Stem Cell Reports

Supplemental Information

Distinct Processes and Transcriptional Targets Underlie CDX2 Requirements in Intestinal Stem Cells and Differentiated Villus Cells

Adrianna K. San Roman, Alessio Tovaglieri, David T. Breault, and Ramesh A. Shivdasani

SUPPLEMENTAL FIGURES

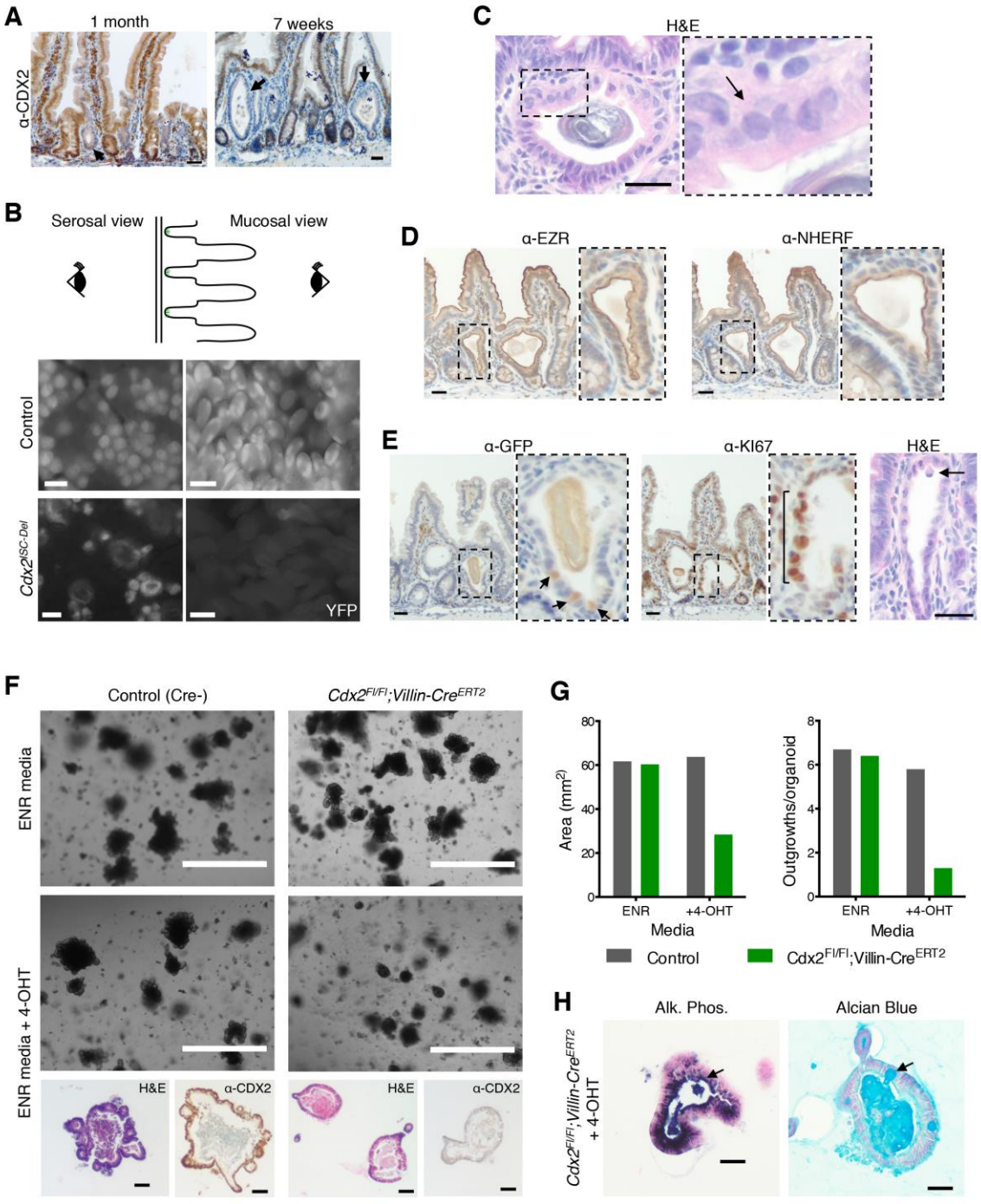
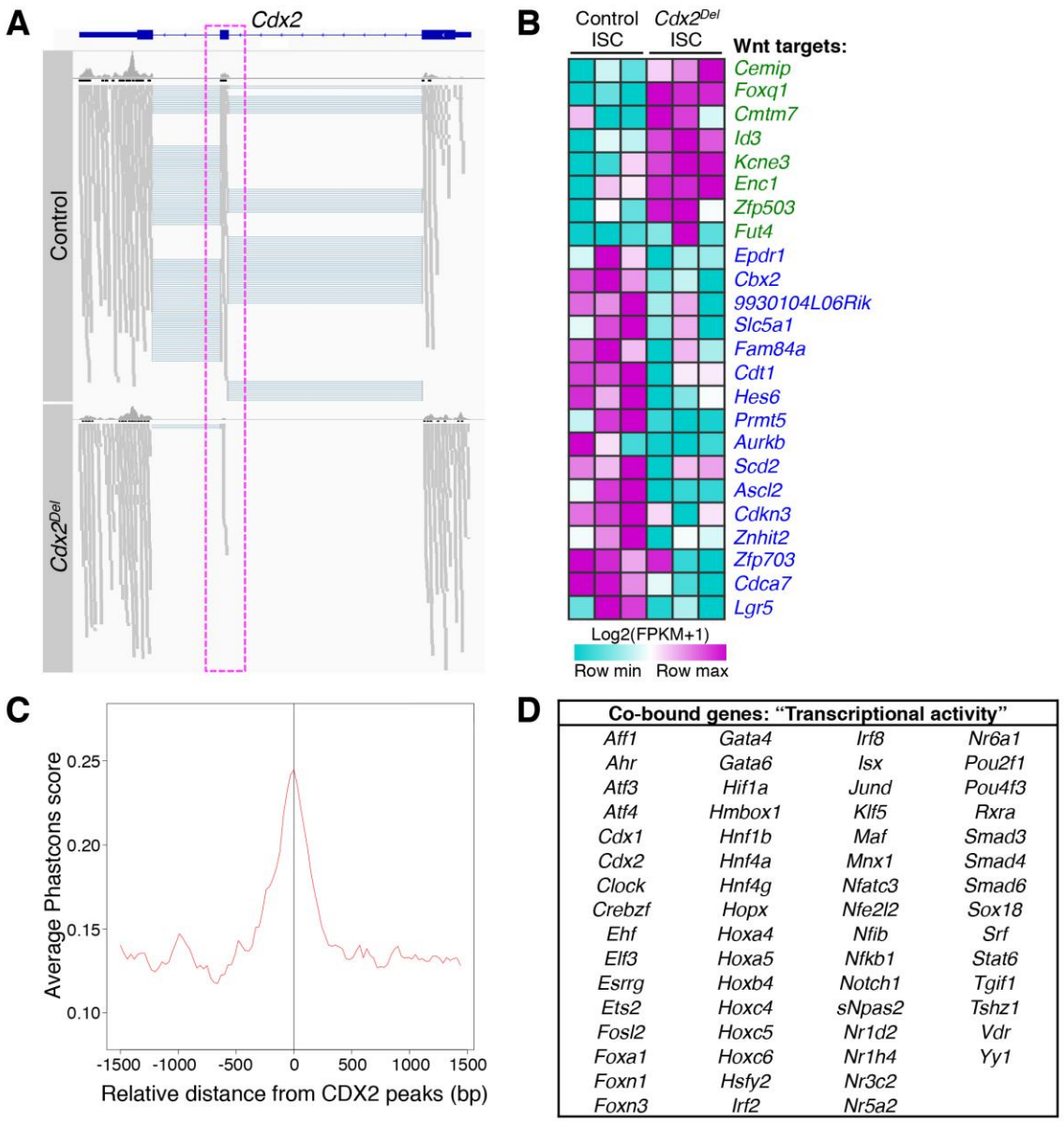


Figure S1. *Cdx2*^{Del} crypts form epithelial cysts in vivo and fail to generate mini-gut organoids in culture, related to Figure 2. (A) CDX2 immunostaining of *Cdx2*$^{ISC-Del}$ intestines 1 month and 7 weeks after *Cdx2* deletion demonstrates earliest cyst formations (arrows), which become more abundant in the subsequent few weeks. Scale bars, 30 μm. **(B)** Whole-mount views of lineage-traced control and *Cdx2*$^{ISC-Del}$ intestines 3 months after *Cdx2* deletion and simultaneous YFP activation. Views from the serosal side of the duodenum (left images) reveal

YFP^{+} halo-like structures that correspond to *Cdx2^{Del}* cysts and contrast with YFP^{+} crypts in control intestines. Abundant ribbons of YFP+ villi seen in mucosal views (right images) of control intestines were never observed in *Cdx2$^{ISC\text{-}Del}$* mice. Scale bars, 50 μm. **(C)** Hematoxylin and eosin (H&E) staining of a representative *Cdx2^{Del}* cyst shows cuboidal cell morphology. The boxed area is magnified to the right, where the arrow points to abnormally round nuclear shape. **(D)** EZR and NHERF immunostaining reveals normal apical (luminal) localization in *Cdx2$^{ISC\text{-}Del}$* cysts, indicating that apico-basal polarity is preserved; dashed area magnified to right. **(E)** In *Cdx2$^{ISC\text{-}Del}$* cysts, LGR5-GFP^{+} cells tend to localize at the base (arrows) and KI67^{+} cells on the sides (bracket), both away from upper areas where cells are extruded into the cyst lumen (arrows). Boxed areas in each panel are magnified to the right of the respective images. Scale bars: GFP and KI67, 30 μm; H&E, 50 μm. **(F)** Organoids from control (Cre-) and *Cdx2$^{Fl/Fl}$;Villin-Cre^{ERT2}* small intestine crypts in normal EGF, Noggin, R-spondin organoid growth medium (ENR) or supplemented with 4-hydroxytamoxifen (4-OHT) from 0-2 days. Scale bars, 1 mm. Bottom, Hematoxylin and Eosin staining of 4-OHT treated mini-guts reveal small, round *Cdx2^{Del}* organoids with few if any crypt outgrowths. Immunohistochemistry for CDX2 confirms loss in *Cdx2$^{Fl/Fl}$;Villin-Cre^{ERT2}* organoids, scale bars, 100 μm. **(G)** Untreated *Cdx2$^{Fl/Fl}$;Villin-Cre^{ERT2}* organoids were comparable to control, while organoids treated with 4-OHT were smaller and had fewer outgrowths. At least 30 organoids from a representative mouse per condition were quantified. **(H)** Staining for mature markers of absorptive enterocytes (alkaline phosphatase) and goblet cells (alcian blue) reveals presence of mature cell types in remaining *Cdx2$^{Fl/Fl}$;Villin-Cre^{ERT2}* organoids treated with 4-OHT. Scale bars, 30 μm.



Co-bound genes: "Transcriptional activity"			
Aff1	*Gata4*	*Irf8*	*Nr6a1*
Ahr	*Gata6*	*Isx*	*Pou2f1*
Atf3	*Hif1a*	*Jund*	*Pou4f3*
Atf4	*Hmbox1*	*Klf5*	*Rxra*
Cdx1	*Hnf1b*	*Maf*	*Smad3*
Cdx2	*Hnf4a*	*Mnx1*	*Smad4*
Clock	*Hnf4g*	*Nfatc3*	*Smad6*
Crebzf	*Hopx*	*Nfe2l2*	*Sox18*
Ehf	*Hoxa4*	*Nfib*	*Srf*
Elf3	*Hoxa5*	*Nfkb1*	*Stat6*
Esrrg	*Hoxb4*	*Notch1*	*Tgif1*
Ets2	*Hoxc4*	*sNpas2*	*Tshz1*
Fosl2	*Hoxc5*	*Nr1d2*	*Vdr*
Foxa1	*Hoxc6*	*Nr1h4*	*Yy1*
Foxn1	*Hsfy2*	*Nr3c2*	
Foxn3	*Irf2*	*Nr5a2*	

Figure S2. Validation and analysis of RNA-seq and CDX2 ChIP-seq in ISCs, related to Figures 3 and 4. (A) Aligned RNA-seq reads in control and *Cdx2*^{Del} ISCs confirms excision of *Cdx2* exon 2 (magenta box). Grey bars represent single RNA-seq reads and teal lines indicate reads that span two exons. **(B)** Heatmap of significantly down- (blue) and up- (green) regulated Wnt target genes from RNA-seq of *Cdx2*^{Del} vs. control ISCs. **(C)** Plot of average PhastCons score around CDX2 binding sites identified in ISCs, indicating their evolutionary conservation. **(D)** Amongst the genes that are bound by CDX2 within 30 kb in both ISCs and Villus cells, genes related to the Gene Ontology term 'Transcriptional activity' are enriched, listed alphabetically.

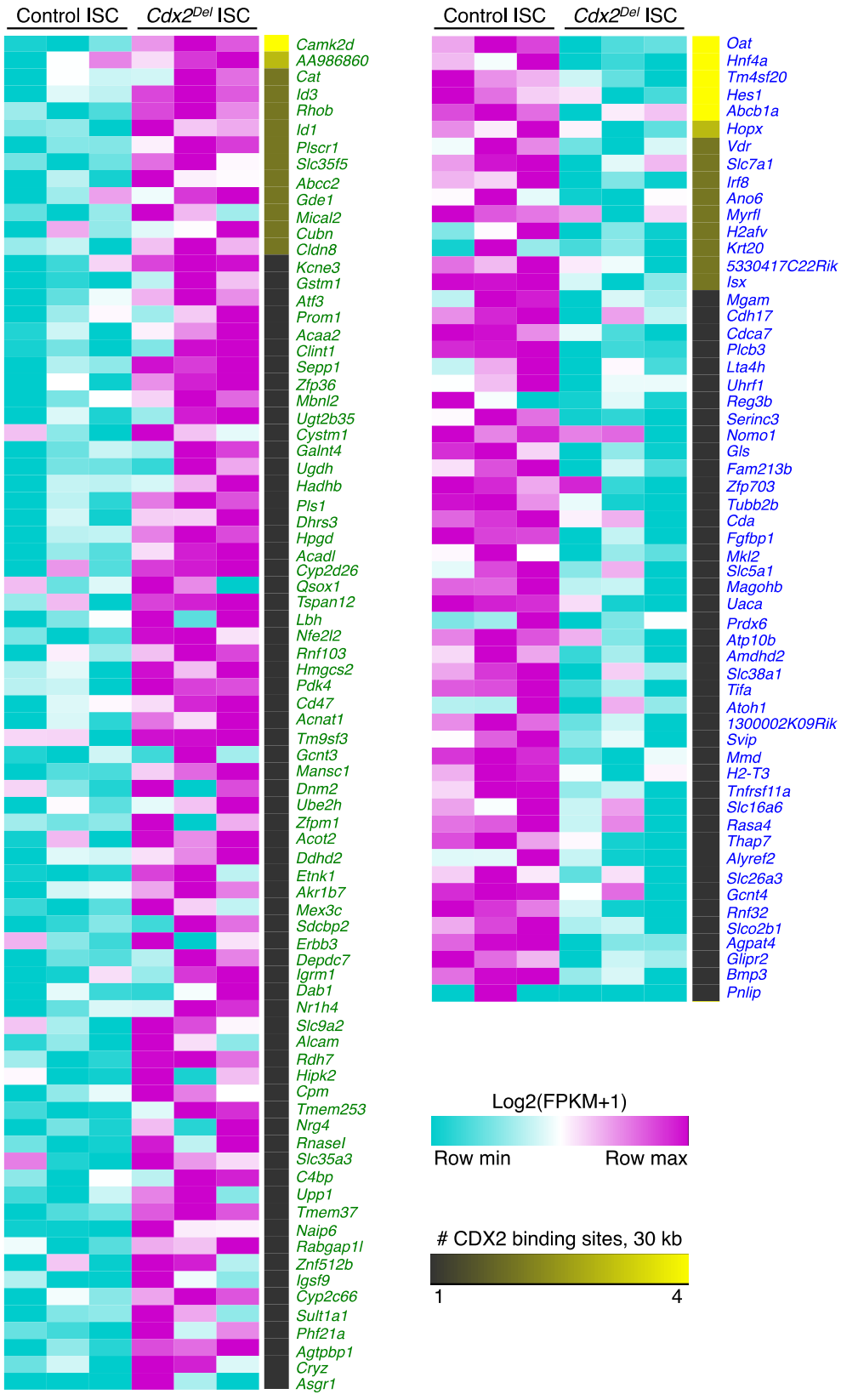

Control ISC
Cdx2Del ISC
Camk2d
AA986860
Cat
Id3
Rhob
Id1
Plscr1
Slc35f5
Abcc2
Gde1
Mical2
Cubn
Cldn8
Kcne3
Gstm1
Atf3
Prom1
Acaa2
Clint1
Sepp1
Zfp36
Mbnl2
Ugt2b35
Cystm1
Galnt4
Ugdh
Hadhb
Pls1
Dhrs3
Hpgd
Acadl
Cyp2d26
Qsox1
Tspan12
Lbh
Nfe2l2
Rnf103
Hmgcs2
Pdk4
Cd47
Acnat1
Tm9sf3
Gcnt3
Mansc1
Dnm2
Ube2h
Zfpm1
Acot2
Ddhd2
Etnk1
Akr1b7
Mex3c
Sdcbp2
Erbb3
Depdc7
Igrm1
Dab1
Nr1h4
Slc9a2
Alcam
Rdh7
Hipk2
Cpm
Tmem253
Nrg4
RnaseI
Slc35a3
C4bp
Upp1
Tmem37
Naip6
Rabgap1l
Znf512b
Igsf9
Cyp2c66
Sult1a1
Phf21a
Agtpbp1
Cryz
Asgr1
Control ISC
Cdx2Del ISC
Oat
Hnf4a
Tm4sf20
Hes1
Abcb1a
Hopx
Vdr
Slc7a1
Irf8
Ano6
Myrfl
H2afv
Krt20
5330417C22Rik
Isx
Mgam
Cdh17
Cdca7
Plcb3
Lta4h
Uhrf1
Reg3b
Serinc3
Nomo1
Gls
Fam213b
Zfp703
Tubb2b
Cda
Fgfbp1
Mkl2
Slc5a1
Magohb
Uaca
Prdx6
Atp10b
Amdhd2
Slc38a1
Tifa
Atoh1
1300002K09Rik
Svip
Mmd
H2-T3
Tnfrsf11a
Slc16a6
Rasa4
Thap7
Alyref2
Slc26a3
Gcnt4
Rnf32
Slco2b1
Agpat4
Glipr2
Bmp3
Pnlip
Log2(FPKM+1)
Row min
Row max
CDX2 binding sites, 30 kb
1
4

Figure S3. Candidate direct CDX2 target genes in ISCs, related to Figure 4. Heatmap of RNA-seq data comparing control and *Cdx2*^{Del} ISCs. Significantly up (left; green) and down-regulated (right; blue) genes are listed with the corresponding number of nearby CDX2 binding sites in ISCs represented by the black to yellow heatmaps to the right.

SUPPLEMENTAL TABLES

Table S1. Gastric marker gene expression in *Cdx2*^{Del} ISCs, compared to control, with nearby CDX2 binding, Related to Figure 3.

Gene	Log2 Fold-Change *Cdx2*^{Del} vs Control ISC	Q-value	ISC binding#	Villus binding#
*Atp4a**	2.4	0.0012	0	0
Cldn18	FPKM<1	N/A	0	0
Col11a2	FPKM<1	N/A	0	0
Gast	FPKM<1	N/A	0	0
Gif	0.44	0.51	0	0
Muc1	FPKM<1	N/A	0	0
Muc5ac	FPKM<1	N/A	0	0
Muc6	-0.97	0.35	0	0
*Pgc**	4.3	0.00024	0	1
Shh	-0.21	0.84	0	0
Tff2	2.65	0.33	0	0

*Genes significantly increased in *Cdx2*^{Del} ISCs.
#Number of CDX2 peaks within 30 kb of the transcription start site

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Mouse strains, conditional gene disruption, and genetic lineage tracing. *Cdx2*$^{Fl/Fl}$ mice (Verzi et al., 2010) were crossed with either *Lgr5-EGFP-IRES-Cre*^{ERT2} or *Villin-Cre*^{ERT2} mice (Barker et al., 2007; el Marjou et al., 2004) to generate ISC-specific or whole-epithelium conditional *Cdx2* knockout mice, respectively. To isolate ISCs, *Cdx2*$^{Fl/Fl}$*;Lgr5-EGFP-IRES-Cre*^{ERT2} and control (*Cdx2*$^{+/+}$*;Lgr5-EGFP-IRES-Cre*^{ERT2}) mice were injected intraperitoneally (IP) with 2 mg tamoxifen (TAM) for 4 consecutive days and LGR5-GFP^{HI} cells, were harvested the next day by fluorescence-activated cell sorting. To induce intestinal epithelial *Cdx2* loss, we injected *Cdx2*$^{Fl/Fl}$*;Villin-Cre*^{ERT2} and control (*Cdx2*$^{Fl/Fl}$*;*Cre-negative) mice with 1 mg TAM for 5 days and harvested intestines 3 days later. For lineage tracing experiments, we crossed *Cdx2*$^{Fl/Fl}$*;Lgr5-EGFP-IRES-Cre*^{ERT2} or control (*Cdx2*$^{+/+}$*;Lgr5-EGFP-IRES-Cre*^{ERT2}) mice with the ubiquitously expressed *Rosa26-lox-STOP-lox-YFP* Cre reporter line (Srinivas et al., 2001). To begin lineage tracing, we injected *Cdx2*$^{Fl/Fl}$*;Lgr5-EGFP-IRES-CreERT2;Rosa26-lox-STOP-lox-*

YFP and control (*Cdx2*$^{+/+}$*;Lgr5-EGFP-IRES-CreERT2;Rosa26-lox-STOP-lox-YFP*) mice with 2 mg TAM for 5 days and harvested intestines 10 days, 1 month, 3 months, and 6 months after the first injection.

Isolation of Lgr5+ ISCs, crypts, and villi. We flushed the proximal 2/3 small bowel of *Lgr5-EGFP-IRES-CreERT2* mice (where we observe a much higher fraction of GFP+ crypts than the distal intestine) with cold phosphate-buffered saline (PBS), opened the intestine, gently scraped the surface with a glass slide to remove most villi, and incubated the tissue in 5 mM EDTA in PBS at 4°C for 30 min on a shaker with periodic shaking by hand. Tissues were transferred to fresh 5 mM EDTA in PBS at 4°C for an additional 15 min and shaken by hand for 2 min to detach crypts. Epithelium pooled from both steps was disaggregated with 4X TrypLE (Invitrogen) in dye-free Dulbecco's Modified Eagle Medium (Mediatech) at 37°C for 45 min on a rotating platform. Cells were washed in PBS, stained with Live/Dead Cell Viability Dye (Invitrogen), and sorted on a FACSAria III instrument (BD Biosciences) to collect single, live, LGR5-GFP^{HI} cells.

Cell cycle profiling of ISCs. ISC isolation protocol was followed with the addition of 10 μg/mL Hoechst dye (Invitrogen) during the TrypLE disaggregation step. Cell cycle was analyzed 3 times on different cell isolates by gating on single, live LGR5-GFP^{HI} cells, followed by detection of Hoechst Dye with a UV laser. Data were analyzed with FlowJo software (Tree Star) using the Watson Pragmatic Test to determine cell cycle phase distribution.

RNA-seq and analysis. For ISCs, ≥ 10^5 cells were collected from the proximal 2/3 small intestine by flow cytometry from single mice. mRNA was extracted using oligo $d(T)_{25}$ magnetic beads (New England Biolabs), followed by DNase treatment with the RNeasy kit (Qiagen), and was quantified using Ribogreen (Life Technologies). 0.5 ng mRNA was used for cDNA synthesis and library preparation using the Encore Complete RNA-Seq Library System (Nugen), modified for ½ volume reactions. After isolating 200-450 base pair (bp) fragments using Pippin Prep (Sage Science), 50 bp, single-end reads were sequenced using the Illumina Hi-Seq 2000 instrument.

For villus cells, total RNA was isolated from small intestine villi using TRIzol reagent and the RNeasy kit, and genomic DNA was removed with the Turbo DNA-free kit (Ambion). RNA-sequencing libraries were prepared with the TruSeq RNA Sample Preparation Kit (Illumina) according to the manufacturers specifications. 75 bp single-end reads were sequenced on an Illumina NextSeq 500 instrument.

The Tuxedo software package was used to align reads to the *Mus musculus* reference genome build 9 (mm9), assemble transcripts, and determine differences in transcript levels using a false discovery rate of 0.05 (Trapnell et al., 2012). Three independent samples were sequenced and compared for each condition. To adjust for 3' bias in the samples, we calculated FPKM values using the 3' tag counting method (Sigurgeirsson et al., 2014), in which reads were aligned to the 3' 1000 base pairs of the most abundant gene isoform transcript. A threshold of FPKM ≥ 1 was applied to consider a gene "expressed". All fold-changes were calculated by adding 1 to the FPKM value of both samples (to avoid dividing by 0) and the log2 value was calculated. Notably, our results

on wild-type ISCs were highly concordant with reported expression data (Munoz et al., 2012; Sheaffer et al., 2014).

Average Log2(FPKM+1) expression data was plotted in Graphpad Prism and heatmaps of expression data were generated with normalized Log2(FPKM+1) expression values for each sample in Gene-E software (http://www.broadinstitute.org/cancer/software/GENE-E/).

Gene Ontology (GO) analysis. Significantly dysregulated gene lists were analyzed using DAVID functional annotation clustering software (default options, medium stringency) to identify significantly enriched clusters (scores > 1.3) (Huang da et al., 2009). Representative GO terms are listed with the corresponding enrichment scores.

Histology, immunofluorescence and immunohistochemistry. Intestine segments were opened and fixed overnight in 4% paraformaldehyde at 4°C, washed in PBS, dehydrated in 70% ethanol, and processed for paraffin embedding as "swiss rolls". Staining with hematoxylin and eosin (H&E), alkaline phosphatase, alcian blue, and various antibodies (Ab) followed standard procedures. For immunohistochemistry, tissues were pre-treated in a pressure cooker in 10 mM sodium citrate buffer (pH 6) for antigen retrieval followed by incubation in methanol with 0.5% H_2O_2 to block endogenous peroxidases. Tissues were blocked in 5% fetal bovine serum and incubated overnight at 4°C with one of the following antibodies:

Name	Species	Dilution	Company	Catalog #
CDX2	Mouse	1:20	Biogenex	MU392A-UC
KI67	Mouse	1:1000	Vector Labs	VP-K452
GFP	Mouse	1:100	Santa Cruz	SC-9996
EZR	Mouse	1:100	Neomarker	MS-661-P1
NHERF1	Rabbit	1:100	Abcam	Ab3452

After washing, tissue sections were treated with species-specific, biotin conjugated anti-IgG secondary Ab (1:300, Vector Laboratories). Antigens were detected using the Vectastain Elite ABC Kit (Vector Laboratories) with diaminobenzidine (Sigma) substrate and nuclei were counterstained with Harris' Hematoxylin (Electron Microscopy Science).

Quantification of KI67+ ISCs and TA cells was performed on at least 4 mice by counting 30-50 crypts. Only crypts with contiguous epithelium from the crypt base to the villus where Paneth cells were clearly visible were considered. Crypt-base ISCs were defined as cells between Paneth cells, while TA cells are above Paneth cells from the +5 position to the crypt-villus junction. The average number of KI67+ CBCs or TA cells per crypt was calculated for each animal and statistically significant differences were compared between control and *Cdx2*$^{E\text{-}Del}$ mice using a *t*-test and Graphpad Prism software.

For frozen sections, intestines were fixed as above, incubated in 30% sucrose overnight and embedded in Optimal Cutting Temperature Compound (Tissue-Tek). Cryosections were rinsed in PBS and mounted with DAPI (Vectashield, Vector Laboratories) for imaging of native YFP signal.

For whole mount visualization of YFP-tracing signals, ~1cm segments of small bowel were dissected and fixed as above. Following washes in PBS, tissues were placed on slides either mucosa-up or serosa-up, mounted with coverslips, and native YFP was imaged.

Electron microscopy. Mouse ilea were flushed with PBS, fixed at least overnight in EM fixative (2% formaldehyde, 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer, pH 7.4) and embedded in Taab812 Resin (Marivac Ltd., Nova Scotia, Canada). 80 nm sections were stained with 0.2% Lead Citrate, viewed, and imaged with a Philips Technai BioTwin Spirit electron microscope at an accelerating voltage of 80 kV.

Intestinal organoid culture. Crypts isolated from the proximal small bowel of *Cdx2*$^{Fl/Fl}$*;Villin-Cre*^{ERT2} and control (*Cdx2*$^{Fl/Fl}$;Cre^{-}) mice were cultured in matrigel and organoid growth factor medium lacking Wnt, as previously described (Sato et al., 2009). 4-OH-Tamoxifen (4-OHT; Sigma; 1 μm) was added for the first two days to activate CRE^{ERT2}. Thereafter, media without 4-OHT was replaced every other day. Organoids were visualized by dissecting microscope daily to monitor differences. On day 8 organoids were fixed for 40 min in 4% PFA, rinsed in PBS and 70% ethanol, and resuspended in HistoGel (Thermo Scientific). Gel-embedded organoids were processed as described above for fresh tissues, embedded in paraffin, sectioned and stained with Ab as described above. Organoid area and outgrowths were assessed in at least 30 organoids from one mouse per condition.

ChIP-seq and analysis. Isolated LGR5-GFP^{HI} cells were collected and fixed in 1% formaldehyde for 25 minutes at room temperature, washed in PBS and snap-frozen in liquid nitrogen for pooling. 5 x 10^6 ISCs were pooled for immunoprecipitation, using previously published methods (Verzi et al., 2013) modified for ½ reactions to adjust for low chromatin input. 3 μg of CDX2 antibody (Bethyl, rabbit) was used. CDX2 ChIP-seq and input (sonicated ISC DNA not immunoprecipitated) libraries were prepared using the ThruPLEX-FD Prep Kit (Rubicon Genomics) and sequenced (50 bp, single-end reads) on an Illumina Hi-Seq 2000 instrument. 47,263,894 unique CDX2 sequence tags were aligned to the *Mus musculus* reference genome build 9 (mm9) and significant CDX2 binding sites ($P<0.0001$) were identified using MACS 1.4.2 software (Zhang et al., 2008), considering the input DNA and local background. Sequence motifs were identified using SeqPos; and phylogenetic conservation using Conservation Plot, all in the Cistrome project (Liu et al., 2011). Previously published ChIP-Seq datasets (CDX2 ChIP-Seq in villus (Verzi et al., 2013) GEO accession #: GSM851117, H3K4me2 ChIP-Seq in ISCs (Kim et al., 2014); GEO accession #: GSM1256037; and H3K27ac ChIP-Seq in ISCs (Sheaffer et al., 2014); ArrayExpress accession #: E-MTAB-2350) were re-analyzed in parallel. To visualize ChIP-seq signals, genome-wide signal tracks were normalized for sequencing depth using Wiggler and these were visualized in the Integrated Genome Viewer (Broad Institute). Heatmaps of ChIP-Seq signals surrounding CDX2 peak summits in ISCs were generated in Deeptools (Ramirez et al., 2014).

Receptor tyrosine kinase (RTK) arrays. LGR5-GFP^{HI} ISCs were isolated and pooled from control and *Cdx2*$^{ISC\text{-}Del}$ mice as described above. The Proteome Profiler Mouse Phospho-RTK

Array Kit was used for protein isolation and array protocols according to the manufacturers instructions. Total protein isolated from ISCs was quantified with the Pierce BCA Protein Assay Kit and an equal amount of protein (at least 100 μg) was applied to the array membranes. This experiment was performed on two biological replicates for each genotype. A section of the array from one representative pair is shown in Figure 3D.

Integrated analysis of ChIP-seq and RNA-seq data. CDX2 peaks were first classified into ISC only, villus only or overlapping (if the peaks intersect by at least 1 bp). To compare binding and RNA-seq data, CDX2 peaks were associated with the single nearest gene within 30 kb using GREAT (McLean et al., 2010). Genes considered bound in both populations either had a shared peak or both ISC only and villus only peaks within 30 kb. To determine whether bound genes are associated with cell-type specific gene expression, the ISC-only, villus-only or co-bound genes were associated with genes differentially expressed in ISCs or villus cells, and a Chi Square test was used to determine significant differences in proportions. To visualize enrichment of CDX2 binding near genes that change in expression during cell differentiation, all differentially expressed genes (fold change>2) were binned into groups of 100 and the average number CDX2 binding sites was calculated for these genes. Binding enrichment was calculated by dividing the average number of peaks in each bin by the average number of binding sites near all expressed genes. To investigate the association of CDX2 binding sites in ISCs with genes perturbed upon its loss, differentially expressed genes were binned into groups of 25 and the CDX2 binding enrichment over background was calculated. A Fisher's exact test was used to assess for differences in the proportion of bound genes that increase or decrease in expression.

SUPPLEMENTAL REFERENCES

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Beghtel, H., Peters, P.J.*, et al.* (2007). Identification of stem cells in small intestine and colon by marker gene *Lgr5*. Nature *449*, 1003-1007.

el Marjou, F., Janssen, K.P., Chang, B.H., Li, M., Hindie, V., Chan, L., Louvard, D., Chambon, P., Metzger, D., and Robine, S. (2004). Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. Genesis *39*, 186-193.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Kim, T.H., Li, F., Ferreiro-Neira, I., Ho, L.L., Luyten, A., Nalapareddy, K., Long, H., Verzi, M., and Shivdasani, R.A. (2014). Broadly permissive intestinal chromatin underlies lateral inhibition and cell plasticity. Nature *506*, 511-515.

Liu, T., Ortiz, J.A., Taing, L., Meyer, C.A., Lee, B., Zhang, Y., Shin, H., Wong, S.S., Ma, J., Lei, Y.*, et al.* (2011). Cistrome: an integrative platform for transcriptional regulation studies. Genome Biol *12*, R83.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol *28*, 495-501.

Munoz, J., Stange, D.E., Schepers, A.G., van de Wetering, M., Koo, B.K., Itzkovitz, S., Volckmann, R., Kung, K.S., Koster, J., Radulescu, S.*, et al.* (2012). The *Lgr5* intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers. The EMBO journal *31*, 3079-3091.

Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res *42*, W187-191.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.

Sato, T., Vries, R.G., Snippert, H.J., van de Wetering, M., Barker, N., Stange, D.E., van Es, J.H., Abo, A., Kujala, P., Peters, P.J.*, et al.* (2009). Single *Lgr5* stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature *459*, 262-265.

Sheaffer, K.L., Kim, R., Aoki, R., Elliott, E.N., Schug, J., Burger, L., Schubeler, D., and Kaestner, K.H. (2014). DNA methylation is required for the control of stem cell differentiation in the small intestine. Genes Dev *28*, 652-664.

Sigurgeirsson, B., Emanuelsson, O., and Lundeberg, J. (2014). Sequencing degraded RNA addressed by 3' tag counting. PLoS One *9*, e91851.

Srinivas, S., Watanabe, T., Lin, C.S., William, C.M., Tanabe, Y., Jessell, T.M., and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. BMC Dev Biol *1*, 4.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols *7*, 562-578.

Verzi, M.P., Shin, H., He, H.H., Sulahian, R., Meyer, C.A., Montgomery, R.K., Fleet, J.C., Brown, M., Liu, X.S., and Shivdasani, R.A. (2010). Differentiation-specific histone modifications reveal dynamic chromatin interactions and partners for the intestinal transcription factor CDX2. Dev Cell *19*, 713-726.

Verzi, M.P., Shin, H., San Roman, A.K., Liu, X.S., and Shivdasani, R.A. (2013). Intestinal master transcription factor CDX2 controls chromatin access for partner transcription factor binding. Mol Cell Biol *33*, 281-292.

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