Stem Cell Reports
Report
OPEN ACCESS

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

Adrianna K. San Roman,[1,2] Alessio Tovaglieri,[3] David T. Breault,[3,4] and Ramesh A. Shivdasani[1,5,*]
[1]Department of Medical Oncology and Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[2]Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, MA 02215, USA
[3]Division of Endocrinology, Boston Children's Hospital, Boston, MA 02115, USA
[4]Department of Pediatrics, Harvard Medical School, Boston, MA 02215, USA
[5]Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02215, USA
*Correspondence: ramesh_shivdasani@dfci.harvard.edu
http://dx.doi.org/10.1016/j.stemcr.2015.09.006

SUMMARY

Lgr5-expressing intestinal stem cells (ISCs) renew the adult gut epithelium by producing mature villus cells (VCs); the transcriptional basis for ISC functions remains unclear. RNA sequencing analysis identified transcripts modulated during differentiation of $Lgr5^+$ ISCs into VCs, with high expression of the intestine-restricted transcription factor (TF) gene *Cdx2* in both populations. *Cdx2*-deleted mouse ISCs showed impaired proliferation and long-term inability to produce mature lineages, revealing essential ISC functions. Chromatin immunoprecipitation sequencing analysis of CDX2 in $Lgr5^+$ ISCs, coupled with mRNA profiling of control and $Cdx2^{-/-}$ ISCs, identified features of CDX2 regulation distinct from VCs. Most CDX2 binding in ISCs occurs in anticipation of future gene expression, but whereas CDX2 primarily activates VC genes, direct ISC targets are activated and repressed. Diverse CDX2 requirements in stem and differentiated cells may reflect the versatility of TFs that specify a tissue in development and control the same tissue in adults.

INTRODUCTION

Intestinal stem cells (ISCs) produce diverse mature cell types. The best characterized and most abundant ISC pool consists of $Lgr5^+$ cells at the crypt base (Barker et al., 2007). These cells replicate daily, giving rise to highly proliferative transit-amplifying (TA) progenitors, which occupy higher crypt tiers and eventually produce terminally differentiated villus cells (VCs), including absorptive enterocytes and secretory (goblet and enteroendocrine) cells. $Lgr5^+$ ISCs contribute to all mature VC lineages for months (Barker et al., 2007) and generate "organoids" that self-renew in culture (Sato et al., 2009). Other ISC populations at different crypt positions replicate less frequently, respond to tissue injury, and express markers such as *Bmi1*, *Hopx*, *Lrig1*, and *Tert* (Clevers, 2013); $Lgr5^+$ ISCs may co-express the same marker genes (Itzkovitz et al., 2012; Muñoz et al., 2012). Indefinite self-renewal and the potential to yield distinct daughter lineages require particular transcriptional programs, but the transcription factors (TFs) that drive them and the associated chromatin states are insufficiently understood. Although knockout mice have implicated several TFs, including ASCL2, KLF5, and YY1, in control of various ISC functions (Nandan et al., 2015; Perekatt et al., 2014; van der Flier et al., 2009), their target genes and mechanisms remain unclear, in part because it is difficult to isolate enough ISCs for biochemical studies.

Using RNA sequencing (RNA-seq) to profile transcripts in intestinal VCs and $Lgr5^+$ ISCs, we observed that *Cdx2*, a homeobox TF gene implicated in gut epithelial development (Gao et al., 2009; Grainger et al., 2010) and adult enterocyte maturation (Verzi et al., 2010b), was highly and almost equally expressed in both cell populations. Conditional null mice revealed an autonomous proliferation deficit in $Cdx2^{-/-}$ ISCs, and instead of producing mature intestinal cell types, these cells formed undifferentiated cystic structures in vivo and failed to generate proper organoids in culture. Analysis of CDX2 occupancy in wild-type ISCs, integrated with transcriptional changes in purified $Cdx2^{-/-}$ ISCs, identified candidate direct target genes and revealed a likely dual role in gene activation and repression. Most CDX2 binding in ISCs, however, occurs in anticipation of gene activity in future cell generations.

RESULTS

Differential Gene Expression in ISCs and VCs

To identify gene expression changes during intestinal epithelial differentiation, we isolated $Lgr5^+$ ISCs from $Lgr5^{EGFP\text{-}CreERT2}$ knockin mice ($Lgr5^{KI}$) (Barker et al., 2007) and mature VCs, comprised mainly of enterocytes. As RNA-seq revealed modest 3′ bias in ISC samples, we applied a 3′ tag counting method to restrict FPKM (fragments per kilobase of exon per million fragments mapped) calculations to reads mapping to the most 3′ 1,000 bp of each gene (Sigurgeirsson et al., 2014). We observed significant differences in 3,317 VC-enriched and 2,889 ISC-enriched transcripts ($Q < 0.05$, fold change > 2; Figure 1A).

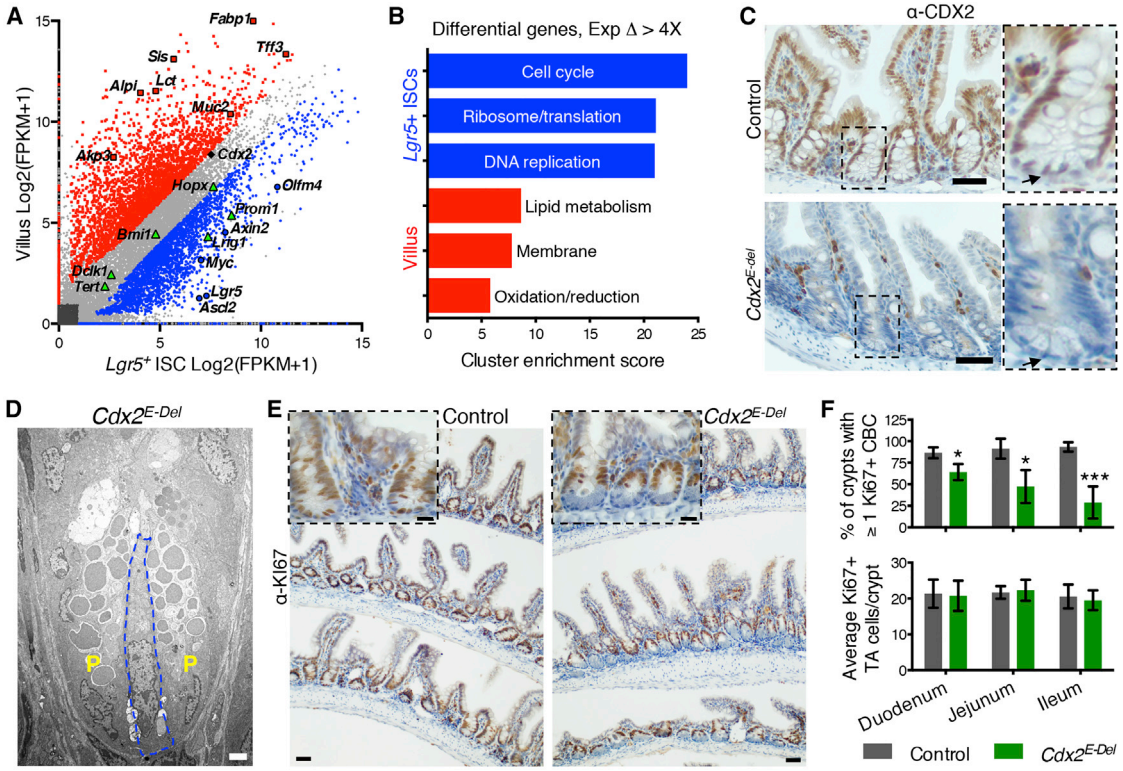


Figure 1. Loss of *Cdx2* mRNA Selectively Impairs ISC Proliferation

(A) RNA-seq data from three biological replicates of *Lgr5*$^+$ ISCs and VCs identify modulated genes. Grey, 6,752 genes with comparable expression in both cell types; red, 3,317 VC-enriched genes; blue, 2,889 ISC-enriched genes (>2-fold; $Q < 0.05$). Markers of enterocytes or goblet cells (red), *Lgr5*$^+$ ISCs (blue), and other ISC populations (green) are highlighted. *Cdx2* is highly expressed in both cell types.

(B) Gene Ontology clusters of genes with >4-fold change in VCs and ISCs.

(C) CDX2 immunohistochemistry (IHC) in control and *Cdx2*$^{E\text{-}Del}$ mouse ileum 3 days after final TAM treatment, confirming complete epithelial loss, including ISCs (magnified to right; arrows).

(D) Electron microscopy of *Cdx2*$^{E\text{-}Del}$ ileum confirms the presence of ISCs (blue) among Paneth (P) cells.

(E) KI67 IHC reveals loss of proliferation in ISCs (magnified in Inset), but not in TA cells.

(F) Fraction of crypts with at least one KI67$^+$ ISC in different regions of control and *Cdx2*$^{E\text{-}Del}$ intestines (*$p < 0.01$; ***$p < 0.0001$; t test). The average number of KI67$^+$ TA cells is not significantly altered. At least 30–50 crypts were counted in each of at least four biological replicates.

Scale bars represent 50 μm (C and E), 20 μm (inset in E), and 2 μm (D).

VC-enriched genes included enterocyte (*Sis*, *Alpi/Akp3*, *Lct*, *Fabp1*) and goblet-cell (*Muc2*, *Tff3*) markers, while genes such as *Lgr5*, *Ascl2*, *Axin2*, *Myc*, and *Olfm4* were enriched in ISCs (Figure 1A). Genes enriched >4-fold in ISCs belonged in Gene Ontology terms related to cell-cycle control, protein translation, and DNA replication, while VC-enriched terms included lipid metabolism, membrane transport, and oxidation-reduction (Figure 1B), indicating that the data are reliable. Genes previously ascribed to alternative ISC populations—*Bmi1*, *Hopx*, *Lrig1*, and *Prom1*—were highly expressed in both *Lgr5*$^+$ ISCs and VCs (Figure 1A), supporting recent reports that these markers are not exclusive (Itzkovitz et al., 2012; Muñoz et al., 2012). Levels of *Cdx2*, an intestine-restricted TF gene necessary for enterocyte maturation (Verzi et al., 2010b), were almost as high in purified ISCs as in VCs (Figure 1A); immunostaining confirmed CDX2 expression in ISCs (Figure 1C).

Loss of CDX2 Markedly and Selectively Impairs ISC Proliferation and Differentiation

To investigate a possibly separate CDX2 role in ISCs, we examined *Cdx2*$^{Fl/Fl}$;*Villin-Cre*^{ERT2} mice, where tamoxifen (TAM) deletes *Cdx2* in all small intestine epithelial cells (*Cdx2*$^{E\text{-}Del}$ for epithelium deleted; Figure 1C). Although ISCs with characteristic ultrastructural features were present at the crypt base (Figure 1D), the number of crypts with at least one KI67$^+$ ISC was reduced throughout the *Cdx2*$^{E\text{-}Del}$ small intestine, most significantly in the ileum

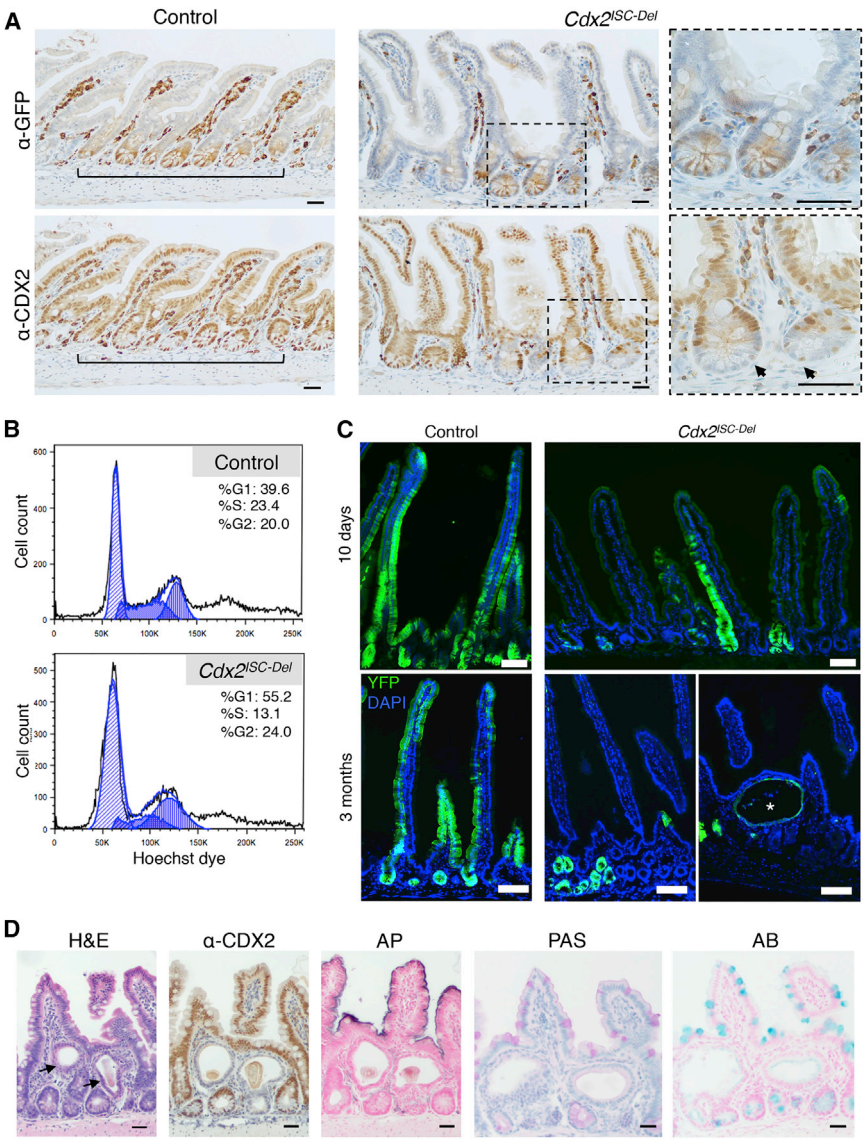


Figure 2. CDX2 Is Required for *Lgr5*$^{+}$ ISC Proliferation and Differentiation

(A) Serial tissue sections from control and *Cdx2*$^{\text{ISC-Del}}$ intestines 4 days after TAM, showing CDX2 loss in patches of LGR5-GFP-Cre^{+} crypts (dashed areas magnified to right). CDX2 is retained in some Paneth cells in these crypts (arrows).

(B) Representative cell-cycle analysis of *Lgr5*$^{+}$ ISCs from control and *Cdx2*$^{\text{ISC-Del}}$ proximal small bowel. Flow cytometry of *Cdx2*$^{\text{Del}}$ ISCs stained with Hoechst dye reveals accumulation in G1, with a reduced S-phase fraction.

(C) *Cdx2*$^{\text{Fl/Fl}}$;*Lgr5*$^{\text{KI}}$;*R26R*$^{\text{YFP}}$ and control mice were injected with TAM for 5 days and examined at the indicated times. YFP fluorescence at 10 days indicates that traced *Cdx2*$^{\text{Del}}$ crypts retain some ability to contribute to villi. At 90 days, tracing in *Cdx2*$^{\text{ISC-Del}}$ intestines is confined to cysts (asterisk) and unproductive crypts.

(D) H&E staining of the cysts shows luminal debris. IHC confirms that traced cysts lack CDX2 and are surrounded by CDX2^{+} crypts that did not express the *Lgr5*$^{\text{KI}}$ allele. Alkaline phosphatase (AP), periodic acid-Schiff (PAS), and Alcian blue (AB) staining show absence of mature enterocytes and goblet cells in the cystic structures.

Scale bars represent 30 μm (A and D) and 50 μm (C). See also Figure S1.

(Figures 1E and 1F), where CDX2 levels are normally highest (James and Kazenwadel, 1991). This defect was confined to crypt base ISCs; TA cells in higher tiers showed a normal mitotic fraction (Figures 1E and 1F).

Cdx2$^{\text{Fl/Fl}}$;*Lgr5*$^{\text{KI}}$ mice (*Cdx2*$^{\text{ISC-Del}}$ for ISC deleted) showed CDX2 loss from all GFP^{+} crypts after TAM treatment (Figure 2A). As *Lgr5*$^{\text{KI}}$ mice are mosaic (Barker et al., 2007), many crypts retained CDX2 in all cells, and because Paneth cells do not express *Lgr5* (Barker et al., 2007) and turn over slowly (Ireland et al., 2005), they retained CDX2 (Figure 2A). In the proximal small intestine, where mosaic *Lgr5*$^{\text{KI}}$ expression is highest (Barker et al., 2007), we observed fewer LGR5-GFP$^{\text{HI}}$ cells in S-phase and more cells in G1 (Figure 2B). Thus, independent measures reveal a significant defect in replication of *Cdx2* null ISCs. Because CDX2 is expressed only in epithelial cells and not in the lamina propria, this defect must be cell autonomous.

Intact TA cell replication suggested that early events following ISC differentiation may not require CDX2. To test this possibility, we crossed a *Rosa26*$^{\text{lox-STOP-lox-YFP}}$ allele (*R26R*$^{\text{YFP}}$ [Srinivas et al., 2001]) onto the *Cdx2*$^{\text{Fl/Fl}}$;*Lgr5*$^{\text{KI}}$ background to enable lineage tracing. After treating mice with TAM, we followed YFP expression serially (Figure 2C), noting again that mosaic *Lgr5*$^{\text{KI}}$ expression would delete *Cdx2* and activate *Yfp* in a minority of crypts. As TAM induction of the *Lgr5-Cre*$^{\text{ERT2}}$ irreversibly marks cells with YFP, continuous *Lgr5* expression is not required to monitor the lineage tracing. Ten days later, YFP was present in crypts and villi in control (*Cdx2*$^{+/+}$;*Lgr5*$^{\text{KI}}$;*R26R*$^{\text{YFP}}$) and, at slightly lower frequency, in *Cdx2*$^{\text{ISC-Del}}$ mice (Figure 2C). However,

starting at 1 month and continuing for at least 6 months after Cre activation, YFP expression in *Cdx2*$^{ISC\text{-}Del}$ intestines was confined to crypts and cystic inclusions, and never detected in villi (Figures 2C, S1A, and S1B). These cysts, similar to those previously noted in *Cdx2*$^{-/-}$ crypts (Stringer et al., 2012), showed total absence of CDX2 and of mature VC features such as alkaline phosphatase, which marks enterocytes, or avidity for periodic acid-Schiff or Alcian blue, which marks goblet cells (Figure 2D). Cells lining the cysts had a cuboidal morphology, with round nuclei (Figure S1C); EZR and NHERF localization indicated intact apico-basal polarity (Figure S1D). LGR5-GFP expression was typically confined to one side of the cysts, near proliferating TA-like cells, which is reminiscent of normal crypt organization, and non-replicating cells commonly were sloughed from the upper cyst wall, collecting as luminal debris (Figure S1E). Thus, *Cdx2*^{Del} ISCs proliferate poorly, and although they spawn replication-competent TA cells, maturation of their progeny is severely impaired. Of note, intestinal crypt cells also express CDX1 (James and Kazenwadel, 1991), a homolog with redundant activity in TA cells (Verzi et al., 2011). The specific defects in *Cdx2*$^{-/-}$ crypts reveal a crucial, non-redundant requirement in the stem cell compartment.

Lgr5^{+} ISCs generate organoid bodies when cultured in defined media (Sato et al., 2009). To test whether *Cdx2* loss affects this property, we treated *Cdx2*$^{Fl/Fl}$*;Villin-Cre*^{ERT2} intestinal crypts with TAM in culture. Eight days later, organoids from control *Cre*$^{-}$ crypts were large and showed crypt-like evaginations, whereas *Cdx2*-deficient crypts formed few organoids; when present, these were small and spherical, with few protrusions, and totally lacked CDX2, but showed signs of multilineage differentiation in the form of mature enterocyte and goblet cell markers (Figures S1F–S1H). Thus, significant deficits in cell replication, lineage tracing, organoid formation, and generation of mature intestinal cells reveal a CDX2 requirement in ISC functions, ostensibly distinct from CDX2's activity in mature enterocytes or its redundant functions with CDX1 in TA cells.

CDX2 Loss in ISCs Affects Genes that Regulate Cell Proliferation

To identify genes dysregulated in CDX2 null ISCs, we isolated GFP^{+} cells from *Cdx2*$^{ISC\text{-}Del}$ and control (*Cdx2*$^{+/+}$*; Lgr5*^{KI}) intestinal crypts after 4 days of TAM exposure. RNA-seq confirmed loss of most transcripts containing *loxP*-flanked exon 2 in *Cdx2*^{Del} ISCs (Figure S2A). In addition, 412 transcripts were decreased and 639 were increased at least 1.5-fold ($Q < 0.05$; Figure 3A). Genes with reduced expression were enriched for terms related to mitosis, consistent with reduced ISC replication, whereas genes that increased in expression were enriched for terms related to drug and lipid metabolism (Figure 3B). Because CDX2 may mediate intestine-specific Wnt signaling (Verzi et al., 2010a), we assessed the consequences of CDX2 loss on a curated list of 203 intestinal Wnt target genes (van de Wetering et al., 2002). These transcripts showed higher average expression in control ISCs than in VCs, but average expression was unperturbed in *Cdx2*^{Del} ISCs (Figure 3C); only 24 individual Wnt target genes were significantly affected, with no obvious pattern (Figure S2B). Although 217 CDX2 binding sites contain a TCF/LEF motif (Figure 3C), the data suggest that CDX2 activity in ISCs is largely Wnt independent, possibly compensated by the partially redundant factor CDX1. We also probed arrays containing 39 phosphorylated receptor tyrosine kinases with protein lysates from purified control and *Cdx2*^{Del} ISCs. Strong phosphorylation signals for the epidermal growth factor receptor (EGFR) and ERBB2 were reproducibly lower in *Cdx2*^{Del} ISCs (Figure 3D). Together, these findings reveal that CDX2 controls select ISC genes, some of which mediate cell surface events in EGFR signaling. A single dominant target gene is, however, unlikely to explain CDX2 requirements in vivo. We therefore sought to determine features of CDX2 regulation by mapping its binding sites in ISCs.

Gene Targets of CDX2 Regulation in ISCs

In contrast with >12,000 sites previously detected in VCs (Verzi et al., 2013), chromatin immunoprecipitation sequencing (ChIP-seq) on purified wild-type *Lgr5*$^{+}$ ISCs identified 2,371 regions enriched for CDX2 binding (e.g., Figure 3E). The distribution of CDX2 binding sites in promoters and distant sites in *Lgr5*$^{+}$ ISCs resembled that in VCs (Figure 3F); these areas showed evolutionary conservation (Figure S2C), and the CDX2 consensus motif was the most recurring (Figure 3G). Moreover, binding sites in ISCs coincided with presence of H3K4me2 (Kim et al., 2014) and H3K27ac (Sheaffer et al., 2014) histone marks, indicating that they represent bona fide active *cis*-regulatory elements (Figures 3E and 3H).

Most CDX2 binding sites in VCs (Verzi et al., 2013) gave weak or no signals in *Lgr5*$^{+}$ ISCs (Figure 4A), reflecting either truly absent binding or the challenges inherent in detecting low-level occupancy. Conversely, at least 60% of CDX2 binding sites in purified ISCs were also evident in mature VCs (Figures 4A and 4B), likely representing occupancy at genes expressed in both cell types or ISC binding in anticipation of VC activity. The remaining ~1,000 binding sites in *Lgr5*$^{+}$ ISCs were considerably weaker or absent in VCs (Figures 4A and 4B) and may represent sites of ISC-specific activity. To distinguish among these possibilities, we first assigned CDX2 binding sites to the nearest gene within 30 kb. Of the 1,156 genes showing at least 1 binding site in ISCs, 883 genes were also occupied

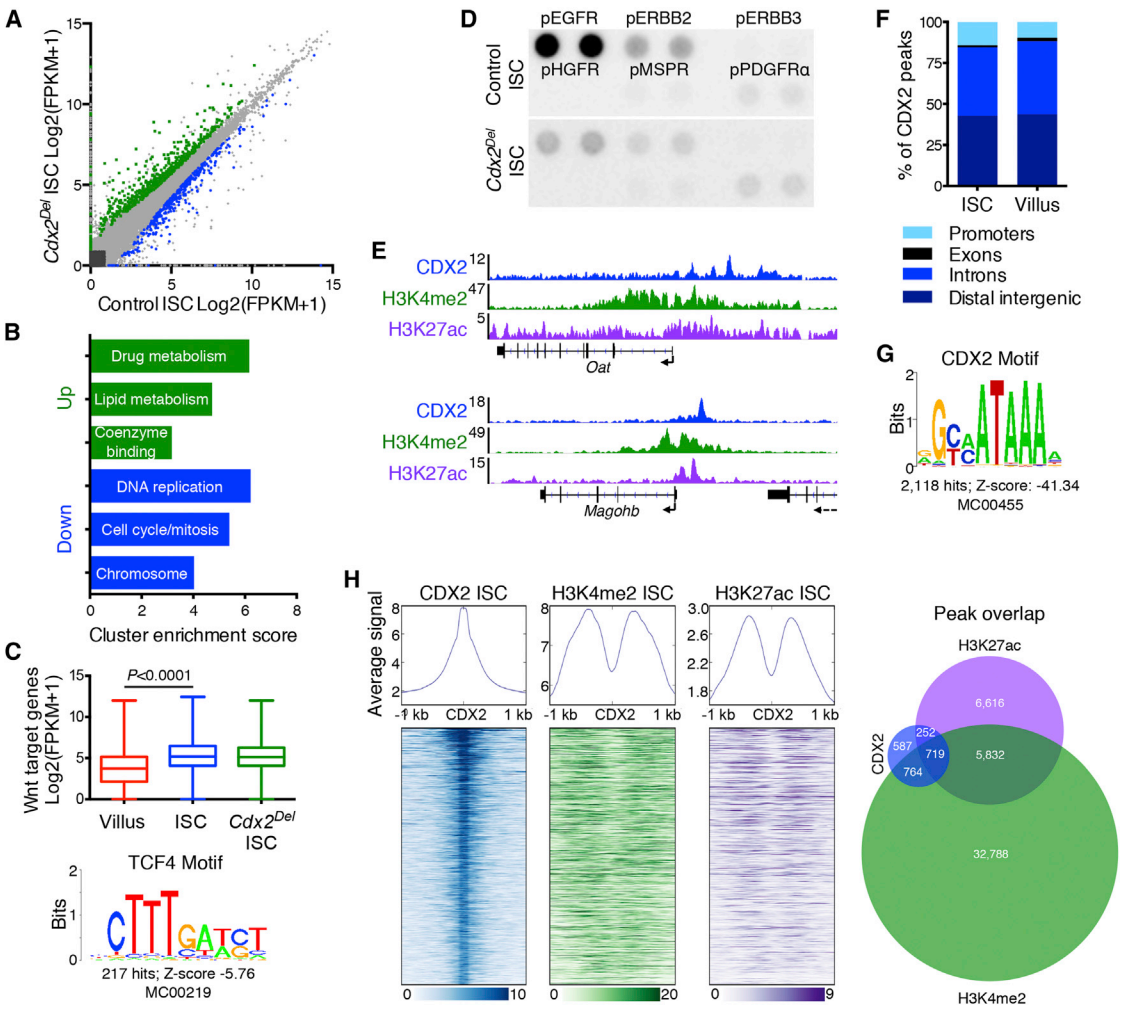


Figure 3. Genes Perturbed by CDX2 Loss and Delineation of CDX2 Binding in ISCs

(A) RNA-seq data from isolated control and *Cdx2*^{Del} ISCs, using three biological replicates. 639 transcripts are elevated in *Cdx2*$^{-/-}$ ISCs (green), and 412 transcripts are reduced (blue). Grey, genes with <1.5-fold change or $Q > 0.05$.

(B) Gene Ontology cluster enrichment among genes reduced or increased in *Cdx2*^{Del} ISCs.

(C) Average RNA-seq FPKM values of 203 intestinal Wnt target genes in control VCs and ISCs and in *Cdx2*^{Del} ISCs. Significance was assessed by a t test. Bottom: the TCF4 motif found in 217 CDX2 binding sites.

(D) Phospho-receptor tyrosine kinase arrays probed with lysates from purified control or *Cdx2*^{Del} ISCs. Signals for pEGFR and pERBB2 are reduced in *Cdx2*^{Del} ISCs.

(E) ChIP-seq signals for CDX2, H3K4me2, and H3K27ac near the *Oat* and *Magohb* genes; the y axis represents aligned ChIP-seq tag counts.

(F) Distribution of CDX2 binding sites among genomic regions is similar in ISCs and VCs.

(G) The consensus CDX2 motif is the most enriched (Z score: −41.34) in CDX2-bound sites.

(H) Heatmaps of ChIP signals for CDX2, H3K4me2, and H3K27ac around CDX2 binding sites in purified ISCs. Right: overlap of CDX2 binding sites with called H3K4me2 and H3K27ac peaks, which underestimate the overlap evident in heatmaps.

See also Figure S2 and Table S1.

in VCs (Figure 4C). We first approached these findings from the perspective of genes with nearby CDX2 binding (Figure 4D). The small proportion of genes bound only in ISCs was similar among ISC-enriched and VC-expressed genes, whereas genes bound only in VCs or in both ISCs and VCs were more highly expressed in VCs. Approaching the data next from the perspective of genes differentially expressed in ISCs and VCs (Figure 4E), we see that villus-expressed genes are enriched for CDX2 binding in both ISCs and VCs. In contrast, genes with unique CDX2 binding in ISCs are distributed across the expression spectrum (Figure 4E). Notably, genes bound in both ISCs and VCs

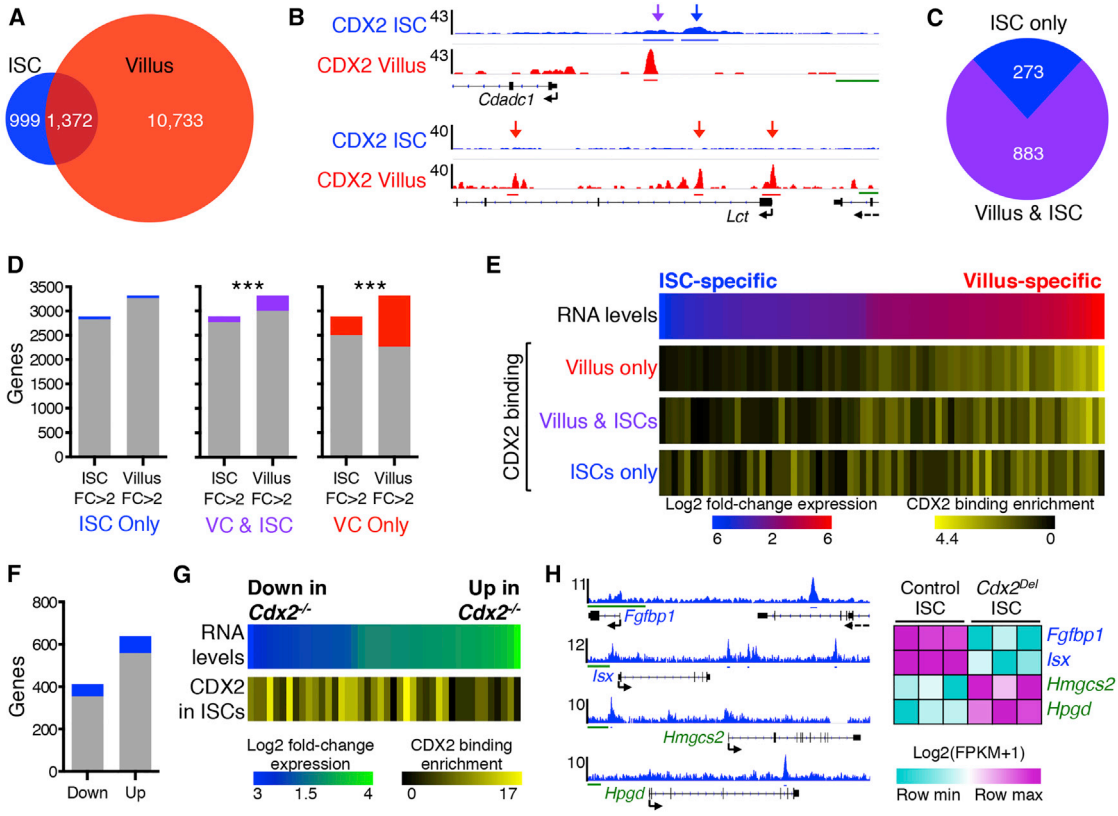


Figure 4. Contrasting Features of CDX2 Binding in ISCs and VCs

(A) Number of CDX2 binding sites identified exclusively in ISCs (blue), only in VCs (red), or in both cell types with ≥1-bp overlap (purple).

(B) Illustrative data traces for CDX2 binding sites of each class (arrows). y axis, ChIP-seq tag counts; x axis: genome coordinates.

(C) Among the 1,156 genes that bind CDX2 within 30 kb in ISC, 273 are bound only in ISCs and 883 are also bound in VCs.

(D) Genes bound only in VCs (red) or in both VCs and ISCs (purple) are mostly expressed in VCs (***$p < 0.0001$; χ^2 test); genes bound only in ISCs (blue) are not enriched for expression in either compartment.

(E) Genes ordered by the degree of differential expression in ISCs and VCs (>2-fold, $Q < 0.05$, from data in Figure 1A) were binned into groups of 100 (blue–red heatmap). Yellow heatmaps represent average number of CDX2 binding sites in VCs only, in ISCs only, or in both populations within 30 kb of genes in each bin, normalized for average CDX2 binding across all expressed genes.

(F) CDX2 binds near similar proportions of genes ($p = 0.57$, Fisher's exact test) that show reduced or increased transcript levels in *Cdx2*$^{-/-}$ ISCs.

(G) Genes ordered by the degree of differential expression in *Cdx2*^{Del} versus control ISCs (from data in Figure 3A) were binned into groups of 25 (blue–green heatmap). Below: average number of CDX2 binding sites in ISCs within 30 kb of the genes in each bin, divided by the average CDX2 binding across all expressed genes.

(H) CDX2 binding data at representative upregulated (*Hmgcs2* and *Hpgd*) and downregulated (*Fgfpb1* and *Vdr*) genes. Right: heatmap of corresponding RNA-seq data.

Scale bars represent 1 kb (green line in B) and 5 kb (H). See also Figures S2 and S3.

are highly enriched for "transcriptional activity" (Gene Ontology cluster enrichment score: 3.38) and include *Cdx1/2*, *Gata4/6*, *Hnf4a/g*, *Hnf1b*, *Klf5*, *Vdr*, and *Yy1* (Figure S2D). Thus, much of the CDX2 binding in ISCs persists in differentiated cells, near genes that are highly expressed in the latter. This binding thus precedes future VC activity rather than serving functions unique to ISCs.

Because direct transcriptional targets in ISCs should show both CDX2 binding and perturbed mRNA levels in its absence, we considered ISC binding in the light of differential gene expression in *Cdx2*$^{ISC\text{-}Del}$ ISCs. Most CDX2-bound genes were unaffected in transcript levels, implying that CDX2 is dispensable for their regulation. Among the 137 genes with CDX2 binding and altered mRNA levels in *Cdx2*$^{-/-}$ ISCs, 80 genes were increased and 57 were decreased (Figures 4F–4H and S3). In VCs, by contrast, CDX2 binds almost exclusively at loci that are downregulated upon *Cdx2* loss (Verzi et al., 2013). Thus, whereas

CDX2 predominantly activates VC genes, in ISCs it may activate some genes and repress others.

DISCUSSION

We demonstrate an early, cell-autonomous role for CDX2 in promoting ISC replication in vivo and *Cdx2*$^{-/-}$ cells' steadily diminished ability to produce differentiated VCs. Previous studies of *Cdx2*^{Del} cysts and organoid cultures have suggested that *Cdx2* loss activates stomach epithelial marker genes and hence compromises intestinal identity (Hryniuk et al., 2012; Simmini et al., 2014; Stringer et al., 2012). However, our RNA-seq analysis of *Cdx2*^{Del} ISCs revealed elevated levels of only 2 of 11 common stomach-specific transcripts (Table S1), and we did not detect CDX2 binding near these genes in ISCs. Thus, gastric transdifferentiation does not seem to be a prominent early feature of *Cdx2*$^{-/-}$ ISCs, though other target genes, including transcriptional and epigenetic regulators, may perturb stomach-specific genes over time.

Feedback mechanisms for ISC homeostasis are poorly understood. By revealing different CDX2 requirements in various crypt cell populations, our findings provide useful insights. First, combined loss of CDX1 and CDX2 markedly affects TA cell replication (Verzi et al., 2011), but these cells withstand loss of CDX2 alone. Second, the outcome when *Lgr5*$^{+}$ ISCs are incapacitated by CDX2 loss—but remain viable—differs from that when *Lgr5*$^{+}$ ISCs are killed using *Diphtheria* toxin. Whereas reserve Bmi1^{+} ISCs are rapidly recruited into an active role in the latter case (Tian et al., 2011), *Cdx2*$^{-/-}$ cysts persist for months, avoiding replacement by CDX2^{+} cells, such as the progeny of quiescent ISCs in *Cdx2*$^{ISC\text{-}Del}$ mice. One possibility is that leaky *Lgr5-Cre* expression in quiescent ISCs triggered deletion of *Cdx2*, which is also required in this ISC pool, but *Lgr5* expression is notably absent in quiescent ISCs in *Lgr5*^{KI} mice (Barker et al., 2007; Tian et al., 2011). We therefore favor the alternative explanation that persistent ISCs and TA cells in *Cdx2*$^{ISC\text{-}Del}$ intestines are sufficient to avoid the feedback signals that elicit compensatory activity from quiescent ISCs.

This study advances understanding of CDX2 mechanisms in ISC maintenance and differentiation. CDX2 occupies hundreds of genomic sites exclusively in ISCs and thousands of additional sites in VCs. Unlike in VCs, however, CDX2 binding in ISCs correlates poorly with ISC-expressed genes. Instead, much of it overlaps with binding sites in VCs, occurring at genes activated in subsequent cell generations. Moreover, while CDX2 behaves almost exclusively as a transactivator in VCs, it may directly activate or repress ISC genes. This could occur through interactions with different co-factors, alternative post-translational modifications, or differential chromatin access. Studies of other TFs and chromatin in purified ISCs are necessary to resolve these possibilities.

EXPERIMENTAL PROCEDURES

Mice

Cdx2$^{Fl/Fl}$, *Villin-Cre*^{ERT2}, *Lgr5-EGFP-IRES-Cre*^{ERT2}, and *Rosa26-lox-STOP-lox-YFP* mice were described previously (Barker et al., 2007; el Marjou et al., 2004; Srinivas et al., 2001; Verzi et al., 2010b). For recombination of conditional alleles, mice were injected intraperitoneally for 4–5 days with 1–2 mg tamoxifen (Sigma). All procedures followed protocols approved by an Institutional Animal Care and Use Committee.

Epithelial Cell Isolation

Crypts from the proximal small intestine were isolated by scraping off villi and incubating in 5 mM EDTA. Crypts were trypsinized into a single cell suspension and passed through a FACSAria III instrument (BD Biosciences) to sort single live LGR5-GFP^{HI} cells. Alternatively, intestines were incubated in EDTA and VCs were captured using 70-μm filters.

RNA-Seq

More than 10^5 control ISCs, *Cdx2*^{Del} ISCs, or VCs were collected from single mice, and isolated RNA was sequenced (see the Supplemental Experimental Procedures for RNA isolation, library preparation, sequencing, and analysis protocols). Three biological replicates for each condition were compared, and significantly dysregulated genes were analyzed for Gene Ontology term enrichment.

Histology

Intestine segments were fixed overnight in 4% paraformaldehyde at 4°C, and they were dehydrated for paraffin embedding or incubated overnight in 30% sucrose and frozen in optimum cutting temperature (OCT) compound (Tissue-Tek). 5-μm paraffin sections were stained with H&E, alkaline phosphatase, alcian blue, periodic acid-Schiff, or specific antibodies (Ab) to CDX2, KI67, GFP, EZR, or NHERF (more information in the Supplemental Experimental Procedures). 10-μm cryosections were mounted with DAPI to image native YFP.

ChIP-Seq

5×10^6 LGR5-GFP^{HI} ISCs were pooled from several mice and cross-linked in 1% formaldehyde at room temperature for 25 min, followed by lysis, sonication, and immunoprecipitation with 3 μg CDX2 Ab (Bethyl Laboratories). See the Supplemental Experimental Procedures for ChIP, sequencing, and analysis protocols. CDX2 binding sites were identified from aligned sequencing reads and assigned to the single nearest gene within 30 kb. Published ChIP-seq data were analyzed in parallel.

Additional details of protocols can be found in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for all sequencing data reported in this paper is GEO: GSE70766.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2015.09.006.

ACKNOWLEDGMENTS

This work was supported by NIH awards R01DK082889 (R.A.S.), R01DK084056 (D.T.B.), and F31CA180784 (A.K.S.R.) and a NSF graduate research fellowship (A.K.S.R.). We thank S. Robine for *Villin-Cre*^{ERT2} mice; P. Cejas and T.-H. Kim for sharing expertise and reagents; and M. Verzi and N. O'Neill for valuable discussions.

Received: July 14, 2015
Revised: September 9, 2015
Accepted: September 10, 2015
Published: October 15, 2015

REFERENCES

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

Clevers, H. (2013). The intestinal crypt, a prototype stem cell compartment. Cell *154*, 274–284.

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.

Gao, N., White, P., and Kaestner, K.H. (2009). Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2. Dev. Cell *16*, 588–599.

Grainger, S., Savory, J.G.A., and Lohnes, D. (2010). Cdx2 regulates patterning of the intestinal epithelium. Dev. Biol. *339*, 155–165.

Hryniuk, A., Grainger, S., Savory, J.G.A., and Lohnes, D. (2012). Cdx function is required for maintenance of intestinal identity in the adult. Dev. Biol. *363*, 426–437.

Ireland, H., Houghton, C., Howard, L., and Winton, D.J. (2005). Cellular inheritance of a Cre-activated reporter gene to determine Paneth cell longevity in the murine small intestine. Dev. Dyn. *233*, 1332–1336.

Itzkovitz, S., Lyubimova, A., Blat, I.C., Maynard, M., van Es, J., Lees, J., Jacks, T., Clevers, H., and van Oudenaarden, A. (2012). Single-molecule transcript counting of stem-cell markers in the mouse intestine. Nat. Cell Biol. *14*, 106–114.

James, R., and Kazenwadel, J. (1991). Homeobox gene expression in the intestinal epithelium of adult mice. J. Biol. Chem. *266*, 3246–3251.

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.

Muñoz, 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. EMBO J. *31*, 3079–3091.

Nandan, M.O., Ghaleb, A.M., Bialkowska, A.B., and Yang, V.W. (2015). Krüppel-like factor 5 is essential for proliferation and survival of mouse intestinal epithelial stem cells. Stem Cell Res. (Amst.) *14*, 10–19.

Perekatt, A.O., Valdez, M.J., Davila, M., Hoffman, A., Bonder, E.M., Gao, N., and Verzi, M.P. (2014). YY1 is indispensable for *Lgr5*+ intestinal stem cell renewal. Proc. Natl. Acad. Sci. USA *111*, 7695–7700.

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., and Clevers, H. (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., Schübeler, 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.

Simmini, S., Bialecka, M., Huch, M., Kester, L., van de Wetering, M., Sato, T., Beck, F., van Oudenaarden, A., Clevers, H., and Deschamps, J. (2014). Transformation of intestinal stem cells into gastric stem cells on loss of transcription factor Cdx2. Nat. Commun. *5*, 5728.

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.

Stringer, E.J., Duluc, I., Saandi, T., Davidson, I., Bialecka, M., Sato, T., Barker, N., Clevers, H., Pritchard, C.A., Winton, D.J., et al. (2012). Cdx2 determines the fate of postnatal intestinal endoderm. Development *139*, 465–474.

Tian, H., Biehs, B., Warming, S., Leong, K.G., Rangell, L., Klein, O.D., and de Sauvage, F.J. (2011). A reserve stem cell population in small intestine renders *Lgr5*-positive cells dispensable. Nature *478*, 255–259.

van de Wetering, M., Sancho, E., Verweij, C., de Lau, W., Oving, I., Hurlstone, A., van der Horn, K., Batlle, E., Coudreuse, D., Haramis, A.P., et al. (2002). The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell *111*, 241–250.

van der Flier, L.G., van Gijn, M.E., Hatzis, P., Kujala, P., Haegebarth, A., Stange, D.E., Begthel, H., van den Born, M., Guryev, V., Oving, I., et al. (2009). Transcription factor achaete scute-like 2 controls intestinal stem cell fate. Cell *136*, 903–912.

Verzi, M.P., Hatzis, P., Sulahian, R., Philips, J., Schuijers, J., Shin, H., Freed, E., Lynch, J.P., Dang, D.T., Brown, M., et al. (2010a). TCF4

and CDX2, major transcription factors for intestinal function, converge on the same cis-regulatory regions. Proc. Natl. Acad. Sci. USA *107*, 15157–15162.

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. (2010b). 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., Ho, L.-L., Liu, X.S., and Shivdasani, R.A. (2011). Essential and redundant functions of caudal family proteins in activating adult intestinal genes. Mol. Cell. Biol. *31*, 2026–2039.

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.

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

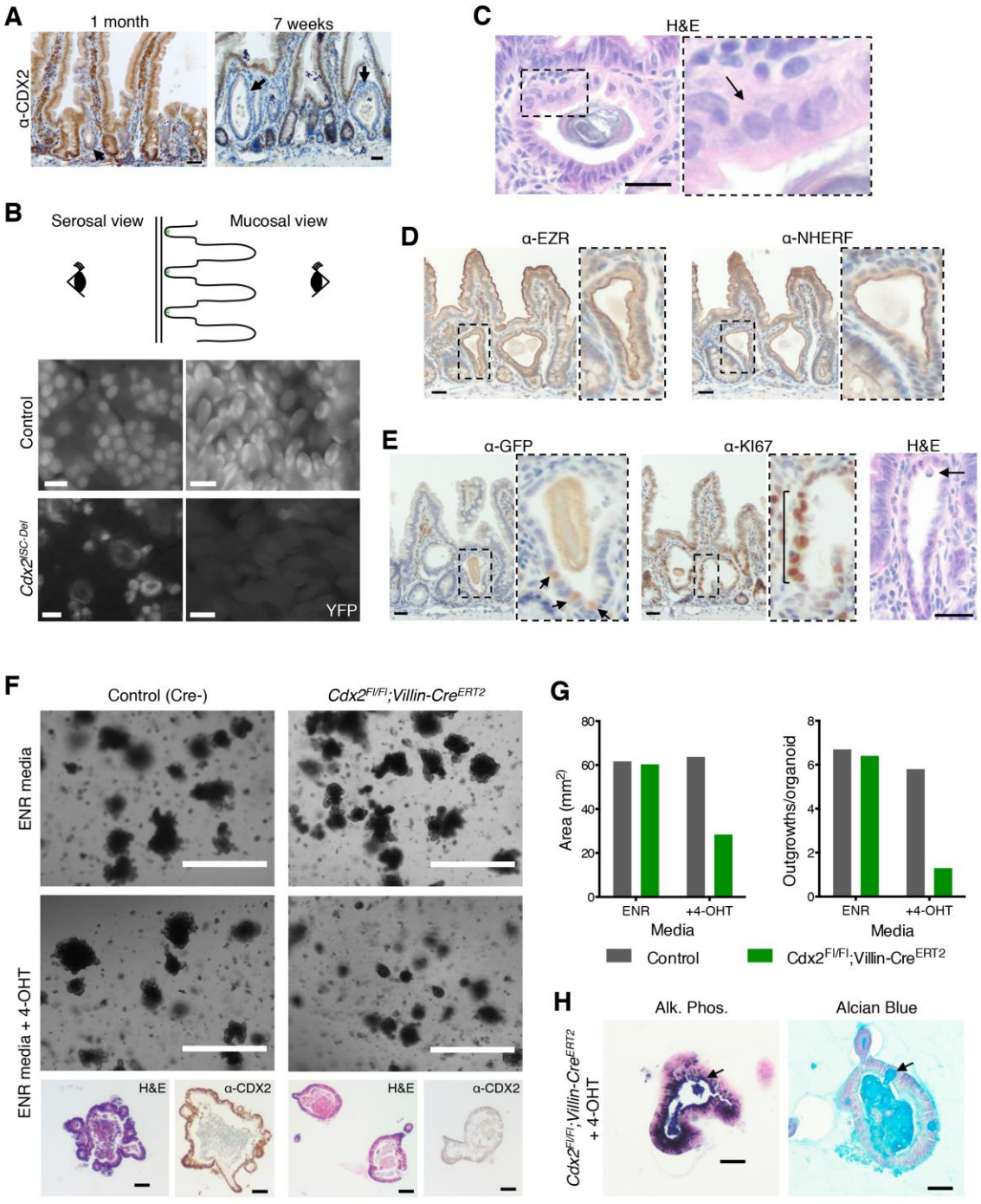
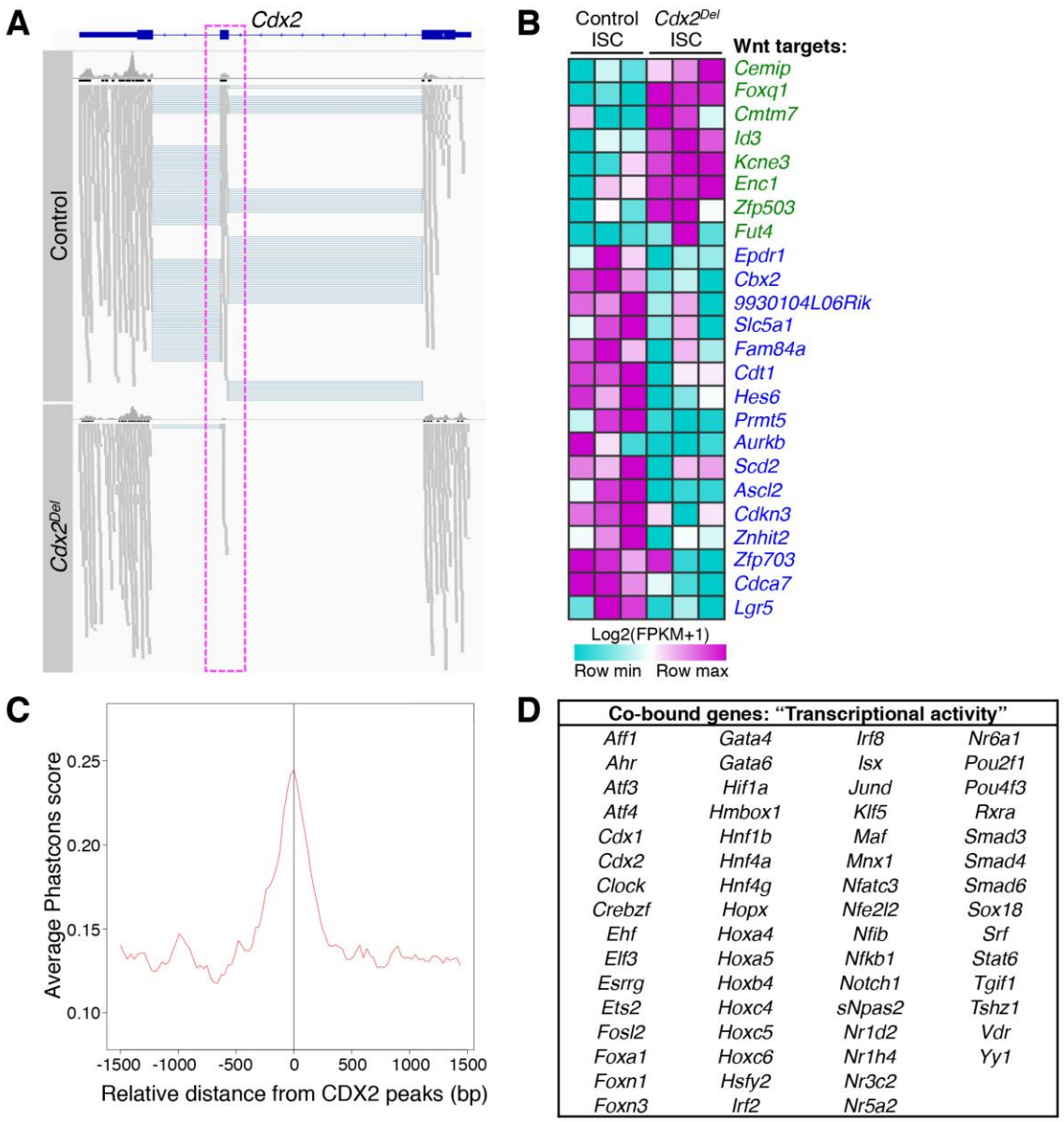


Figure S1. *Cdx2*^{Del} crypts form epithelial cysts in vivo and fail to generate mini-gut organoids in culture, related to Figure 2. (A) CDX2 immunostaining of *Cdx2*$^{ISC-Del}$ intestines 1 month and 7 weeks after *Cdx2* deletion demonstrates earliest cyst formations (arrows), which become more abundant 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.

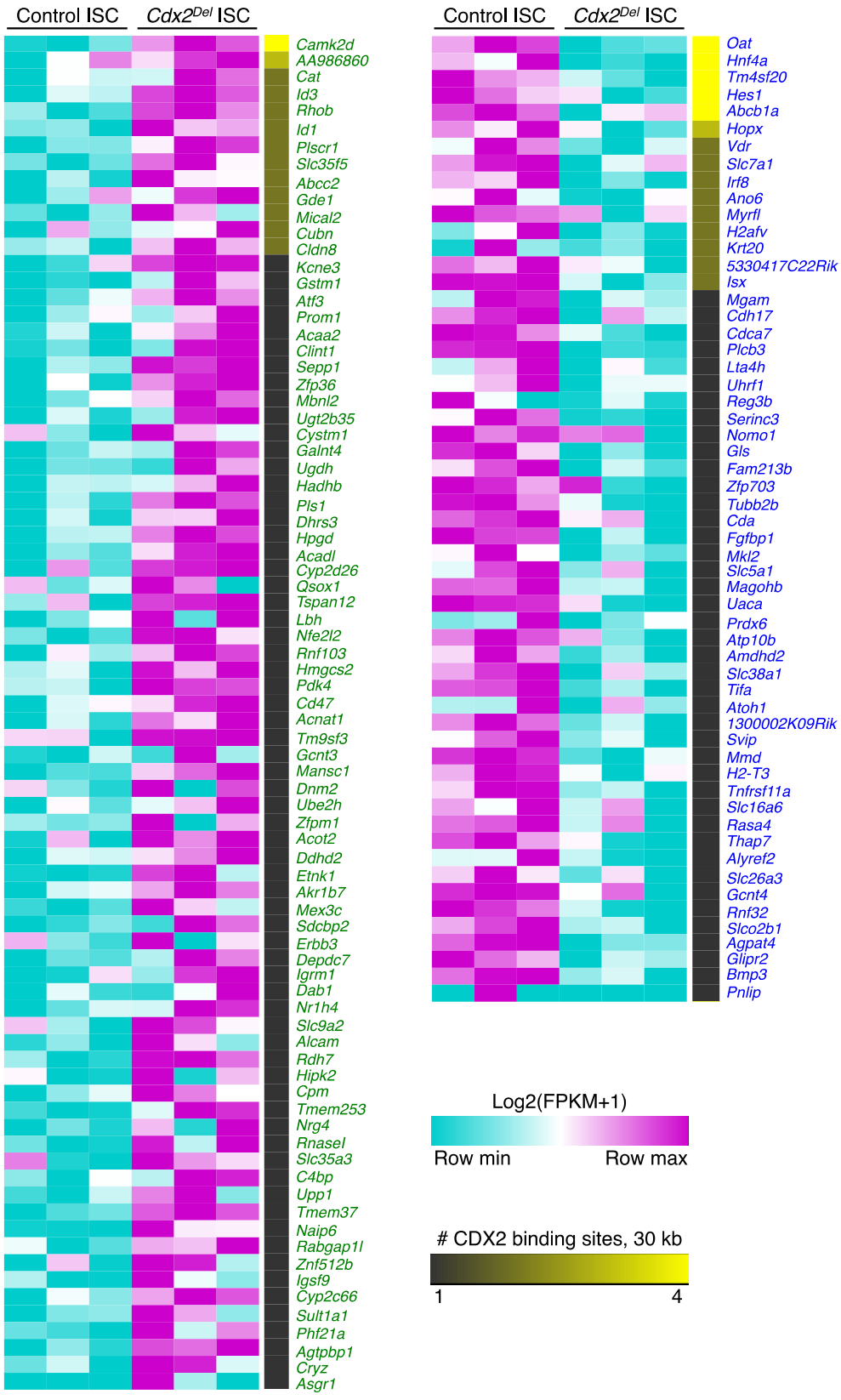

Control ISC
Cdx2Del ISC
Camk2d
AA986860
Cat
Id3
Rhob
Id1
Plscr1
Slc35f5
Abcc2
Gde1
Mical2
Cubn
Cldn8
Kcne3
Gstm1
Atf3
Prom1
Acaa2
Clint1
Sepp1
Zfp36
Mbnl2
Ugt2b35
Cystm1
Galnt4
Ugdh
Hadhb
Pls1
Dhrs3
Hpgd
Acadl
Cyp2d26
Qsox1
Tspan12
Lbh
Nfe2l2
Rnf103
Hmgcs2
Pdk4
Cd47
Acnat1
Tm9sf3
Gcnt3
Mansc1
Dnm2
Ube2h
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.