

## Essential and Redundant Functions of Caudal Family Proteins in Activating Adult Intestinal Genes<sup>∇</sup>

Michael P. Verzi,<sup>1,2,†</sup> Hyunjin Shin,<sup>3,4,†</sup> Li-Lun Ho,<sup>1,2</sup> X. Shirley Liu,<sup>3,4</sup>  
and Ramesh A. Shivdasani<sup>1,2,\*</sup>

*Department of Medical Oncology<sup>1</sup> and Department of Biostatistics and Computational Biology,<sup>3</sup> Dana-Farber Cancer Institute, Boston, Massachusetts, and Harvard School of Public Health<sup>4</sup> and Departments of Medicine, Brigham and Women's Hospital and Harvard Medical School,<sup>2</sup> Boston, Massachusetts*

Received 28 October 2010/Returned for modification 15 December 2010/Accepted 6 March 2011

**Transcription factors that potently induce cell fate often remain expressed in the induced organ throughout life, but their requirements in adults are uncertain and varied. Mechanistically, it is unclear if they activate only tissue-specific genes or also directly repress heterologous genes. We conditionally inactivated mouse *Cdx2*, a dominant regulator of intestinal development, and mapped its genome occupancy in adult intestinal villi. Although homeotic transformation, observed in *Cdx2*-null embryos, was absent in mutant adults, gene expression and cell morphology were vitally compromised. Lethality was significantly accelerated in mice lacking both *Cdx2* and its homolog *Cdx1*, with particular exaggeration of defects in villus enterocyte differentiation. Importantly, *Cdx2* occupancy correlated with hundreds of transcripts that fell but not with equal numbers that rose with *Cdx* loss, indicating a predominantly activating role at intestinal *cis*-regulatory regions. Integrated consideration of a transcription factor's mutant phenotype and cistrome hence reveals the continued and distinct requirement in adults of a critical developmental regulator that activates tissue-specific genes.**

Some transcription factors (TFs) are recognized by tissue-restricted expression and a potent ability to induce lineage-specific conversion of heterologous cells, developmental properties that reflect coordinate regulation of innumerable tissue-specific genes. Molecular mechanisms that direct cell specification are incompletely understood, as are the ongoing requirements for key TFs in mature, established adult tissues. Contributing to this limited understanding, the transcriptional activities and direct targets of most key TFs are unknown.

Two homologous homeodomain proteins of the caudal family, *Cdx1* and *Cdx2*, are restricted to the intestinal epithelium in adult animals, and each can induce intestinal differentiation in transgenic mouse stomachs or human esophageal cells (20, 22, 23, 33). Consistent with this inductive property, mouse embryos lacking *Cdx2* in the primitive gut endoderm develop a foregut type of mucosa in place of the distal intestinal epithelium (11, 12), a homeotic shift that underscores its powerful developmental functions. Despite much effort toward understanding *Cdx* functions, two important questions remain: (i) What is the *Cdx* requirement in adults, long after the intestinal epithelium is specified? (ii) What portion of the intestinal gene expression program do *Cdx* proteins regulate by direct *cis*-element binding? We combined genetic and DNA-binding approaches to answer these questions.

One factor that confounds genetic analysis of the caudal family is the potential redundancy among homologues. In the four distinct spatiotemporal domains that require *Cdx* factors, the trophoderm, blood formation, and embryonic axial skel-

eton and intestine (4, 5, 24, 27, 42), one *Cdx* protein often assumes the function of another (8, 27, 37, 38, 40). In particular, the presence of a normal intestine in *Cdx1* null mice (3) is attributed to compensation by *Cdx2*, which is expressed in an overlapping distribution (6, 18, 34). Although conditional *Cdx2* depletion in embryonic mouse endoderm materially disrupted intestinal development (11), *Cdx1* was undetectable in these animals, leaving uncertain whether the defects resulted from isolated *Cdx2* loss or the fortuitous absence of both *Cdx1* and *Cdx2*. In the adult intestinal epithelium, replicating progenitor cells reside in the crypts of Lieberkühn, and mature, differentiated cells lie along the villus projections. Whereas *Cdx2* is expressed throughout the crypt-villus unit, *Cdx1* is reported to be more prominent in crypts (3, 6, 34); and while studying differential chromatin modifications in gut epithelium, we recently uncovered redundant requirements for *Cdx1* and *Cdx2* in adult crypt cell replication (39). Because *Cdx1* expression is less prominent than that of *Cdx2* in mature villus cells and because absence of *Cdx2* alone caused a gradually fatal defect in enterocyte differentiation, we assumed that *Cdx1* was unavailable to compensate for its absence.

Here, we report that combined loss of *Cdx1* and *Cdx2* in adult mice significantly enhances the effects of isolated *Cdx2* deficiency on mature, postmitotic enterocytes, rapidly accelerating severe malnutrition and death. These effects were accompanied by profound alterations in intestinal gene expression and the morphology of villus enterocytes. Roughly equal numbers of transcripts were increased and decreased, suggesting the possibility of dual, context-dependent activating and repressive functions for *Cdx* proteins, as demonstrated for other TFs (19, 29). However, using chromatin immunoprecipitation with extensive parallel sequencing of immunoprecipitated DNA (ChIP-seq) for analysis of intestinal villi from wild-type mice revealed direct *Cdx2* binding predominantly at loci with

\* Corresponding author. Mailing address: Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115. Phone: (617) 632-5746. Fax: (617) 582-7198. E-mail: ramesh\_shivdasani@dfci.harvard.edu.

† M.P.V. and H.S. contributed equally to this work.

∇ Published ahead of print on 14 March 2011.

reduced expression in *Cdx2*<sup>-/-</sup> and *Cdx1*<sup>-/-</sup>; *Cdx2*<sup>-/-</sup> intestines, indicating a principal function in transcriptional activation. The combination of phenotypic, gene expression, and DNA occupancy analyses hence establishes the *in vivo* functions of a critical regulator of the self-renewing adult gut epithelium.

## MATERIALS AND METHODS

**Histochemical and immunohistochemical analysis of mouse tissues.** Mice 4 to 6 weeks old were injected with 1 mg of tamoxifen for five consecutive days and euthanized on day 6 or 7. All analyses compared mutants to littermate controls, including *Cdx1*<sup>-/-</sup>; *Cre*<sup>+</sup> and heterozygous mice, neither of which showed differences from the wild type. Tissues were flushed with ice-cold phosphate-buffered saline (PBS) and then with 4% paraformaldehyde (PFA) and incubated in PFA overnight at 4°C. They were then rinsed extensively in PBS for 1 h, dehydrated in a series of ethanols, embedded in paraffin, and sectioned at a 5- $\mu$ m thickness. Hematoxylin and eosin (H&E), Alcian blue, and periodic acid-Schiff (PAS) stains were performed using standard methods. For alkaline phosphatase staining, slides were incubated in nitroblue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolylphosphate (BCIP; Roche). For immunostaining, slides were probed with trefoil factor 3 (TFF3) antibody ([Ab] 1:2,000; gift from D. Podolsky, Massachusetts General Hospital). Alternatively, slides were first treated with 10 mM citrate buffer (pH 6) in a pressure cooker with Ab against chromogranin A (1:500; Immunostar), Cdx1 (1:500; gift from J. Lynch, University of Pennsylvania), lysozyme (1:50; Invitrogen), E-cadherin (1:250; Cell Signaling), ZO-1 (1:200; Invitrogen), or Crs4c (1:1,000; gift from A. Ouellette laboratory, University of California, Irvine). After incubation with primary Ab, slides were treated with biotin-conjugated secondary IgG (Vector Labs, Burlingame, CA), and binding was detected with a Vectastain avidin-biotin-peroxidase complex staining kit (Vector) or tyramide signal amplification (TSA) biotin system (PerkinElmer) and diaminobenzidine substrate (Sigma). Numbers of stained cells are expressed as a fraction of all villus epithelial cell nuclei, and *t* tests were applied to estimate the probability of significance.

**Electron microscopy.** One-centimeter segments of the distal ileum were fixed (2.5% paraformaldehyde, 5% glutaraldehyde, 0.06% picric acid, 0.1 M cacodylate, 0.06% CaCl<sub>2</sub>) overnight or longer at 4°C and embedded in Taab 812 resin (Marivac Ltd., Nova Scotia, Canada). Thin (95 nm) sections were stained with 0.2% lead citrate and visualized on a JEOL 1200 electron microscope at an accelerating voltage of 80 kV.

**ChIP, ChIP-seq, and data analysis.** Mouse jejunal villus epithelium was harvested by incubating freshly dissected 1-cm pieces of tissue in 15-ml conical tubes containing phosphate-buffered saline supplemented with 15 mM EDTA. Samples were agitated by vortexing five times (setting 4.5; Vortex Genie 2) for 5 min each, and villus epithelium was retained atop a 70  $\mu$ m filter each time. Fractions were pooled, inspected visually for purity, and cross-linked with 1% formaldehyde for 15 min at 4°C and for 35 min at 25°C. Cross-linked samples were processed for chromatin immunoprecipitation (ChIP) using CDX2 antibody (Bethyl BL3194) as described previously (39). ChIP material was tested for enrichment of expected fragments and amplified, and the DNA was sequenced using the manufacturer's protocols (Illumina). Sequences were mapped to reference genome *Mus musculus* build 9 (mm9) using ELAND tools, allowing 0 to 2 mismatches (Illumina), and binding peaks were identified by model-based analysis of ChIP-seq (MACS) (44) using default parameters and *P* value cutoffs of 10<sup>-10</sup> or 10<sup>-5</sup>.

**Gene expression microarray and analysis.** Mouse jejunal epithelium was harvested as described above, and RNA was extracted using Trizol (Invitrogen). RNA was either reverse transcribed (SuperScript, Invitrogen) and analyzed by quantitative PCR (Applied Biosystems) or labeled and hybridized to Mouse Genome 430 2.0 expression microarrays according to the manufacturer's instructions (Affymetrix). Data were processed as described below.

**Background coercion/normalization and identification of differentially expressed genes from the microarray data sets.** Gene expression level changes resulting from Cdx2 knockout (KO) were compared among early embryonic, late embryonic, and adult stages. The publicly available microarrays of early and late embryonic knockouts (Agilent Technology) (10, 11) were downloaded from the European Bioinformatics Institute (EBI) ArrayExpress (accession numbers E-MTAB-92 and E-MTAB-218) and processed using the R limma (35) package to detect differentially expressed genes between the KO samples and controls. Our mouse adult Cdx2 KO and Cdx1 Cdx2 double KO (DKO) gene expression microarrays (Affymetrix Mouse Genome 430 2.0) were preprocessed for background coercion and normalization using the robust multichip average (RMA)

function (17) within the Bioconductor affy microarray analysis package, and the limma package was also applied to the preprocessed data. Genes that showed false discovery rates (FDRs) smaller than 5% were finally chosen for further analyses. The FDR cutoff of 5% was universally used for detection of differentially expressed genes unless noted otherwise.

**Clustering/GO analysis for adult Cdx2 KO and Cdx1 Cdx2 double KO microarrays.** The *k*-means clustering was applied to identify sets of genes and modules functionally related to Cdx2 loss or Cdx1 Cdx2 double loss. For this, genes that are increased or decreased (FDR  $\leq$  5%) under either of the knockout conditions were selected and clustered by their relative expression level changes across conditions (*k* = 7) (see Fig. 2G). The individual gene groups (G1 to G7) were analyzed using the DAVID (database for annotation, visualization, and integrated discovery) gene ontology tool (<http://david.abcc.ncifcrf.gov/>) to see whether they matched any specific gene ontology (GO) terms significantly. G2 and G3 were combined for GO analysis because these two groups showed similar gene expression patterns.

**Preprocessing and peak calling of *in vivo* Cdx2 ChIP-seq in mouse villi.** The short DNA sequence reads obtained from Cdx2-bound and input DNA fragments were mapped back to the mouse reference genome (University of California, Santa Cruz [UCSC] version mm9) using ELAND. Among the mapped sequence reads, only ones with no more than two mismatches were retained for peak calling using MACS (44) to ensure high detection quality. *P* value cutoffs of 10<sup>-5</sup> and 10<sup>-10</sup> were used to obtain peak sets at two different confidence levels for comparison. The genomic distribution and other summary statistics of the identified Cdx2 binding sites were given using the *cis*-regulatory element annotation system (CEAS) (30) and other analysis tools (e.g., conservation of the binding sites) within the Cistrome pipeline (<http://cistrome.dfci.harvard.edu/ap/>).

**Association of gene expression with *in vivo* Cdx2 binding.** Three different methods were developed to see how Cdx2 is involved in transcriptional regulation in mouse intestine cells based on gene expression and *in vivo* binding ChIP-seq data. First, we visually represented correlations between the degree of mRNA transcript changes with nearby Cdx2 occupancy (see Fig. 7A). Genes were grouped into bins of 100 according to their expression changes with respect to controls, and the average numbers of nearby Cdx2 binding sites (e.g., 20 kb from the transcriptional start site [TSS]) for the genes in each bin were estimated and visually represented in the form of heat maps (see the bottom yellow-black heat maps in Fig. 7A). A color closer to yellow means that the corresponding gene set has more Cdx2 binding sites near its TSSs. Second, we examined the correlation between regulated genes and their nearest Cdx2 binding sites to appreciate the effect of distance on gene regulation (see Fig. 7B). Every gene was matched with its nearest Cdx2 site, and the distance was calculated and summarized using histograms. Third, the differential expression odds ratio was employed to quantify the potential association between Cdx2 binding and the expressions of the target genes (15). The definition of the differential expression odds ratio is as follows: ( $g_k/g_k^r$ )  $\times$  ( $G^r/G$ ), where  $g_k$ ,  $g_k^r$ ,  $G$ , and  $G^r$  represent, respectively, regulated genes (*r*) with *k* nearby Cdx2 binding loci within a given distance (e.g., 20 kb) from its nearest binding site, nonregulated genes (*m*) with *k* sites within the distance, all regulated genes, and all nonregulated genes (whether located within the distance or not). This statistic informs the likelihood ratio between regulated genes within functional binding sites over nonregulated genes with the same number of binding sites, normalized with respect to the genome background. Thus, a high differential expression odds ratio means that the gene set is likely to be directly regulated by its nearby transcription factor binding.

**Microarray data accession number.** All microarray and ChIP-seq data developed in this study were submitted to the Gene Expression Omnibus (GEO) database under accession number GSE24633.

## RESULTS

**Many Cdx2 functions in the adult mouse intestine are distinct from those in embryos.** The defects that occur upon conditional Cre recombinase-dependent Cdx2 depletion at different stages in mouse intestine development fall along a spectrum. Cre expression from the *Foxa3* promoter, which is active in early gut endoderm at embryonic day 7 (E7), converts distal intestinal cells into a squamous type characteristic of the proximal foregut (11). When Cre is expressed from the *Villin* promoter, which is active in nascent crypt-villus units after E12 (21), the intestine ectopically expresses a few stomach genes

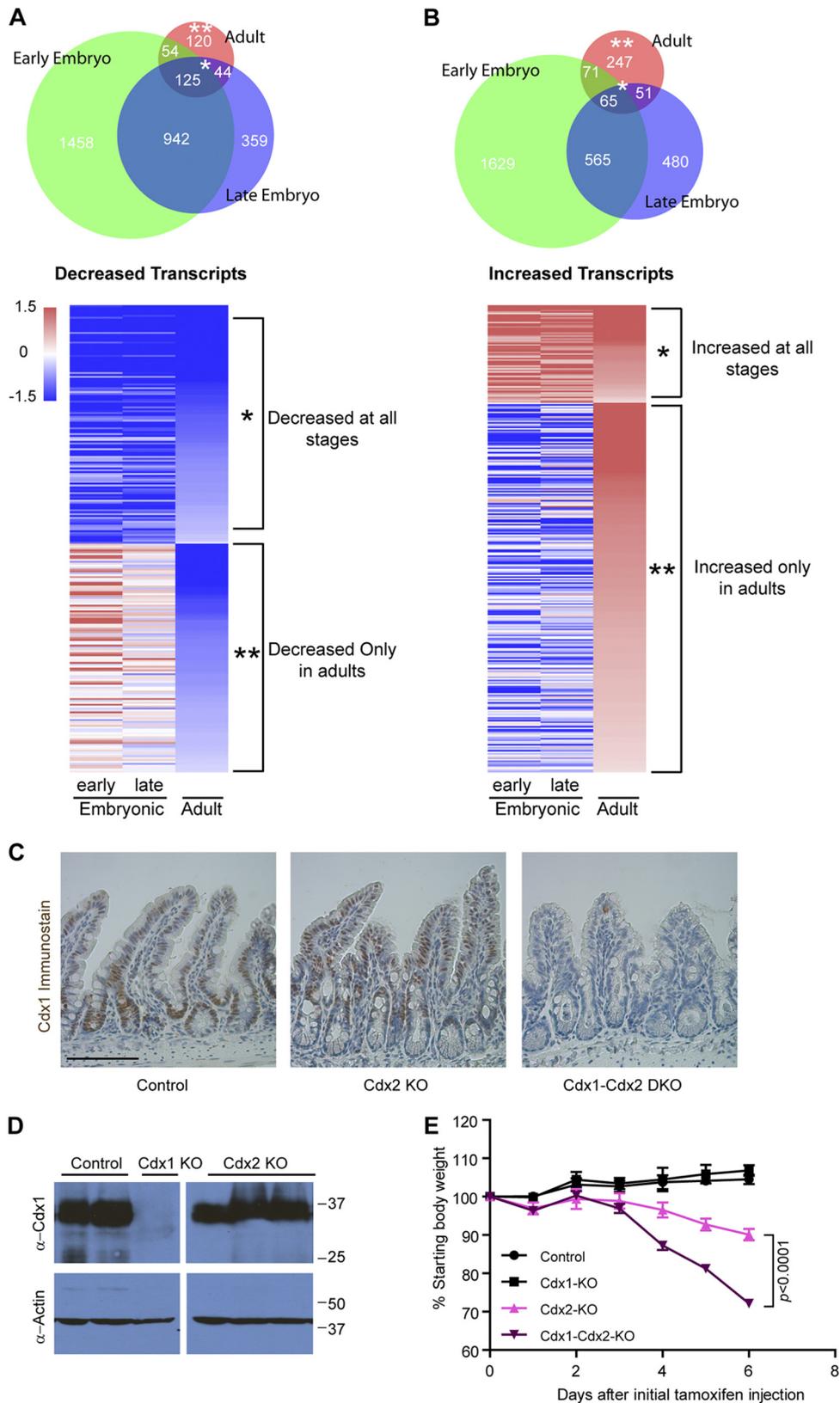


FIG. 1. Stage-specific functions for Cdx2 and functional redundancy with Cdx1 in adult mice. (A and B) Venn diagram representation of genes that decrease (A) or increase (B) in expression with intestine-specific Cdx2 loss in early (11) or late embryos (10), depicted as green and blue circles, respectively, and adult mice (red circles). Transcript numbers confidently altered under each condition are indicated (FDR of <5%). The heat maps display expression levels relative to each sample's internal controls (log<sub>2</sub> scale) for genes dysregulated at all stages (single asterisk) or only in the adult (double asterisk). (C) Cdx1, the only caudal protein coexpressed in the intestine, is detectable by immunohistochemistry in both

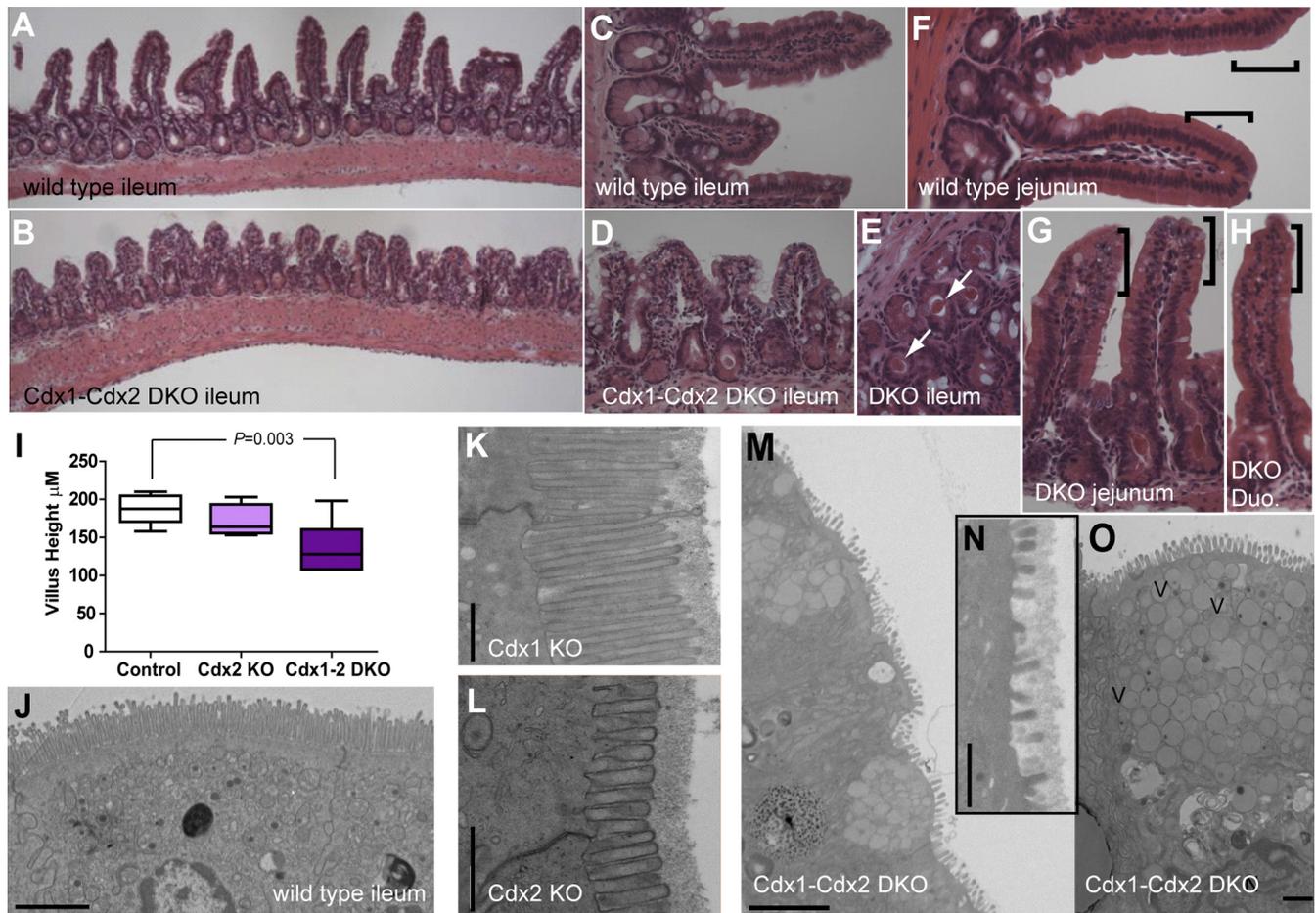


FIG. 2. Additional loss of Cdx1 exacerbates morphological defects in *Cdx2<sup>FV/FV</sup>; Villin-Cre<sup>ER(T2)</sup>* intestines. (A to H) H&E staining of intestinal regions reveals a stunted, disorganized villus epithelium (control in panels A and C versus Cdx1 Cdx2 DKO in panels B, D, and E). DKO crypt lumina are commonly filled with crystalline material (arrows in panel E) never observed in control crypts (A, C, and F). Histology was less severely affected in the proximal DKO intestine (jejunum in panel G and duodenum in panel H) but disorganized, with vacuolated cells with rounded nuclei at the villus tips (brackets). (I) Quantitation of villus height in the ileum from mice of different genotypes; whiskers on the box plots denote maximum and minimum values, and *P* values were determined by a *t* test. (J to O) Transmission electron micrographs of enterocytes from the ileum of controls (wild type in panel J and *Cdx1*<sup>-/-</sup> in panel K), *Cdx2*<sup>-/-</sup> (L), and Cdx1 Cdx2 DKO mice (M, N, and O), highlighting the severely compromised microvillus brush border (M and N) and abundant apical cytoplasmic vesicles (O) in the latter. Microvillus length and density varied among cells depleted of Cdx2 alone but were uniformly scant and stunted in Cdx1 Cdx2 DKO ileum. V, vacuoles in panel O. Scale bars, 2 μm (J and M), 1 μm (K, L, and N), and 0.5 μm (O).

but escapes overt histologic conversion (10). Lastly, conditional *Cdx2* gene inactivation in the adult intestine, stimulated by induction of tamoxifen-responsive Cre expressed under the control of the *Villin* promoter (9), compromises enterocyte function, causing malnutrition and death in about 3 weeks (39). To investigate the basis of these distinct, stage-specific effects, we compared transcriptional profiles from early (*Foxa3-Cre; Cdx2<sup>Flox/Flox</sup>*) and late (*Villin-Cre; Cdx2<sup>Flox/Flox</sup>*) embryonic con-

ditional *Cdx2* null intestines (10, 11) with those induced in adult mice using a tamoxifen-inducible Cre transgene expressed under control of the *Villin* promoter (9) and a distinct floxed *Cdx2* conditional allele (39), *Villin-Cre<sup>ER(T2)</sup>; Cdx2<sup>FV/FV</sup>*. Only a fraction of transcripts was dysregulated at every developmental stage (Fig. 1A and B), representing the minimal set of stage-independent, Cdx2-dependent genes. Many of these genes function in ion transport and lipid metabolism, and they in-

villi and crypts in wild-type mice (left). Villus Cdx1 protein remains detectable in the absence of Cdx2 alone (center), indicating a potential for redundant villus functions in addition to those we previously observed in crypts (39). *Cdx1*<sup>-/-</sup> tissue serves as a negative control for Cdx1 Ab staining. Scale bar, 100 μm. (D) Immunoblot analysis of intestinal epithelium shows similar Cdx1 protein levels in wild-type and Cdx2-deficient mice. *Cdx1*<sup>-/-</sup> intestines provide a negative control, and the blot was reprobed with antiactin Ab as a loading control. (E) Adult mice depleted only of Cdx2 develop clinical signs over 2 to 3 weeks, whereas compound mutants lacking both Cdx1 and Cdx2 lose weight precipitously and invariably succumb within 7 days of the first dose of tamoxifen administered to disrupt *Cdx2* (average of ≥5 animals per set; error bars represent the standard error of the mean; the *P* value was calculated by a *t* test). α, anti.

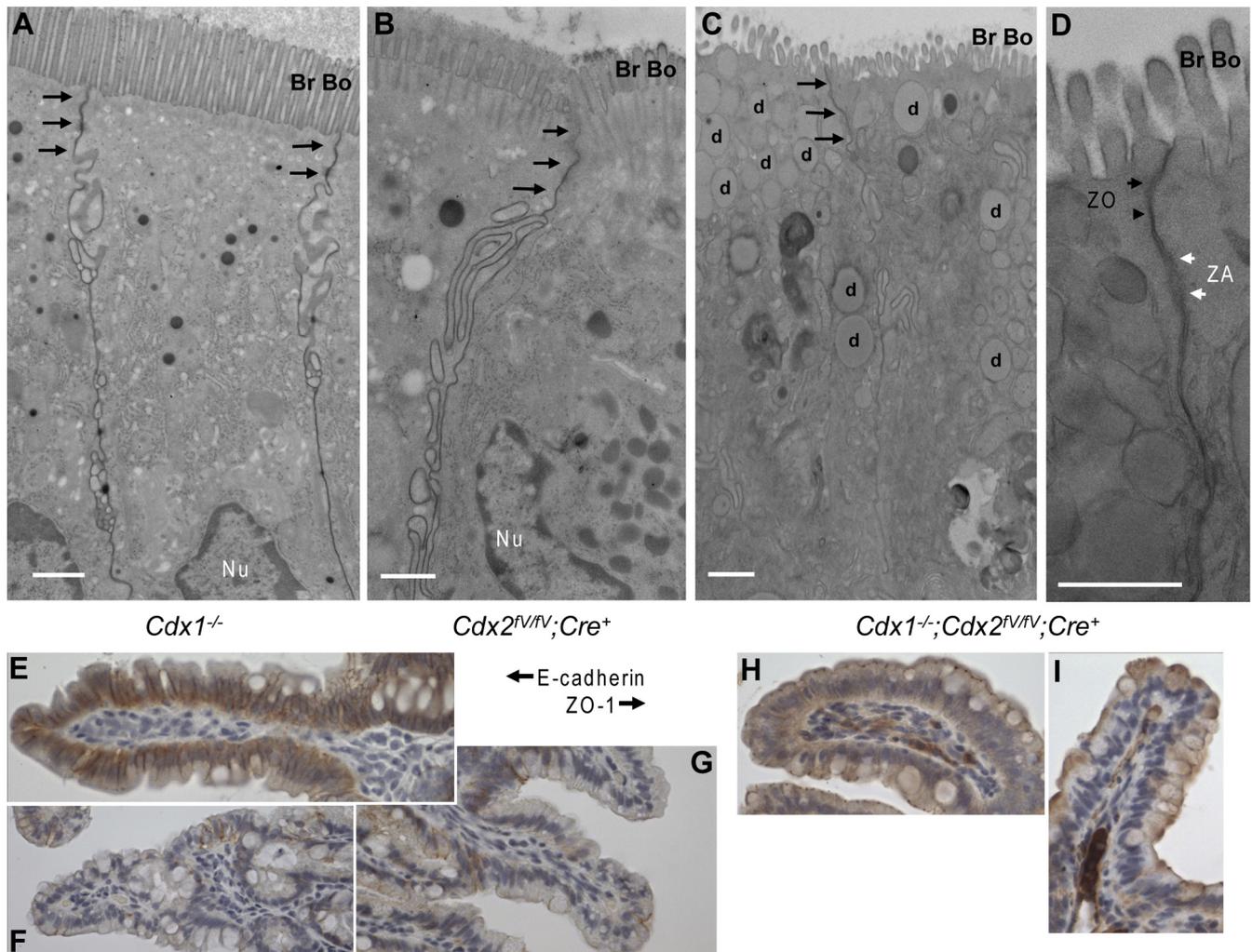


FIG. 3. Intercellular junctions and polarity in adult *Cdx1* *Cdx2* DKO enterocytes. (A to D) In transmission electron micrographs comparing *Cdx1*<sup>-/-</sup> (A) or wild-type (not shown) controls and *Cdx2*-deficient (B) and *Cdx1* *Cdx2* DKO (C and D) enterocytes, apical cell junctions (arrows) appeared overtly intact. The zonula occludens (ZO) and zonula adherens (ZA) are highlighted in the mutant cell shown in panel D. These images further highlight the prominent defect in the apical microvillus brush border (Br Bo) and the presence of innumerable vacuoles (d) in *Cdx1* *Cdx2* DKO cells. (E to I) Immunostaining of control (E and H) and *Cdx1* *Cdx2* DKO (F, G, and I) ileum for the structural proteins E-cadherin and ZO-1. E-cadherin expression is significantly reduced in mutant cells and distributed close to the cell periphery in cells with preserved expression. ZO-1 is not strictly localized to the zonula occludens in mutant cells (I) as it is in control tissue (H). Scale bar, 1  $\mu$ m.

clude the previously proposed *Cdx2* targets *Cdh17* and *Vill* (16, 41). A larger group of transcripts showed *Cdx2* dependence uniquely in the adult intestine (Fig. 1A and B), and some genes changed in opposite directions when *Cdx2* was depleted in fetal or adult intestines, indicating specific and diverse functions in different settings. Thus, *Cdx2* requirements extend beyond development and encompass the continuous regulation of many adult intestine genes.

One reason for discrepant mRNA changes could be that *Cdx1* compensates differently for *Cdx2* loss in adults and embryos. *Cdx1*<sup>-/-</sup> mice have normal intestines (3). *Cdx1* is absent when *Cdx2* function is ablated early in endoderm development (11) but detectable when *Cdx2* is disrupted after the gut epithelium is specified (12). Immunoblot analysis of the adult *Cdx2*-null small intestine showed persistent *Cdx1* protein expression (Fig. 1D). Immunohistochemistry confirmed this result, further revealing *Cdx1* expression in both crypts and villi

(Fig. 1C). These observations indicate that *Cdx1* is available to compensate for *Cdx2* loss in intestinal villi, similar to our previous demonstration of the redundant actions of these two TFs in crypt cell replication (39). We therefore studied *Cdx2*<sup>FV/FV</sup>; *Cdx1*<sup>-/-</sup>; *Villin-Cre*<sup>ER(T2)</sup> mice (here, designated *Cdx1* *Cdx2* DKO) after administration of tamoxifen, hence activating Cre recombinase to disrupt *Cdx2* on a constitutive *Cdx1*-null background. Mice induced to lack all intestinal *Cdx* activity became moribund and lost weight much more rapidly than *Cdx2*<sup>FV/FV</sup>; *Cdx1*<sup>+/-</sup>; *Villin-Cre*<sup>ER(T2)</sup> littermates, indicating potent functional redundancy between the two factors. Whereas *Cdx2* deficiency in the adult intestine allows mice to survive up to 3 weeks, death of all double mutant mice appeared imminent within 2 days after a 5-day course of tamoxifen (Fig. 1E), and necropsy invariably revealed a distended, fluid-filled intestine.

**Epithelial defects in the absence of both caudal proteins.** The intestinal epithelium of adult *Cdx1* *Cdx2* DKO mice was

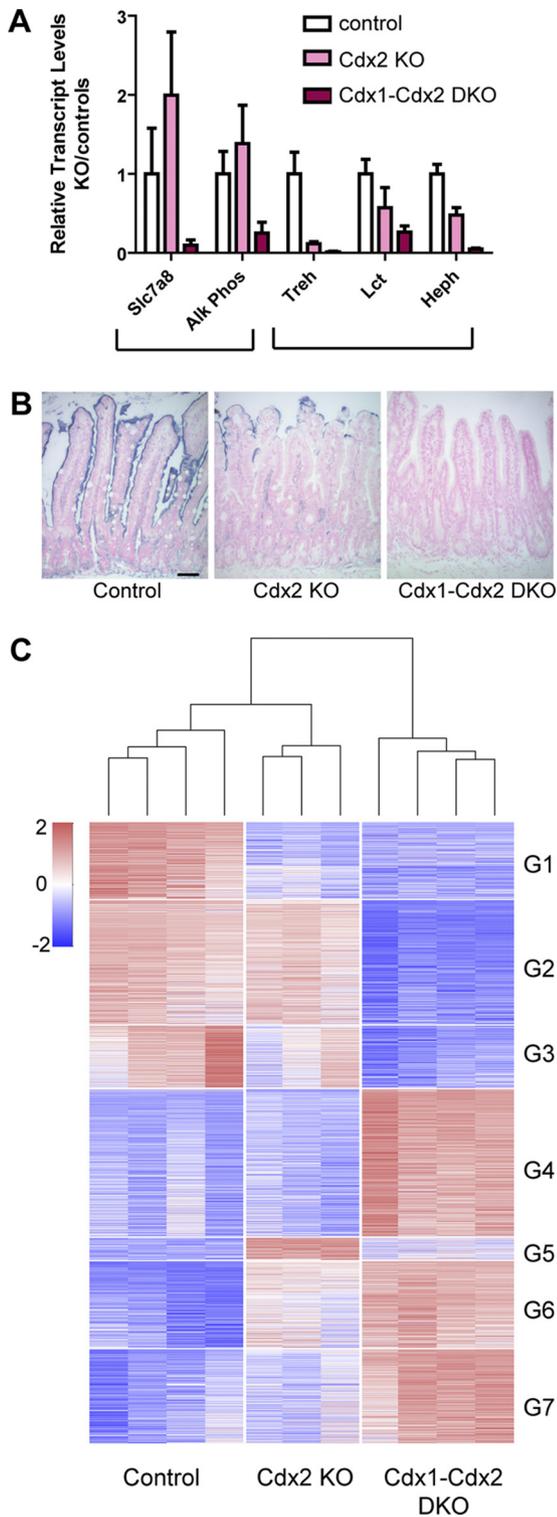
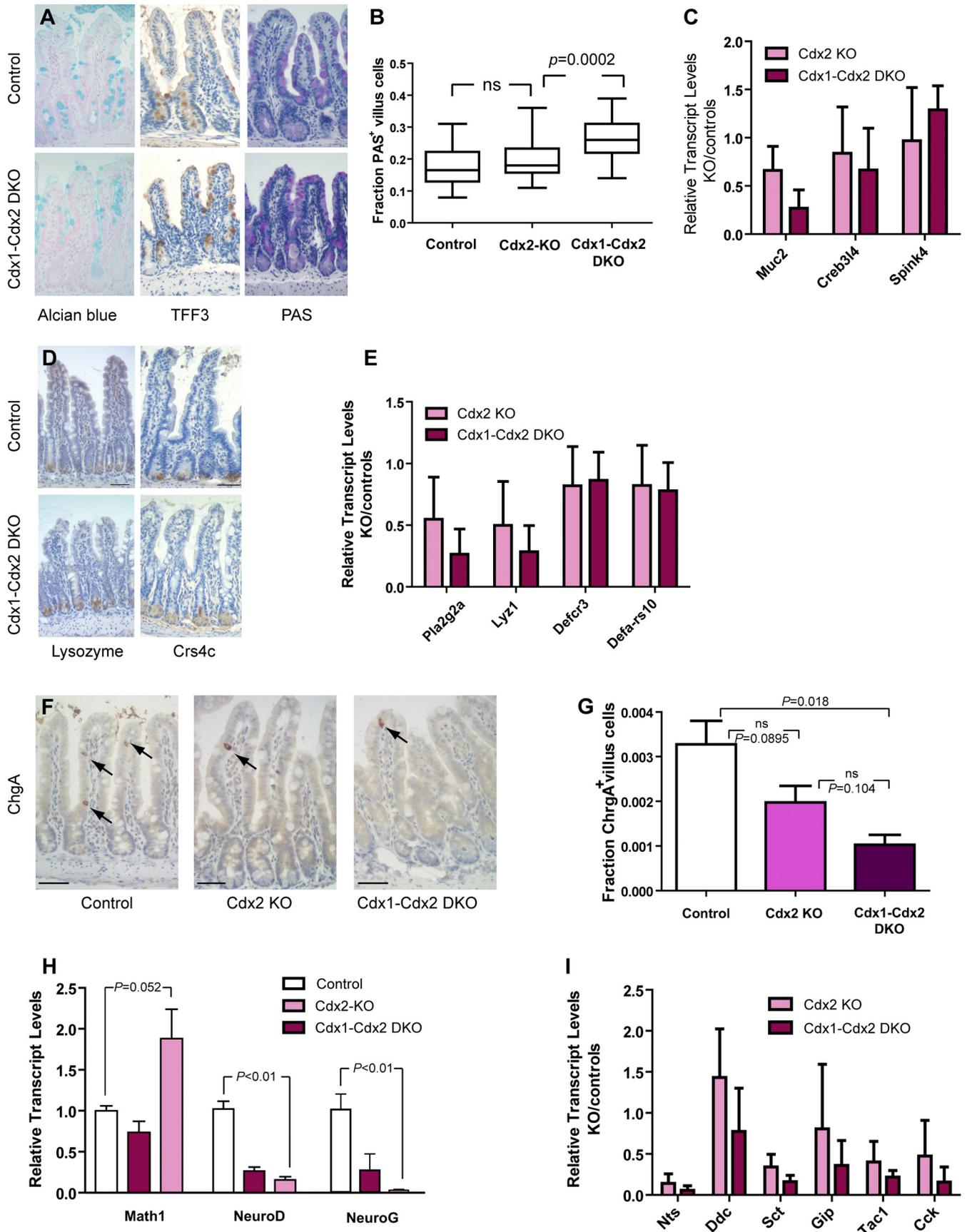


FIG. 4. Additional loss of *Cdx1* exacerbates nearly all transcript decreases observed in *Cdx2<sup>FV/FV</sup>; Villin-Cre<sup>ER(TV)</sup>* intestines. (A) RT-PCR analysis of selected transcripts indicates redundant functions of *Cdx1* and *Cdx2* in transcriptional regulation, with some mRNAs compromised only upon loss of both factors (e.g., *Slc7a8* and *Alpi*) and others affected in *Cdx2* KO but more severely in the double knockout (e.g., *Treh*, *Lct*, and *Heph*). (B) *In situ* alkaline phosphatase activity is diminished in *Cdx2* KO and lost in *Cdx1 Cdx2* DKO jejunum. (C) Analysis of microarray expression data from control, *Cdx2* KO, and *Cdx1 Cdx2* DKO jejunal epithelia using a color scale where red

markedly abnormal, with severe to moderate diminution of villus height in the distal and proximal intestine, respectively (Fig. 2A, B, and I; also see Fig. S1 at <http://research4.dfci.harvard.edu/shivdasani/pubs/supplementary>). In contrast to *Cdx1<sup>-/-</sup>* and *Cdx2<sup>-/-</sup>* intestines, where villus cells appear normal (3, 39) or nearly normal (Fig. 2C and F), cells lining the stunted *Cdx1 Cdx2* DKO villi showed frequent cytoplasmic vacuolation, especially in the top one-third to one-half of the villi, where the epithelium was also disorganized (Fig. 2D, G, and H). These defects were most prominent in the ileum (Fig. 2D), where *Cdx2* expression is normally highest (18), and less severe but clearly discernible in the jejunum (Fig. 2G) and duodenum (Fig. 2H). Crypt lumina throughout the intestine were filled with a crystalline, eosinophilic debris (Fig. 2D and E) that was absent in *Cdx1<sup>-/-</sup>* or *Cdx2<sup>-/-</sup>* intestines. Caspase-3 Ab staining did not reveal excessive apoptosis at the villus tips compared to control intestines (data not shown). Ultrastructural analysis of *Cdx1 Cdx2* DKO enterocytes confirmed the frequent presence of numerous apical multivesicular inclusions (Fig. 2O) and revealed a rudimentary brush border, with substantial reduction in the numbers, height, and density of apical microvilli (Fig. 2M and N). The continuum of structural defects across *Cdx* genotypes was most apparent in the brush border, where microvilli in wild-type (WT) or *Cdx1<sup>-/-</sup>* enterocytes were normal (Fig. 2J and K), and those in *Cdx2<sup>-/-</sup>* enterocytes were generally shorter, wider, and less dense (Fig. 2L), exactly as reported in fetal *Cdx2*-deficient mice (10). However, whereas the adult *Cdx2<sup>-/-</sup>* brush border showed variation from cell to cell, DKO microvilli were uniformly sparse and severely stunted (Fig. 2M and N), with some cells carrying only a handful. This microvillus defect was apparent along the full villus length (Fig. 2M). Our gene expression and histology results combine to suggest that *Cdx1 Cdx2* DKO mice die of starvation as a result of global dysregulation of genes associated with terminal digestion, nutrient absorption, and proper assembly of the apical microvillus brush border.

The combination of reduced villus height, vacuolated cytoplasm, and rudimentary brush border could reflect an underlying defect in cell polarity, as described in fetal mouse intestines lacking *Cdx2* (10). To evaluate cell polarity in *Cdx1 Cdx2* DKO intestinal villus cells, we first examined apical intercellular junctions. Electron microscopy revealed overtly intact zonula occludens and zonula adherens structures (Fig. 3A to D), consistent with the histologic evidence for a well-preserved epithelial barrier (Fig. 2B and D). However, immunohistochemistry revealed reduced staining with E-cadherin Ab (Fig. 3E to G), and, in the few cells with abundant E-cadherin, the

and blue denote high and low expression, respectively. Unsupervised hierarchical clustering by samples (columns) showed high consistency between replicates, as reflected in the dendrogram at the top. *k*-Means clustering (*k* = 7) identified gene sets with similar behaviors across samples, revealing two principal patterns of regulation by caudal proteins. All genes regulated in either *Cdx2* KO or *Cdx1 Cdx2* DKO tissue were selected prior to clustering (FDR ≤ 5%; rows). Transcripts in clusters G1 and G6 change in *Cdx2* KO and more robustly in compound mutants, whereas transcripts in the remaining clusters change only when both *Cdx* factors are inactive. Scale bar, 50 μm.



domain of expression often extended to the cell periphery (Fig. 3F and G). Control tissues (Cdx1 Cdx2 DKO lacking Cre) showed the expected punctate distribution of the zonula occludens marker ZO-1 (Fig. 3H). In contrast, ZO-1 rarely showed the same distribution when both Cdx1 and Cdx2 were missing, and although ZO-1 transcript levels were not reduced, ZO-1 appeared diffusely in association with the apical and basement membranes (Fig. 3I). The sum of these defects resembles the pattern described in Cdx2-deficient fetal intestines (10); taken together with the preservation of intercellular junctions and presence of rudimentary microvilli, the findings implicate Cdx function in selected, but not all, aspects of adult intestinal epithelial cell polarity.

The polarity defect in Cdx2-deficient fetal intestine is attributed in part to Cdx2 control over a transcriptional program for endolysosomal and vacuolar products (10). We therefore assessed RNA expression in adult intestines of the 82 lysosomal and endolysosomal genes that are reduced in mice with fetal Cdx2 deficiency. Consistent with the limited total overlap in gene dysregulation (Fig. 1), only 9 of 40 lysosomal genes and 14 of 43 endolysosomal genes showed reduced expression in adult Cdx2-deficient intestines (see Table S2 at <http://research4.dfci.harvard.edu/shivdasani/pubs/supplementary>). mRNA encoding the transcription factor Tefeb, which is believed to control many of these genes (10), was also unaffected in adults, and some transcripts that were reduced in fetal intestines were increased in adult DKO intestines. Thus, both fetal and adult intestinal villus cells show similar polarity defects in spite of a limited overlap in gene dysregulation.

**Transcriptional redundancies between Cdx1 and Cdx2 in the adult intestine.** Because expression of many intestinal genes is thought to depend on Cdx2 (13), we ascertained Cdx redundancies at the level of gene expression, first using reverse transcription-PCR (RT-PCR). Certain enterocyte transcripts, illustrated by *Treh*, *Lct*, and *Heph*, decreased upon Cdx2 depletion and were further compromised in the absence of Cdx1 (Fig. 4A). Other transcripts, such as *Slc7a8* and *Alpi*, were largely intact in *Cdx2* mutants and were reduced only when both Caudal factors were missing. Alkaline phosphatase histochemistry corroborated the latter result. Although *Alpi* transcript levels are normal in *Cdx2* mutants, the enzyme is expressed inefficiently at the apical membrane, as also reported with fetal *Cdx2* inactivation (10); loss of both Cdx proteins led to markedly reduced transcript levels and undetectable enzyme activity (Fig. 4B). RNA microarray analysis confirmed on the genome scale that absence of Cdx1 enhances the effects of iso-

lated Cdx2 deficiency on gene expression. Compound mutant intestines showed greater numbers of significantly altered transcripts (775 transcripts increased and 691 decreased in Cdx2 KO mice versus 3,899 increased and 3,013 decreased in Cdx1 Cdx2 DKO mice at a false discovery rate [FDR] of  $\leq 5\%$ ) as well as greater magnitudes of change (see Table S2 at <http://research4.dfci.harvard.edu/shivdasani/pubs/supplementary>). An unsupervised machine-learning approach using *k*-means clustering identified seven groups of genes (Fig. 4C). Cluster G1, containing transcripts that decrease upon solitary inactivation of Cdx2 and decline further in double mutants (e.g., *Lct* and *Treh*) (Fig. 4A), is enriched for gene ontology (GO) terms associated with enterocyte functions, including phospholipid and ion transport (see Table S1 at <http://research4.dfci.harvard.edu/shivdasani/pubs/supplementary>). Transcripts in clusters G2 and G3 were barely affected in *Cdx2* single mutants but significantly reduced in double mutant mice (e.g., *Slc7a8* and *Alpi*) (Fig. 4A). Overrepresented in these clusters were GO categories related to cell proliferation, corroborating our prior demonstration of Cdx redundancy in crypt cell replication (39), and mature intestinal functions, further suggesting redundant functions in villus enterocytes. Clusters G1 to G3 highlight the redundant regulation of hundreds of genes and the variable degree to which Cdx1 compensates for Cdx2 loss at individual loci. At every FDR, approximately equal numbers of transcripts were increased and decreased in Cdx1 Cdx2 DKO intestines, as represented in clusters G4 to G7; we return below to consideration of such genes.

**Cdx functions in secretory cell lineages.** The intestinal epithelium carries three secretory cell types: mucin-producing goblet cells, hormone-secreting enteroendocrine (EE) cells, and, at the base of each crypt, microbicidal-secreting Paneth cells (1). Each of these cell types expresses Cdx2 although levels consistently seem lowest in Paneth cells, a lineage recently shown to be suppressed by Cdx2 overexpression (7). Staining with Alcian blue, Trefoil factor 3 (TFF3) antibody, and periodic acid-Schiff (PAS) revealed a small, statistically significant increase in goblet cell numbers in tamoxifen-treated DKO intestines (Fig. 5A and B) although transcript levels of goblet cell genes were barely affected (Fig. 5C). Likewise, Paneth cell-specific mRNA levels were largely intact, as was immunoreactivity for the Paneth cell products lysozyme and *Crs4c* (Fig. 5D and E). The most notable secretory cell defect in Cdx1-Cdx2 DKO intestines was a reduced number of EE cells (Fig. 5F and G). Expression of *Math1*, a TF that controls differentiation of all secretory cells and is induced upon ectopic

FIG. 5. Reduced enteroendocrine cell numbers in the absence of caudal factors. Intestinal secretory cell lineages were examined in Cdx2 KO and Cdx1 Cdx2 DKO mice. Goblet cells, defined by Alcian blue-positive, periodic acid-Schiff-positive (PAS<sup>+</sup>), and anti-TFF3-positive mucous granules (A) were increased slightly in number in Cdx1 Cdx2 DKO mice (B), but transcripts intrinsic to these cells were not significantly altered (C). Cell counts are represented as box plots with whiskers representing the highest and lowest counts; error bars indicate standard deviations. Representative Paneth cell proteins lysozyme and *Crs4c* (D) and mRNAs *Pla2g2a*, *Lyz1*, *Defcr3*, and *Defa-rs10* (E) were also expressed at nearly similar levels in control, Cdx2 KO, and Cdx1 Cdx2 DKO intestines. Bars indicate standard deviations. In contrast, chromogranin A (ChgA) Ab staining (F; arrows) revealed modestly reduced numbers of enteroendocrine cells in Cdx2 KO ileum and significantly reduced numbers in Cdx1 Cdx2 DKO ileum (G). Average fractions of enteroendocrine cells are plotted  $\pm$  standard errors of the means. This reduction was accompanied by lower expression of the endocrine cell TF mRNAs *NeuroD1* and *Neurogenin3*, whereas levels of *Math1*, a TF that controls all secretory cell differentiation, were not reduced (H). Bars represent standard errors of the means, and the statistical significance of differences between tissues of different genotypes is indicated. Transcripts encoding enteroendocrine cell products were also suppressed (I). Bars represent the standard deviations. Transcript levels were assessed by quantitative RT-PCR and graphed relative to littermate controls for at least five animals of each mutant genotype. The *P* value was calculated by a *t* test. Scale bar, 50  $\mu$ m.

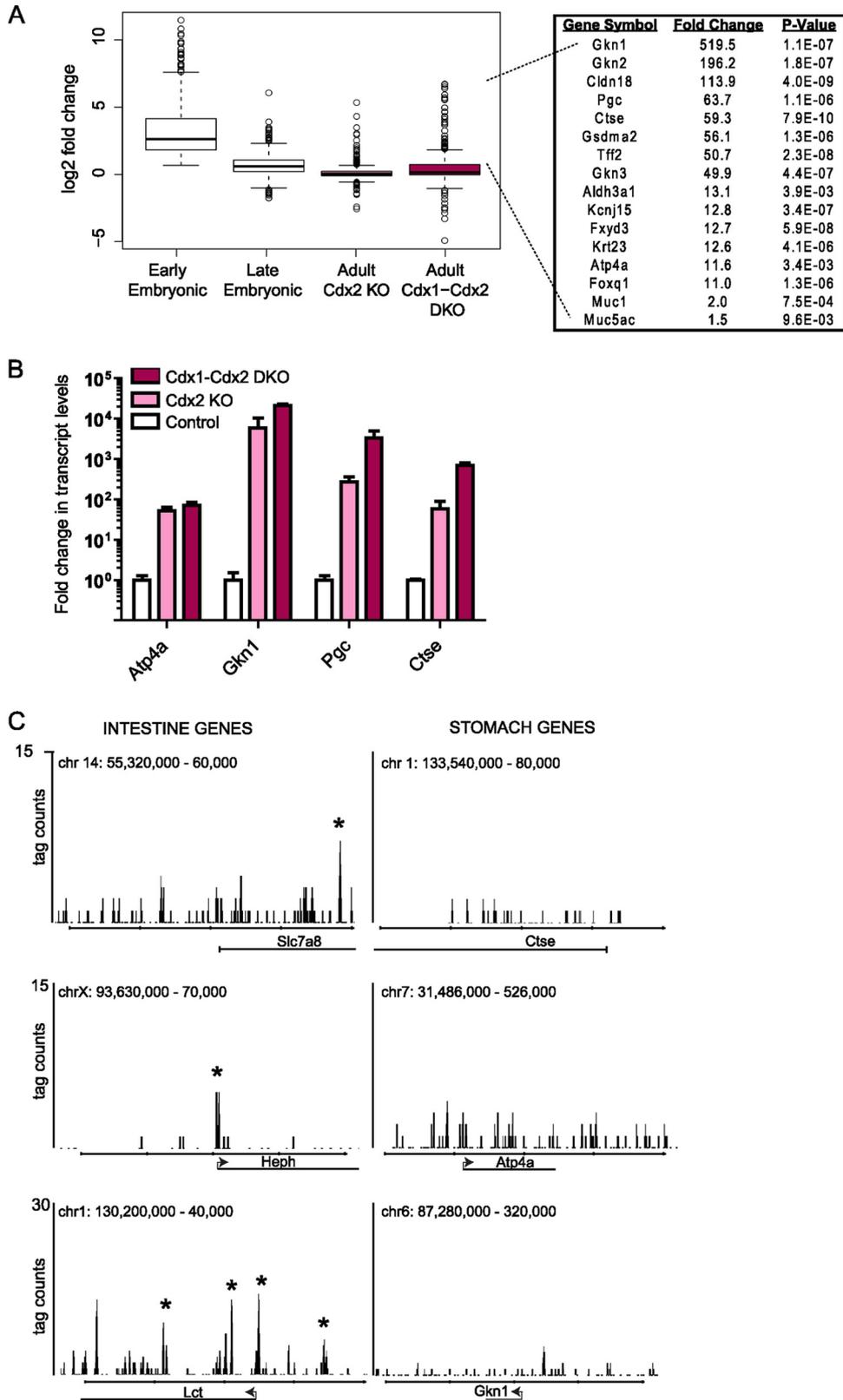


FIG. 6. Limited plasticity for homeotic conversion of the intestinal epithelium and indirect repression of foregut genes. (A) Comparison in early (11) and late (10) embryos and adult Cdx2 KO and Cdx1 Cdx2 DKO mice of the expression levels of the fifth percentile of transcripts that increased in Foxa3-Cre Cdx2 knockout intestines (here, designated early embryos; these transcripts are enriched for ectopically expressed foregut genes). Gene expression relative to control mice is plotted on a log<sub>2</sub> scale. Each box indicates the interquartile range (25% from the median), and the whiskers indicate 2.5 times the interquartile range. Most transcripts that increase in early embryos were unaffected in the adult, but a small fraction,

Cdx2 expression in the stomach (23, 31), was preserved. However, mRNAs encoding *Ngn3* and *NeuroD*, TFs that operate downstream of *Math1* in EE cells, and several endocrine end products were reduced (Fig. 5H and I). EE cell deficiency was statistically significant in Cdx1 Cdx2 DKO but not in Cdx2 single-mutant intestines (Fig. 5G), suggesting that Cdx1 and Cdx2 function redundantly to regulate EE cell differentiation. In sum, loss of caudal proteins has a modest effect on adult intestinal secretory cells, with the most significant effects occurring in cell numbers and specific products of the EE lineage.

**In vivo role of Cdx proteins in plasticity of adult gut epithelial gene expression.** In early Cdx2-deficient mouse embryos, the distal intestine adopts foregut features, including high expression of keratins and other esophageal genes (11). This homeotic effect is not apparent when Cdx2 is inactivated later in gestation (10, 12), suggesting that it reflects a short window, between E8 and E12, when gut endodermal fate is pliable and Cdx2 imposes intestinal identity. Indeed, we observed that transcripts with the largest quantitative increase in *Foxa3-Cre*; *Cdx2<sup>Flox/Flox</sup>* embryos, the fifth percentile that is heavily enriched for foregut genes (11), were largely silent or similar to controls in both *Villin-Cre*; *Cdx2<sup>Flox/Flox</sup>* fetuses (10) and tamoxifen-treated *Cdx2<sup>FV/FV</sup>*; *Villin-Cre<sup>ER(T2)</sup>* adults on *Cdx1<sup>-/-</sup>* or *Cdx1<sup>+/+</sup>* backgrounds (Fig. 6A). Furthermore, acute Cdx loss in adults did not induce foregut squamous morphology in the ileum (Fig. 2B). Expression analysis nevertheless revealed high expression levels of several outlier stomach-specific transcripts in adult Cdx-depleted intestines, represented in the dots that fall outside the range of the box plots and highlighted in the adjoining table (Fig. 6A). Compared to the tight silencing of such genes in wild-type mice, RT-PCR analysis confirmed significant expression of representative examples in the absence of Cdx2 and further increases upon additional loss of Cdx1 (Fig. 6B; note  $\log_{10}$  scale on the y axis). These results indicate that the adult intestine retains plasticity sufficient to express a limited number of heterologous genes and that Cdx TFs normally suppress this latent potential.

**Analysis of genome-wide Cdx2 binding in vivo suggests that it functions principally to activate genes.** Similar numbers of transcripts increased and decreased in both Cdx2 single knockout and Cdx1 Cdx2 DKO intestines at all FDRs (Fig. 4C and data not shown), and the foregoing analysis showed that some increased transcripts correspond to foregut genes. Taken together, these results point to the possibility of alternative activating or repressive Cdx functions at different loci. Indeed, silencing of heterologous genes is as important in cell differentiation as activation of tissue-specific genes; critical regulators of cellular identity, in particular, might serve both functions, as is known for some TFs (26, 43).

To identify direct targets of Cdx2 activity and to resolve

whether Cdx2 functions as an activator, repressor, or both, we used chromatin immunoprecipitation (ChIP) and massively parallel sequencing of immunoprecipitated DNA (ChIP-seq) to map its physical occupancy in the genome of primary intestinal epithelial cells in wild-type mice. We focused attention on intestinal villi because mutant enterocyte phenotypes were prominent (Fig. 2 to 4) and because uncontaminated crypt cell numbers are limiting for accurate whole-genome ChIP. A lack of antibodies suitable for ChIP precluded separate mapping of Cdx1 occupancy, and, in any event, the redundant phenotypes and RNA expression profiles in knockout mice (Fig. 4) indicate that the two factors bind many of the same sites. Cdx2 ChIP-seq in primary mouse intestinal villus cells yielded  $6.8 \times 10^6$  sequence tags, giving  $5.95 \times 10^6$  uniquely mappable reads, and identified 8,775 occupied sites (at a  $P$  value of  $\leq 10^{-5}$ ; 1,976 sites at a  $P$  value of  $\leq 10^{-10}$ ) (see Fig. S3 at <http://research4.dfci.harvard.edu/shivdasani/pubs/supplementary>). Cdx2 binding sites were conserved across species and heavily enriched for the known Cdx-binding motif (2), indicating detection of bona fide Cdx2 occupancy (see Fig. S3). Clusters of sequence tags mapped near many enterocyte genes that showed reduced expression in the absence of Cdx2 (Fig. 6C, left column), suggesting that it might activate such genes directly. Conversely, we rarely observed Cdx2 binding within 40 kb of genes that increase in expression when Cdx2 is missing (Fig. 6C, right column), including foregut-specific genes. These data hint that Cdx2 rarely represses gene expression directly.

To determine if these selected results apply generally, we examined the global frequency of Cdx2 occupancy near genes that decline in expression (candidate targets for gene activation), increase expression (repressed genes), or remain stable in Cdx2 KO or Cdx1 Cdx2 DKO tissue. In the heat map shown at the top of Fig. 7A, each segment represents bins of 100 dysregulated genes, ranked from those containing genes with the most reduced expression (blue) to the most increased (red; black is neutral) in Cdx2 KO jejunum. The companion heat map plots the frequency of Cdx2 binding sites identified at high confidence ( $P$  value of  $\leq 10^{-10}$ ) within 20 kb of genes in the corresponding expression bin. Cdx2 binding was strongly associated with the set of downregulated genes (Fig. 7A, left side) and not at all with genes that increase expression (right side), providing powerful, genome-scale evidence that Cdx2 functions primarily, and perhaps exclusively, to transactivate intestinal target genes. Associations between Cdx2 occupancy and gene regulation were stronger in Cdx1 Cdx2 DKO intestines (Fig. 7A, bottom) than in single Cdx2-null intestines (Fig. 7A, top) when changes in expression were considered although the trends were similar in both cases.

To strengthen this conclusion and to better define Cdx2 transactivation preferences, we first examined the relationship

---

represented by the small circles falling outside the range of about 62.5% from the median, is notably increased, consistent with gene derepression. These genes are highlighted in the adjacent table with  $P$  values and fold changes observed in the adult Cdx1 Cdx2 DKO mice. (B) Expression of representative stomach-specific transcripts assessed by quantitative RT-PCR on four independent adult intestine samples. Note the  $\log_{10}$  scale on the y axis. (C) Cdx2 ChIP-seq analysis in wild-type intestine revealed no Cdx2 binding at such derepressed foregut-specific loci (stomach genes), suggesting that Cdx-dependent effects on their transcription may be indirect. In contrast, ChIP-Seq readily revealed Cdx2 occupancy at many loci that reduce expression in KO intestines (intestine genes). Each representation of sequence tag counts depicts a 40-kb window of the mouse genome (build mm9) with genes indicated by a horizontal black line. Arrows denote transcriptional start sites, and the asterisks mark regions in which Cdx2 binding significantly exceeds the experimental background.

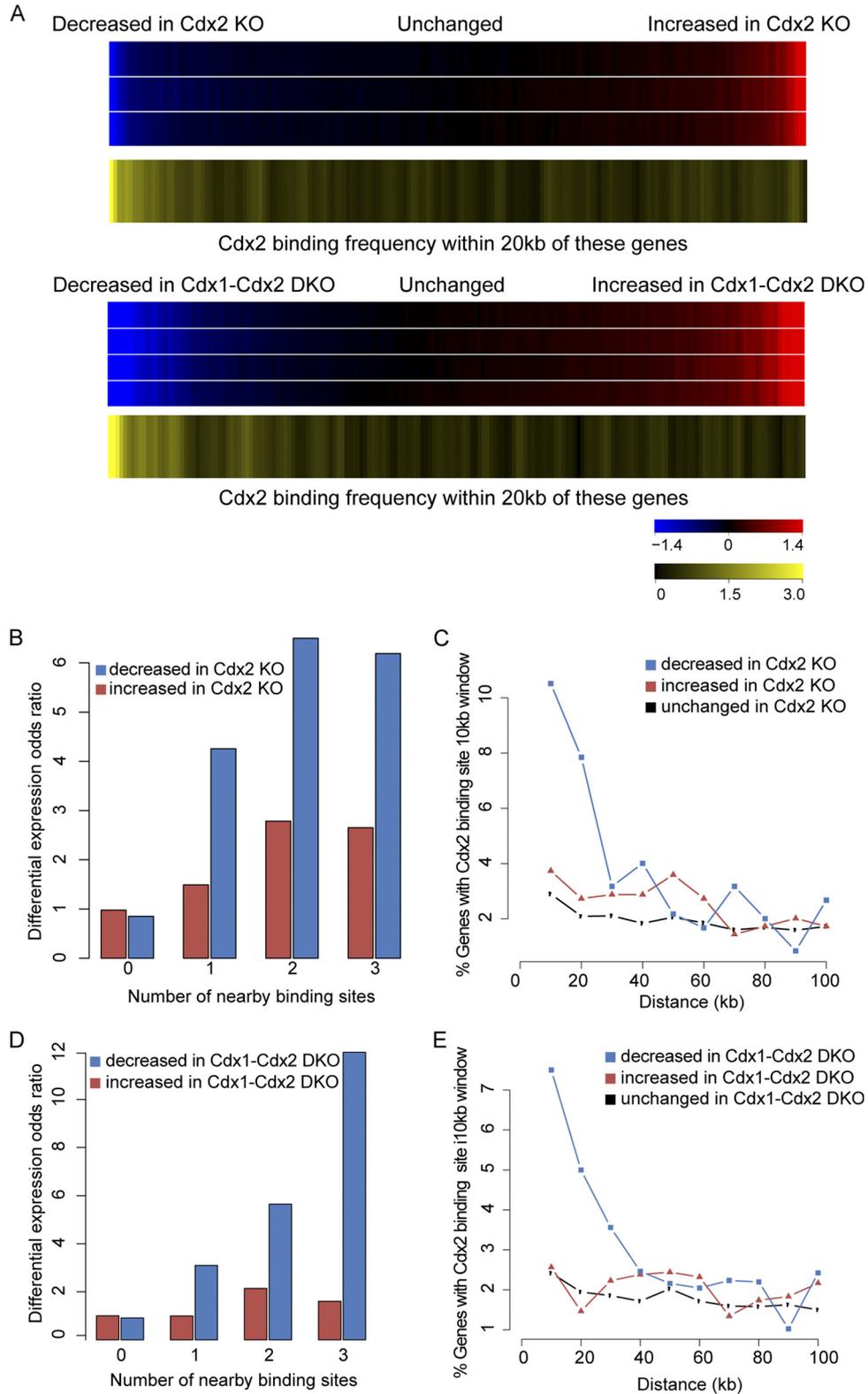


FIG. 7. Cdx2 functions principally as a transcriptional activator in adult intestinal villi. (A) The upper heat maps display  $\log_2$  fold change in gene expression in three or four replicates of Cdx2 KO or Cdx1 Cdx2 DKO intestinal epithelium compared to wild-type tissue, respectively. Each heat map spans the spectrum from a marked decrease to a marked increase. Each vertical bin includes 100 genes, and its color indicates the average relative expression level for the group. The lower heat maps plot the average number of Cdx2 binding sites located within 20 kb of the TSSs of the corresponding genes in each bin displayed in the expression heat map, encoded by the intensity of yellow shading. Genes that are significantly reduced in expression in mutant epithelia tend to accommodate much greater Cdx2 binding (left), whereas genes that increase expression in the mutant intestine show no such tendency. These results indicate that Cdx2 functions primarily to activate rather than repress transcription. (B and

between the number of binding sites near a gene and the likelihood of its dependence on Cdx2. Even a single nearby binding site within 20 kb of a TSS increased the odds ratio of differential expression for genes that decline upon loss of caudal factors; multiple Cdx2 binding sites (e.g., *Lct*, represented in the left side of Fig. 6C) conferred increasing likelihood of reduced expression in the absence of Cdx proteins (Fig. 7B and D). Conversely, we observed little association between numbers of Cdx2 binding sites and the set of upregulated genes, a result that held true for genes that alter expression with loss of Cdx2 alone (Fig. 7B) or upon combined loss of Cdx1 and Cdx2 (Fig. 7D). Next, we evaluated the distribution of distances between Cdx2-bound regions and annotated transcription start sites (TSSs). Only 3.6% of Cdx2-occupied sites fell <1 kb from a TSS, within classical promoters (MACS,  $P$  value  $\leq 10^{-5}$ ) (see Fig. S3 at <http://research4.dfci.harvard.edu/shivdasani/pubs/supplementary>); most sites lay in introns or intergenic regions at a median distance of 38.6 kb from the nearest TSS (see Fig. S3). The observation that genes downregulated in Cdx2 KO or Cdx1 Cdx2 DKO intestines were more frequently bound by Cdx2 than genes that increased or stayed the same was apparent at distances of >20 kb from TSSs (Fig. 7C and E). Not only was Cdx2 binding near increased and nonregulated genes less frequent, but it also distributed randomly in relation to TSSs (Fig. 7C and E). Genes that increased or stayed the same upon Cdx loss showed similar patterns, distinguishing likely direct transcriptional targets from all other genes. Distribution of binding distances indicated that Cdx2 typically occupies regulated loci 2 kb to 15 kb from the TSS although binding also occurs at far greater distances and within promoters (data not shown). Details on the number and distances of Cdx2-bound sites near regulated genes are publicly available (see "Microarray data accession number" above). These results collectively argue that Cdx2 principally activates intestinal epithelial genes through binding at distal regulatory regions.

## DISCUSSION

Although the intestine-restricted homeodomain protein Cdx2 is required to specify intestinal epithelium during development, its functions in the adult organ were unclear. Here, we demonstrate redundant Cdx protein requirements in adult intestinal structure, function, and gene expression, and we relate these requirements to genome-wide Cdx2 binding in primary mouse intestinal villus cells. This analysis allowed delineation of likely direct transcriptional targets and suggests that Cdx2 operates mainly as a transcriptional activator rather than a repressor. The combination of *in vivo* function, gene expres-

sion, and DNA occupancy provides an unprecedented view of regulation of diverse genes through *cis*-regulatory regions that depend on one or both Cdx proteins for their full activity. These TFs act primarily at distal regulatory regions to activate genes that control many core aspects of terminal intestinal epithelial differentiation and are necessary for adult nutrition and survival.

Loss of Cdx1 alone does not affect bowel structure or function (3), and its requirement is unmasked only upon loss of Cdx2. It therefore follows that Cdx2 function in the adult intestine is more vital than that of Cdx1. One possibility is that Cdx2 function prevails by virtue of differences in protein structure. However, Cdx1 and Cdx2 have similar sequences, particularly within the DNA-binding homeodomain, and recognize the same DNA sequence (2). Although differences outside the homeodomain include a serine 60 residue in Cdx2 that can be phosphorylated to regulate transcriptional activity (25), *Cdx2* knock-in into the *Cdx1* locus rescues vertebral patterning defects in *Cdx1* mutant embryos, indicating functional equivalence (27). The simpler explanation, therefore, is that Cdx2 is more abundant and also expressed throughout the crypt-villus unit, whereas Cdx1 expression in wild-type mice is more prominent in crypt than in villus cells (32, 34). It is also conceivable that the profound defects we observe in Cdx1 Cdx2 DKO villus cells have their origins in crypt cell precursors that express both TFs abundantly. However, Cdx1 is present in wild-type and *Cdx2*<sup>-/-</sup> crypts (Fig. 1C), and candidate target genes are almost always more severely affected in Cdx1 Cdx2 DKO intestines than in Cdx2 single-mutant intestines (Fig. 4C). Moreover, we uncovered clear associations between gene expression in Cdx1 Cdx2 DKO intestines and Cdx2 binding in villus cells (Fig. 7). Taken together, these observations suggest simple redundancy as the best explanation for the compound mutant phenotype, just as Cdx2 levels increase in the absence of Cdx1 (3); Cdx2 probably assumes all the functions of Cdx1. Future availability of Cdx1 antibodies suitable for ChIP might allow a rigorous test of this idea. Dosage of caudal genes has an important role in elongation of the embryonic posterior axis as heterozygous loss of each factor adds to the defects that occur with complete loss of the other (37). Although similar gene dosage effects proved difficult to quantify in the intestine, we suspect that *Cdx1*<sup>+/-</sup>; *Cdx2*<sup>FV/FV</sup> intestinal defects are intermediate between those in the *Cdx2*<sup>-/-</sup> and Cdx1 Cdx2 DKO organ.

Cdx2 inactivation in embryos caused homeotic conversion of the distal intestinal epithelium (11, 12), whereas ectopic expression in the stomach induced intestinal properties (22, 27, 33). Cdx2 manipulation in each of these studies commenced

---

D) Genes that decline in Cdx-deficient epithelium are more likely to harbor multiple Cdx2 binding events than genes that increase expression. Odds ratios for differential expression in single (B) or double (D) mutant tissues were evaluated in relation to the number of Cdx2 binding sites within 20 kb of TSSs of decreased (blue) or increased (red) genes, hence quantifying direct Cdx2 binding at loci that respond to Cdx protein loss compared to the genome-wide background for Cdx2 binding (refer to Materials and Methods for mathematical details). (C and E) Distribution of TSS distances in 10-kb windows from the nearest Cdx2 binding sites. Genes with an FDR of  $\leq 5\%$  were determined as either decreased or increased in Cdx2 KO (C) or Cdx1 Cdx2 DKO (E) intestines compared to controls, and the remaining genes were considered nonregulated. The graph plots the fraction of all genes in the indicated category containing a Cdx2 binding site within the indicated window. Downregulated genes are more likely to have nearby Cdx2 binding sites than up- or nonregulated genes. All analysis was done using Cdx2 binding sites predicted by MACS at a  $P$  value of  $\leq 10^{-10}$ .

before birth, possibly during a window of developmental plasticity. In adult mice, in contrast, even combined Cdx gene disruption did not induce widespread activation of foregut genes or a homeotic morphological shift, and gene expression differed substantially according to the developmental context of *Cdx2* inactivation (Fig. 1). Thus, although Cdx2 dominantly controls intestinal differentiation in both embryos and adults, its target genes in the two contexts show surprisingly little overlap, with marked suppression of anterior foregut genes in embryos and direct control over a large portion of the enterocyte transcriptional program in adults. Nevertheless, expression of some foregut genes increased by several orders of magnitude over a nearly silent baseline in adults (Fig. 6), indicating that caudal proteins continually suppress a modicum of residual developmental plasticity in the adult gut. Because our analysis of genome occupancy suggests a primary activating function, it follows that silencing of anterior genes likely occurs indirectly through one or more intermediary factors. The distinct effects of *Cdx2* gene inactivation at different stages probably reflect significantly different interactions of a dominant TF with diverse chromatin features and coregulatory factors on the embryonic and adult genomes. For example, our ChIP analysis in adult intestinal villus cells showed Cdx2 binding within 20 kb of the TSS at only 5 of 43 endolysosomal genes and at 12 of 40 lysosomal genes that are implicated in the cell polarity defects observed in the Cdx2-deficient fetal intestine (data not shown).

Combined loss of Cdx1 and Cdx2 in the intestine reduced expression of more than 3,000 transcripts and increased expression of an even larger number. In this light, it is difficult to attribute the cellular phenotypes we highlight in this report to any single gene or group of genes. Rather, the complex deficits in cell morphology and vacuolation, paucity of the microvillus brush border, and aberrant expression and distribution of polarity proteins likely result from the simultaneous dysregulation of hundreds of genes of diverse function. In the EE cell lineage, however, where cell numbers are appreciably reduced in Cdx1 Cdx2 DKO intestines, the data do allow a stricter interpretation. *Math1* mRNA levels are unaffected, whereas TFs that operate lower in the EE transcriptional hierarchy (28) and hormonal end products are reduced (Fig. 5). These results suggest that Cdx function in this lineage might occur downstream or in parallel to *Math1*, impinging on control of *Ngn3* and *NeuroD* levels.

The functions of few other dominant lineage-determining genes have been investigated in the same detail in early and late embryos and in adults. MyoD and Myf5 are essential for skeletal muscle development, but their inactivation in adult mice has few consequences (36). In contrast, *Gata1* is important in both embryonic and adult erythropoiesis (14). One possibility is that regenerative adult tissues like the intestine and blood depend more on the function of dominant regulators than postmitotic tissues such as muscle. Resolution of this question can occur with examination of other TFs that drive fate choice in embryos and remain expressed in the same sites in adults. However, although Cdx mutant mouse phenotypes define temporal genetic requirements, they do not alone distinguish between direct and secondary targets or illuminate the transcriptional basis of Cdx functions. Our synthesis of mouse genetic and cistrome analyses overcomes the inherent limita-

tions of each approach, permits a deeper appreciation of transcriptional mechanisms in the intestine, and points the way for integration of experimental approaches to elucidate the specific activities of Cdx in the embryo or, more generally, of key TFs in other tissues.

#### ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grants R01DK082889 (R.A.S.), R01HG4069 (X.S.L.), and 1K01DK088868-01 (M.P.V.), a fellowship from the Crohn's and Colitis Foundation (grant 1987 to M.P.V.), and core resources provided under the aegis of center grant P50CA127003.

We thank John Lynch for Cdx1 antiserum, Andre Ouellette for Crs4c antibody, Dan Podolsky for TFF3 antiserum, Peter Gruss for sharing *Cdx1*<sup>-/-</sup> mice and Sylvie Robine for sharing *Villin-Cre*<sup>ER(T2)</sup> transgenic mice.

#### REFERENCES

- Barker, N., M. van de Wetering, and H. Clevers. 2008. The intestinal stem cell. *Genes Dev.* **22**:1856–1864.
- Berger, M. F., et al. 2008. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133**:1266–1276.
- Bonhomme, C., et al. 2008. Cdx1, a dispensable homeobox gene for gut development with limited effect in intestinal cancer. *Oncogene* **27**:4497–4502.
- Chawengsaksophak, K., W. de Graaff, J. Rossant, J. Deschamps, and F. Beck. 2004. Cdx2 is essential for axial elongation in mouse development. *Proc. Natl. Acad. Sci. U. S. A.* **101**:7641–7645.
- Chawengsaksophak, K., R. James, V. E. Hammond, F. Kontgen, and F. Beck. 1997. Homeosis and intestinal tumours in Cdx2 mutant mice. *Nature* **386**:84–87.
- Crissey, M. A., et al. 2008. The homeodomain transcription factor Cdx1 does not behave as an oncogene in normal mouse intestine. *Neoplasia* **10**:8–19.
- Crissey, M. A., et al. 2011. Cdx2 levels modulate intestinal epithelium maturity and paneth cell development. *Gastroenterology* **140**:517–528 e8.
- Davidson, A. J., and L. I. Zon. 2006. The caudal-related homeobox genes *cdx1a* and *cdx4* act redundantly to regulate *hox* gene expression and the formation of putative hematopoietic stem cells during zebrafish embryogenesis. *Dev. Biol.* **292**:506–518.
- el Marjou, F., et al. 2004. Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. *Genesis* **39**:186–193.
- Gao, N., and K. H. Kaestner. 2010. Cdx2 regulates endo-lysosomal function and epithelial cell polarity. *Genes Dev.* **24**:1295–1305.
- Gao, N., P. White, and K. H. Kaestner. 2009. Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2. *Dev. Cell* **16**:588–599.
- Grainger, S., J. G. Savory, and D. Lohnes. 2010. Cdx2 regulates patterning of the intestinal epithelium. *Dev. Biol.* **339**:155–165.
- Guo, R. J., E. R. Suh, and J. P. Lynch. 2004. The role of Cdx proteins in intestinal development and cancer. *Cancer Biol. Ther.* **3**:593–601.
- Gutierrez, L., et al. 2008. Ablation of *Gata1* in adult mice results in aplastic crisis, revealing its essential role in steady-state and stress erythropoiesis. *Blood* **111**:4375–4385.
- He, H. H., et al. 2010. Nucleosome dynamics define transcriptional enhancers. *Nat. Genet.* **42**:343–347.
- Hinoi, T., et al. 2002. CDX2 regulates liver intestine-cadherin expression in normal and malignant colon epithelium and intestinal metaplasia. *Gastroenterology* **123**:1565–1577.
- Irizarry, R. A., et al. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**:249–264.
- James, R., T. Erler, and J. Kazenwadel. 1994. Structure of the murine homeobox gene *cdx-2*. Expression in embryonic and adult intestinal epithelium. *J. Biol. Chem.* **269**:15229–15237.
- Latchman, D. S. 2001. Transcription factors: bound to activate or repress. *Trends Biochem. Sci.* **26**:211–213.
- Liu, T., et al. 2007. Regulation of Cdx2 expression by promoter methylation, and effects of Cdx2 transfection on morphology and gene expression of human esophageal epithelial cells. *Carcinogenesis* **28**:488–496.
- Madison, B. B., et al. 2002. Cis elements of the villin gene control expression in restricted domains of the vertical (crypt) and horizontal (duodenum, cecum) axes of the intestine. *J. Biol. Chem.* **277**:33275–33283.
- Mutoh, H., et al. 2002. Conversion of gastric mucosa to intestinal metaplasia in Cdx2-expressing transgenic mice. *Biochem. Biophys. Res. Commun.* **294**:470–479.
- Mutoh, H., et al. 2004. Cdx1 induced intestinal metaplasia in the transgenic mouse stomach: comparative study with Cdx2 transgenic mice. *Gut* **53**:1416–1423.

24. Niwa, H., et al. 2005. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123**:917–929.
25. Rings, E. H., et al. 2001. Phosphorylation of the serine 60 residue within the Cdx2 activation domain mediates its transactivation capacity. *Gastroenterology* **121**:1437–1450.
26. Rosenfeld, M. G., and C. K. Glass. 2001. Coregulator codes of transcriptional regulation by nuclear receptors. *J. Biol. Chem.* **276**:36865–36868.
27. Savory, J. G., et al. 2009. Cdx1 and Cdx2 are functionally equivalent in vertebral patterning. *Dev. Biol.* **330**:114–122.
28. Schonhoff, S. E., M. Giel-Moloney, and A. B. Leiter. 2004. Minireview: development and differentiation of gut endocrine cells. *Endocrinology* **145**:2639–2644.
29. Scully, K. M., et al. 2000. Allosteric effects of Pit-1 DNA sites on long-term repression in cell type specification. *Science* **290**:1127–1131.
30. Shin, H., T. Liu, A. K. Manrai, and X. Liu. 2009. CEAS: cis-regulatory element annotation system. *Bioinformatics* **25**:2605–2606.
31. Shroyer, N. F., D. Wallis, K. J. Venken, H. J. Bellen, and H. Y. Zoghbi. 2005. Gfi1 functions downstream of Math1 to control intestinal secretory cell subtype allocation and differentiation. *Genes Dev.* **19**:2412–2417.
32. Silberg, D. G., et al. 1997. CDX1 protein expression in normal, metaplastic, and neoplastic human alimentary tract epithelium. *Gastroenterology* **113**:478–486.
33. Silberg, D. G., et al. 2002. Cdx2 ectopic expression induces gastric intestinal metaplasia in transgenic mice. *Gastroenterology* **122**:689–696.
34. Silberg, D. G., G. P. Swain, E. R. Suh, and P. G. Traber. 2000. Cdx1 and Cdx2 expression during intestinal development. *Gastroenterology* **119**:961–971.
35. Smyth, G. K. 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**:Article 3.
36. Ustanina, S., J. Carvajal, P. Rigby, and T. Braun. 2007. The myogenic factor Myf5 supports efficient skeletal muscle regeneration by enabling transient myoblast amplification. *Stem Cells* **25**:2006–2016.
37. van den Akker, E., et al. 2002. Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development* **129**:2181–2193.
38. van Nes, J., et al. 2006. The Cdx4 mutation affects axial development and reveals an essential role of Cdx genes in the ontogenesis of the placental labyrinth in mice. *Development* **133**:419–428.
39. Verzi, M. P., et al. 2010. Differentiation-specific histone modifications reveal dynamic chromatin interactions and alternative partners for the intestinal transcription factor CDX2. *Dev. Cell* **19**:713–726.
40. Wang, Y., et al. 2008. Cdx gene deficiency compromises embryonic hematopoiesis in the mouse. *Proc. Natl. Acad. Sci. U. S. A.* **105**:7756–7761.
41. Yamamichi, N., et al. 2009. Cdx2 and the Brm-type SWI/SNF complex cooperatively regulate villin expression in gastrointestinal cells. *Exp. Cell Res.* **315**:1779–1789.
42. Young, T., et al. 2009. Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev. Cell* **17**:516–526.
43. Yu, M., et al. 2009. Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Mol. Cell* **36**:682–695.
44. Zhang, Y., et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**:R137. <http://genomebiology.com/content/9/9/R137>.