. DeMartino, X. Zheng, N. Staliarova, T. Dayton, H. Begthel, W. J. van de Wetering, E. Bodewes, J. van Zon, S. Tans, C. Lopez-Iglesias, P. J. Peters, W. Wu, D. Kotlarz, C. Klein, T. Margaritis, F. Holstege, H. Clevers, Optimized human intestinal organoid model reveals interleukin-22-dependency of paneth cell formation. *Cell Stem Cell* **29**, 1333–1345.e6 (2022). [doi:10.1016/j.stem.2022.08.002](https://doi.org/10.1016/j.stem.2022.08.002) [Medline](#)
20. J. G. Doench, N. Fusi, M. Sullender, M. Hegde, E. W. Vaimberg, K. F. Donovan, I. Smith, Z. Tothova, C. Wilen, R. Orchard, H. W. Virgin, J. Listgarten, D. E. Root, Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016). [doi:10.1038/nbt.3437](https://doi.org/10.1038/nbt.3437) [Medline](#)
21. S. A. Lambert, A. Jolma, L. F. Campitelli, P. K. Das, Y. Yin, M. Albu, X. Chen, J. Taipale, T. R. Hughes, M. T. Weirauch, The human transcription factors. *Cell* **172**, 650–665 (2018). [doi:10.1016/j.cell.2018.01.029](https://doi.org/10.1016/j.cell.2018.01.029) [Medline](#)

22. L. Lin, B. Holmes, M. W. Shen, D. Kammeron, N. Geijsen, D. K. Gifford, R. I. Sherwood, Comprehensive mapping of key regulatory networks that drive oncogene expression. *Cell Rep.* **33**, 108426 (2020). [doi:10.1016/j.celrep.2020.108426](https://doi.org/10.1016/j.celrep.2020.108426) [Medline](#)
23. N. F. Shroyer, D. Wallis, K. J. T. Venken, H. J. Bellen, H. Y. Zoghbi, Gfi1 functions downstream of Math1 to control intestinal secretory cell subtype allocation and differentiation. *Genes Dev.* **19**, 2412–2417 (2005). [doi:10.1101/gad.1353905](https://doi.org/10.1101/gad.1353905) [Medline](#)
24. T. K. Noah, B. Donahue, N. F. Shroyer, Intestinal development and differentiation. *Exp. Cell Res.* **317**, 2702–2710 (2011). [doi:10.1016/j.yexcr.2011.09.006](https://doi.org/10.1016/j.yexcr.2011.09.006) [Medline](#)
25. T. Nakaya, S. Ogawa, I. Manabe, M. Tanaka, M. Sanada, T. Sato, M. M. Taketo, K. Nakao, H. Clevers, M. Fukayama, M. Kuroda, R. Nagai, KLF5 regulates the integrity and oncogenicity of intestinal stem cells. *Cancer Res.* **74**, 2882–2891 (2014). [doi:10.1158/0008-5472.CAN-13-2574](https://doi.org/10.1158/0008-5472.CAN-13-2574) [Medline](#)
26. V. Korinek, N. Barker, P. Moerer, E. van Donselaar, G. Huls, P. J. Peters, H. Clevers, Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nat. Genet.* **19**, 379–383 (1998). [doi:10.1038/1270](https://doi.org/10.1038/1270) [Medline](#)
27. E. Batlle, J. T. Henderson, H. Beghtel, M. M. W. van den Born, E. Sancho, G. Huls, J. Meeldijk, J. Robertson, M. van de Wetering, T. Pawson, H. Clevers, β -catenin and TCF mediate cell positioning in the intestinal epithelium by controlling the expression of EphB/ephrinB. *Cell* **111**, 251–263 (2002). [doi:10.1016/S0092-8674\(02\)01015-2](https://doi.org/10.1016/S0092-8674(02)01015-2) [Medline](#)
28. M. van de Wetering, E. Sancho, C. Verweij, W. de Lau, I. Oving, A. Hurlstone, K. van der Horn, E. Batlle, D. Coudreuse, A.-P. Haramis, M. Tjon-Pon-Fong, P. Moerer, M. van den Born, G. Soete, S. Pals, M. Eilers, R. Medema, H. Clevers, The β -catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**, 241–250 (2002). [doi:10.1016/S0092-8674\(02\)01014-0](https://doi.org/10.1016/S0092-8674(02)01014-0) [Medline](#)
29. J. H. van Es, A. Haegebarth, P. Kujala, S. Itzkovitz, B.-K. Koo, S. F. Boj, J. Korving, M. van den Born, A. van Oudenaarden, S. Robine, H. Clevers, A critical role for the Wnt effector Tcf4 in adult intestinal homeostatic self-renewal. *Mol. Cell. Biol.* **32**, 1918–1927 (2012). [doi:10.1128/MCB.06288-11](https://doi.org/10.1128/MCB.06288-11) [Medline](#)
30. K. Minamide, T. Sato, Y. Nakanishi, H. Ohno, T. Kato, J. Asano, T. Ohteki, IRF2 maintains the stemness of colonic stem cells by limiting physiological stress from interferon. *Sci. Rep.* **10**, 14639 (2020). [doi:10.1038/s41598-020-71633-3](https://doi.org/10.1038/s41598-020-71633-3) [Medline](#)
31. T. Sato, S. Ishikawa, J. Asano, H. Yamamoto, M. Fujii, T. Sato, K. Yamamoto, K. Kitagaki, T. Akashi, R. Okamoto, T. Ohteki, Regulated IFN signalling preserves the stemness of intestinal stem cells by restricting differentiation into secretory-cell lineages. *Nat. Cell Biol.* **22**, 919–926 (2020). [doi:10.1038/s41556-020-0545-5](https://doi.org/10.1038/s41556-020-0545-5) [Medline](#)
32. A. D. Gracz, L. A. Samsa, M. J. Fordham, D. C. Trotier, B. Zwarycz, Y.-H. Lo, K. Bao, J. Starmer, J. R. Raab, N. F. Shroyer, R. L. Reinhardt, S. T. Magness, Sox4 promotes Atoh1-independent intestinal secretory differentiation toward tuft and enteroendocrine fates. *Gastroenterology* **155**, 1508–1523.e10 (2018). [doi:10.1053/j.gastro.2018.07.023](https://doi.org/10.1053/j.gastro.2018.07.023) [Medline](#)
33. M. Bjerknes, H. Cheng, Cell lineage metastability in Gfi1-deficient mouse intestinal epithelium. *Dev. Biol.* **345**, 49–63 (2010). [doi:10.1016/j.ydbio.2010.06.021](https://doi.org/10.1016/j.ydbio.2010.06.021) [Medline](#)

34. H. S. Najafabadi, S. Mnaimneh, F. W. Schmitges, M. Garton, K. N. Lam, A. Yang, M. Albu, M. T. Weirauch, E. Radovani, P. M. Kim, J. Greenblatt, B. J. Frey, T. R. Hughes, C2H2 zinc finger proteins greatly expand the human regulatory lexicon. *Nat. Biotechnol.* **33**, 555–562 (2015). [doi:10.1038/nbt.3128](https://doi.org/10.1038/nbt.3128) [Medline](#)
35. A. Sakaue-Sawano, H. Kurokawa, T. Morimura, A. Hanyu, H. Hama, H. Osawa, S. Kashiwagi, K. Fukami, T. Miyata, H. Miyoshi, T. Imamura, M. Ogawa, H. Masai, A. Miyawaki, Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell* **132**, 487–498 (2008). [doi:10.1016/j.cell.2007.12.033](https://doi.org/10.1016/j.cell.2007.12.033) [Medline](#)
36. S. Aibar, C. B. González-Blas, T. Moerman, V. A. Huynh-Thu, H. Imrichova, G. Hulselmans, F. Rambow, J.-C. Marine, P. Geurts, J. Aerts, J. van den Oord, Z. K. Atak, J. Wouters, S. Aerts, SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017). [doi:10.1038/nmeth.4463](https://doi.org/10.1038/nmeth.4463) [Medline](#)
37. M. H. Sham, C. Vesque, S. Nonchev, H. Marshall, M. Frain, R. D. Gupta, J. Whiting, D. Wilkinson, P. Charnay, R. Krumlauf, The zinc finger gene Krox20 regulates HoxB2 (Hox2.8) during hindbrain segmentation. *Cell* **72**, 183–196 (1993). [doi:10.1016/0092-8674\(93\)90659-E](https://doi.org/10.1016/0092-8674(93)90659-E) [Medline](#)
38. J. Ghislain, C. Desmarquet-Trin-Dinh, P. Gilardi-Hebenstreit, P. Charnay, M. Frain, Neural crest patterning: Autoregulatory and crest-specific elements co-operate for Krox20 transcriptional control. *Development* **130**, 941–953 (2003). [doi:10.1242/dev.00318](https://doi.org/10.1242/dev.00318) [Medline](#)
39. A. Desmazières, P. Charnay, P. Gilardi-Hebenstreit, Krox20 controls the transcription of its various targets in the developing hindbrain according to multiple modes. *J. Biol. Chem.* **284**, 10831–10840 (2009). [doi:10.1074/jbc.M808683200](https://doi.org/10.1074/jbc.M808683200) [Medline](#)
40. S. L. Dunwoodie, D. Henrique, S. M. Harrison, R. S. P. Beddington, Mouse Dll3: A novel divergent Delta gene which may complement the function of other Delta homologues during early pattern formation in the mouse embryo. *Development* **124**, 3065–3076 (1997). [doi:10.1242/dev.124.16.3065](https://doi.org/10.1242/dev.124.16.3065) [Medline](#)
41. R. M. Henke, D. M. Meredith, M. D. Borromeo, T. K. Savage, J. E. Johnson, Ascl1 and Neurog2 form novel complexes and regulate Delta-like3 (Dll3) expression in the neural tube. *Dev. Biol.* **328**, 529–540 (2009). [doi:10.1016/j.ydbio.2009.01.007](https://doi.org/10.1016/j.ydbio.2009.01.007) [Medline](#)
42. B. E. Bernstein, T. S. Mikkelsen, X. Xie, M. Kamal, D. J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S. L. Schreiber, E. S. Lander, A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006). [doi:10.1016/j.cell.2006.02.041](https://doi.org/10.1016/j.cell.2006.02.041) [Medline](#)
43. P. Voigt, W.-W. Tee, D. Reinberg, A double take on bivalent promoters. *Genes Dev.* **27**, 1318–1338 (2013). [doi:10.1101/gad.219626.113](https://doi.org/10.1101/gad.219626.113) [Medline](#)
44. A. B. Osipovich, Q. Long, E. Manduchi, R. Gangula, S. B. Hipkens, J. Schneider, T. Okubo, C. J. Stoeckert Jr., S. Takada, M. A. Magnuson, Insm1 promotes endocrine cell differentiation by modulating the expression of a network of genes that includes Neurog3 and Ripply3. *Development* **141**, 2939–2949 (2014). [doi:10.1242/dev.104810](https://doi.org/10.1242/dev.104810) [Medline](#)
45. A. Beucher, E. Gjernes, C. Collin, M. Courtney, A. Meunier, P. Collombat, G. Gradwohl, The homeodomain-containing transcription factors Arx and Pax4 control enteroendocrine

- subtype specification in mice. *PLOS ONE* **7**, e36449 (2012).
[doi:10.1371/journal.pone.0036449](https://doi.org/10.1371/journal.pone.0036449) [Medline](#)
46. P. Collombat, A. Mansouri, J. Hecksher-Sørensen, P. Serup, J. Krull, G. Gradwohl, P. Gruss, Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes Dev.* **17**, 2591–2603 (2003). [doi:10.1101/gad.269003](https://doi.org/10.1101/gad.269003) [Medline](#)
47. P. Collombat, X. Xu, P. Ravassard, B. Sosa-Pineda, S. Dussaud, N. Billestrup, O. D. Madsen, P. Serup, H. Heimberg, A. Mansouri, The ectopic expression of Pax4 in the mouse pancreas converts progenitor cells into α and subsequently β cells. *Cell* **138**, 449–462 (2009). [doi:10.1016/j.cell.2009.05.035](https://doi.org/10.1016/j.cell.2009.05.035) [Medline](#)
48. G. Bejerano, M. Pheasant, I. Makunin, S. Stephen, W. J. Kent, J. S. Mattick, D. Haussler, Ultraconserved elements in the human genome. *Science* **304**, 1321–1325 (2004).
[doi:10.1126/science.1098119](https://doi.org/10.1126/science.1098119) [Medline](#)
49. P. Collombat, J. Hecksher-Sørensen, V. Broccoli, J. Krull, I. Ponte, T. Mundiger, J. Smith, P. Gruss, P. Serup, A. Mansouri, The simultaneous loss of Arx and Pax4 genes promotes a somatostatin-producing cell fate specification at the expense of the α - and β -cell lineages in the mouse endocrine pancreas. *Development* **132**, 2969–2980 (2005).
[doi:10.1242/dev.01870](https://doi.org/10.1242/dev.01870) [Medline](#)
50. D. E. Dickel, A. R. Ypsilanti, R. Pla, Y. Zhu, I. Barozzi, B. J. Mannion, Y. S. Khin, Y. Fukuda-Yuzawa, I. Plajzer-Frick, C. S. Pickle, E. A. Lee, A. N. Harrington, Q. T. Pham, T. H. Garvin, M. Kato, M. Osterwalder, J. A. Akiyama, V. Afzal, J. L. R. Rubenstein, L. A. Pennacchio, A. Visel, Ultraconserved enhancers are required for normal development. *Cell* **172**, 491–499.e15 (2018). [doi:10.1016/j.cell.2017.12.017](https://doi.org/10.1016/j.cell.2017.12.017) [Medline](#)
51. R. Elmentait, N. Kumasaka, K. Roberts, A. Fleming, E. Dann, H. W. King, V. Kleshchevnikov, M. Dabrowska, S. Pritchard, L. Bolt, S. F. Vieira, L. Mamanova, N. Huang, F. Perrone, I. Goh Kai’En, S. N. Lisgo, M. Katan, S. Leonard, T. R. W. Oliver, C. E. Hook, K. Nayak, L. S. Campos, C. Domínguez Conde, E. Stephenson, J. Engelbert, R. A. Botting, K. Polanski, S. van Dongen, M. Patel, M. D. Morgan, J. C. Marioni, O. A. Bayraktar, K. B. Meyer, X. He, R. A. Barker, H. H. Uhlig, K. T. Mahbubani, K. Saeb-Parsy, M. Zilbauer, M. R. Clatworthy, M. Haniffa, K. R. James, S. A. Teichmann, Cells of the human intestinal tract mapped across space and time. *Nature* **597**, 250–255 (2021).
[doi:10.1038/s41586-021-03852-1](https://doi.org/10.1038/s41586-021-03852-1) [Medline](#)
52. J. Burclaff, R. J. Bliton, K. A. Breau, M. T. Ok, I. Gomez-Martinez, J. S. Ranek, A. P. Bhatt, J. E. Purvis, J. T. Woosley, S. T. Magness, A proximal-to-distal survey of healthy adult human small intestine and colon epithelium by single-cell transcriptomics. *Cell. Mol. Gastroenterol. Hepatol.* **13**, 1554–1589 (2022). [doi:10.1016/j.jcmgh.2022.02.007](https://doi.org/10.1016/j.jcmgh.2022.02.007) [Medline](#)
53. R. Elmentait, A. D. B. Ross, K. Roberts, K. R. James, D. Ortmann, T. Gomes, K. Nayak, L. Tuck, S. Pritchard, O. A. Bayraktar, R. Heuschkel, L. Vallier, S. A. Teichmann, M. Zilbauer, Single-cell sequencing of developing human gut reveals transcriptional links to childhood Crohn’s disease. *Dev. Cell* **55**, 771–783.e5 (2020).
[doi:10.1016/j.devcel.2020.11.010](https://doi.org/10.1016/j.devcel.2020.11.010) [Medline](#)
54. A. B. Osipovich, K. D. Dudek, E. Greenfest-Allen, J.-P. Cartailler, E. Manduchi, L. Potter Case, E. Choi, A. G. Chapman, H. W. Clayton, G. Gu, C. J. Stoeckert Jr., M. A.

- Magnuson, A developmental lineage-based gene co-expression network for mouse pancreatic β -cells reveals a role for *Zfp800* in pancreas development. *Development* **148**, dev196964 (2021). [doi:10.1242/dev.196964](https://doi.org/10.1242/dev.196964) [Medline](#)
55. ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012). [doi:10.1038/nature11247](https://doi.org/10.1038/nature11247) [Medline](#)
56. Y. Luo, B. C. Hitz, I. Gabdank, J. A. Hilton, M. S. Kagda, B. Lam, Z. Myers, P. Sud, J. Jou, K. Lin, U. K. Baymuradov, K. Graham, C. Litton, S. R. Miyasato, J. S. Strattan, O. Jolanki, J.-W. Lee, F. Y. Tanaka, P. Adenekan, E. O'Neill, J. M. Cherry, New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. *Nucleic Acids Res.* **48** (D1), D882–D889 (2020). [doi:10.1093/nar/gkz1062](https://doi.org/10.1093/nar/gkz1062) [Medline](#)
57. M. van de Wetering, H. E. Francies, J. M. Francis, G. Bounova, F. Iorio, A. Pronk, W. van Houdt, J. van Gorp, A. Taylor-Weiner, L. Kester, A. McLaren-Douglas, J. Blokker, S. Jaksani, S. Bartfeld, R. Volckman, P. van Sluis, V. S. W. Li, S. Seepo, C. Sekhar Pedamallu, K. Cibulskis, S. L. Carter, A. McKenna, M. S. Lawrence, L. Lichtenstein, C. Stewart, J. Koster, R. Versteeg, A. van Oudenaarden, J. Saez-Rodriguez, R. G. J. Vries, G. Getz, L. Wessels, M. R. Stratton, U. McDermott, M. Meyerson, M. J. Garnett, H. Clevers, Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell* **161**, 933–945 (2015). [doi:10.1016/j.cell.2015.03.053](https://doi.org/10.1016/j.cell.2015.03.053) [Medline](#)
58. M. Arbab, S. Srinivasan, T. Hashimoto, N. Geijsen, R. I. Sherwood, Cloning-free CRISPR. *Stem Cell Reports* **5**, 908–917 (2015). [doi:10.1016/j.stemcr.2015.09.022](https://doi.org/10.1016/j.stemcr.2015.09.022) [Medline](#)
59. M. Fujii, M. Matano, K. Nanki, T. Sato, Efficient genetic engineering of human intestinal organoids using electroporation. *Nat. Protoc.* **10**, 1474–1485 (2015). [doi:10.1038/nprot.2015.088](https://doi.org/10.1038/nprot.2015.088) [Medline](#)
60. A. H. M. Ng, P. Khoshakhlagh, J. E. Rojo Arias, G. Pasquini, K. Wang, A. Swiersy, S. L. Shipman, E. Appleton, K. Kiaee, R. E. Kohman, A. Vernet, M. Dysart, K. Leeper, W. Saylor, J. Y. Huang, A. Graveline, J. Taipale, D. E. Hill, M. Vidal, J. M. Melero-Martin, V. Busskamp, G. M. Church, A comprehensive library of human transcription factors for cell fate engineering. *Nat. Biotechnol.* **39**, 510–519 (2021). [doi:10.1038/s41587-020-0742-6](https://doi.org/10.1038/s41587-020-0742-6) [Medline](#)
61. M. W. Shen, M. Arbab, J. Y. Hsu, D. Worstell, S. J. Culbertson, O. Krabbe, C. A. Cassa, D. R. Liu, D. K. Gifford, R. I. Sherwood, Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature* **563**, 646–651 (2018). [doi:10.1038/s41586-018-0686-x](https://doi.org/10.1038/s41586-018-0686-x) [Medline](#)
62. T. Wang, J. J. Wei, D. M. Sabatini, E. S. Lander, Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 [doi:10.1126/science.1246981](https://doi.org/10.1126/science.1246981) (2014). [doi:10.1126/science.1246981](https://doi.org/10.1126/science.1246981) [Medline](#)
63. W. Li, H. Xu, T. Xiao, L. Cong, M. I. Love, F. Zhang, R. A. Irizarry, J. S. Liu, M. Brown, X. S. Liu, MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014). [doi:10.1186/s13059-014-0554-4](https://doi.org/10.1186/s13059-014-0554-4) [Medline](#)

64. C.-H. Chen, T. Xiao, H. Xu, P. Jiang, C. A. Meyer, W. Li, M. Brown, X. S. Liu, Improved design and analysis of CRISPR knockout screens. *Bioinformatics* **34**, 4095–4101 (2018). [doi:10.1093/bioinformatics/bty450](https://doi.org/10.1093/bioinformatics/bty450) [Medline](#)
65. S. Choudhary, R. Satija, Comparison and evaluation of statistical error models for scRNA-seq. *Genome Biol.* **23**, 27 (2022). [doi:10.1186/s13059-021-02584-9](https://doi.org/10.1186/s13059-021-02584-9) [Medline](#)
66. B. Van de Sande, C. Flerin, K. Davie, M. De Waegeneer, G. Hulselmans, S. Aibar, R. Seurinck, W. Saelens, R. Cannoodt, Q. Rouchon, T. Verbeiren, D. De Maeyer, J. Reumers, Y. Saeys, S. Aerts, A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* **15**, 2247–2276 (2020). [doi:10.1038/s41596-020-0336-2](https://doi.org/10.1038/s41596-020-0336-2) [Medline](#)
67. W. X. Tan, C. M. Bok, N. H. J. Ng, A. K. K. Teo, “Chromatin Immunoprecipitation in Human Pluripotent Stem Cell-Derived 3D Organoids to Analyze DNA–Protein Interactions” in *Stem Cell Assays, Methods in Molecular Biology*, vol. 2429, N. Kannan, P. Beer, Eds. (Humana, 2022), pp. 215–232; https://doi.org/10.1007/978-1-0716-1979-7_14.
68. S. Andrews, Q. C. Fast, A quality control tool for high throughput sequence data (2010); <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
69. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012). [doi:10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923) [Medline](#)
70. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021). [doi:10.1093/gigascience/giab008](https://doi.org/10.1093/gigascience/giab008) [Medline](#)
71. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, X. S. Liu, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008). [doi:10.1186/gb-2008-9-9-r137](https://doi.org/10.1186/gb-2008-9-9-r137) [Medline](#)
72. F. Ramírez, D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016). [doi:10.1093/nar/gkw257](https://doi.org/10.1093/nar/gkw257) [Medline](#)
73. J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011). [doi:10.1038/nbt.1754](https://doi.org/10.1038/nbt.1754) [Medline](#)
74. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010). [doi:10.1093/bioinformatics/btq033](https://doi.org/10.1093/bioinformatics/btq033) [Medline](#)
75. G. Yu, L.-G. Wang, Q.-Y. He, ChIPseeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015). [doi:10.1093/bioinformatics/btv145](https://doi.org/10.1093/bioinformatics/btv145) [Medline](#)
76. Q. Wang, M. Li, T. Wu, L. Zhan, L. Li, M. Chen, W. Xie, Z. Xie, E. Hu, S. Xu, G. Yu, Exploring Epigenomic Datasets by ChIPseeker. *Curr. Protoc.* **2**, e585 (2022). [doi:10.1002/cpz1.585](https://doi.org/10.1002/cpz1.585) [Medline](#)