Enhancer, transcriptional, and cell fate plasticity precedes intestinal determination during endoderm development

Kushal K. Banerjee,1,2,7 Madhurima Saxena,1,2,7 Namit Kumar,3,7 Lei Chen,3 Alessia Cavazza,1,2 Natalie H. Toke,3 Nicholas K. O’Neill,1 Shariq Madha,1 Unmesh Jadhav,1,2 Michael P. Verzi,3,4,5 and Ramesh A. Shivdasani1,2,6

1Department of Medical Oncology, Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA; 2Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts 02215, USA; 3Department of Genetics, Rutgers, the State University of New Jersey, Piscataway, New Jersey 08854, USA; 4Cancer Institute of New Jersey, Piscataway, New Jersey 08854, USA; 5Human Genetics Institute of New Jersey, Piscataway, New Jersey 08854, USA; 6Harvard Stem Cell Institute, Cambridge, Massachusetts 02139, USA

After acquiring competence for selected cell fates, embryonic primordia may remain plastic for variable periods before tissue identity is irrevocably determined (commitment). We investigated the chromatin basis for these developmental milestones in mouse endoderm, a tissue with recognizable rostro–caudal patterning and transcription factor (TF)-dependent interim plasticity. Foregut-specific enhancers are as accessible and active in early midgut as in foregut endoderm, and intestinal enhancers and identity are established only after ectopic cis-regulatory elements are decommissioned. Depletion of the intestinal TF CDX2 before this cis element transition stabilizes foregut enhancers, reinforces ectopic transcriptional programs, and hence imposes foregut identities on the midgut. Later in development, as the window of chromatin plasticity elapses, CDX2 depletion weakens intestinal, without strengthening foregut, enhancers. Thus, midgut endoderm is primed for heterologous cell fates, and TFs act on a background of shifting chromatin access to determine intestinal at the expense of foregut identity. Similar principles likely govern other fate commitments.

[Keywords: developmental competence; fate determination; tissue specification; lineage commitment; developmental plasticity; chromatin plasticity; homeodomain transcription factors]

Supplemental material is available for this article.

Received July 18, 2018; revised version accepted September 27, 2018.

After embryonic cells acquire the competence to develop into a limited number of tissues, adult cell fates are determined in phases [Slack 1991]. Possible cell identities are first specified, and developmental potential may remain flexible for some period until cells commit to unilineage differentiation. Each of these steps has a basis in transcription factor (TF) interactions with specific cis-regulatory elements [Shlyueva et al. 2014; Long et al. 2016]. Enhancers recruited early in tissue ontogeny confer competence [Nord et al. 2013; Lara-Astiaso et al. 2014]; cell identities emerge as organ primordia activate these and subsequently additional enhancers [Stergachis et al. 2013; Wang et al. 2015]. For example, when human embryonic stem cells (ESCs) differentiate into pancreatic, liver, or lung lineages, the active histone mark H3K4me1 is found on a large swath of enhancers, signifying broad tissue competence [Wang et al. 2015]. Later, inductive signals drive different cell populations to stabilize a fraction of this early enhancer pool and thereby commit to a single fate.

We investigated the logic underlying enhancer recruitment and dynamic TF interactions with chromatin during development of the mouse alimentary canal, where distinctive epithelia with diversemorphologies and functions arise from undifferentiated gut endoderm. Because cell migration along the rostro–caudal axis is limited, the mucosa of each gut segment descends largely and directly from the corresponding regional endoderm.

© 2018 Banerjee et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genesdev.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License [Attribution-NonCommercial 4.0 International], as described at http://creativecommons.org/licenses/by-nc/4.0/.
The molecular mechanisms are pertinent to human health because gastrointestinal atresias and malformations are among the commonest congenital disorders and because the largest risk for stomach and lower esophageal cancers occurs after native adult epithelia acquire heterologous [intestinal] character.

The pseudostratified epithelium lining the primitive mouse gut tube responds to regional cues to establish foregut (FG), midgut, and hindgut domains [Zorn and Wells 2009, Spence et al. 2011]. Windows of developmental plasticity are evident within these domains before cells adopt definitive squamous [esophagus [Eso] and forestomach [FS]] or columnar epithelial [glandular stomach and intestine] identities [Wells and Melton 1999, Grapin-Botton and Melton 2000]. Some TFs are expressed broadly before approximately embryonic day 9 (E9)—when Wnt, bone morphogenetic protein [BMP], and other signals help establish an anterior–posterior pattern—and are later restricted to specific regions [Wells and Melton 2000]. Moreover, endoderm transplanted before E13 adopts the

Figure 1.  **Transcriptional and chromatin landscapes in the developing and neonatal gut.** (A) Gut epithelia with distinctive morphologies arise from contiguous endodermal regions. (Eso) Esophagus; (FS) forestomach; (HS) hindstomach; (Int) small intestine. (B) Principal component analysis [PCA] of Eso/FS (red), HS (purple), and Int (blue) transcriptomes at embryonic day 12 (E12), E14, and E16. n = 2 per group. The largest sources of variation reflect the emergence of tissue-specific mRNA profiles after E14. Photomicrographs of intestinal endoderm show pseudostratified epithelium at E12, early villus formation at E14, and mature villus structures at E16. (C) Differential chromatin access in postnatal day 1 (P1) endoderm from different regions of the tubal gut. Each cluster of ATAC-seq [assay for transposase-accessible chromatin [ATAC] using sequencing] summits ±1.5 kb represents candidate enhancers selectively accessible in one region or common to the FS and HS (Overlap). In adult intestinal villi, the enhancers identified in P1 intestine carry H3K4me1, H3K27ac, and RNAPII, but these marks are largely missing from enhancers detected in P1 foregut tissues. Integrated Genome Viewer (IGV) tracks show ATAC-seq and RNA sequencing [RNA-seq] data at representative loci from Eso/FS and Int clusters. Numbers represent signal scales, and shaded boxes highlight region-specific enhancers. (D) Gene set enrichment analysis [GSEA] of genes located within 50 kb of sites selectively accessible in the Eso/FS or Int (derived using binding and expression target analysis [BETA]) [Wang et al. 2013], plotted with respect to differential mRNA expression in the E16 Eso/FS and Int. (NES) Normalized enrichment score.
regional identity of the host segment, whereas transplants of more advanced endoderm retain the identity of the donor region [Mizuno and Yasugi 1990; Duluc et al. 1994; Yasugi and Mizuno 2008]. The chromatin and TF bases for initially labile midgut cell fate and subsequent commitment to intestinal differentiation are unclear.

In mid-gestation and beyond, the TF CDX2 is restricted to intestine-specific endoderm [Beck et al. 1995; Silberg et al. 2000] and adult intestinal epithelium [James et al. 1994]. Forced expression in the developing mouse stomach [Mutoh et al. 2002; Silberg et al. 2002], but not in adult esophageal cells [Kong et al. 2011], induces enteric differentiation. Although germline CDX2 loss is lethal, heterozygote animals show focal stomach differentiation in the hindgut (Beck et al. 1999). Conditional loss at approxi-mately E9 causes colonic atresia and replacement of the villous ileal mucosa by a stratified epithelium with Eso hindgut (Beck et al. 1999). Although germline CDX2 deletion at E13 resulted in glandular stomach-like morphology and expression of gastric genes in the duodenum [Graigner et al. 2010], whereas deletion in adult intestines induced weak expression of few stomach genes [Verzi et al. 2010, 2011; Stringer et al. 2012] but precipitated lethal intestinal failure owing to collapse of CDX2-dependent enhancers [Verzi et al. 2013; Saxena et al. 2017]. These studies implicate CDX2 in the highly contextual control of intestinal development and function. We postulated that investigation of CDX2-chromatin interactions during mouse development might illuminate the underpinnings of tissue competence, specification, and determination.

Results

Region-specific gene expression in the developing mouse gut is associated with distinct profiles of open enhancer chromatin

The gut endoderm generates a squamous lining in the Eso and FS and distinctive columnar epithelia in the hind-stomach [HS] and intestine [Zorn and Wells 2009]. To study transcriptional and chromatin dynamics that underlie this rostro-caudal patterning, we purified EPCAM+ endodermal cells [Supplemental Fig. S1A; Sherwood et al. 2007] from discrete regions of the E12, E14, and E16 mouse gut [Fig. 1A]: (1) the prospective FS and Eso, (2) the area between the FS and gastric pylorus [HS], and (3) the tube distal to the pylorus and proximal to the Cecum [midgut or small intestine [Int]]. RNA sequencing [RNA-seq] data from replicate samples [Supplemental Table S1] were highly concordant, and regional markers attested to the purity of cell isolates [Supplemental Fig. S1B]. In principal component analysis [PCA] [Fig. 1B] and correlation analysis [Supplemental Fig. S1C], temporal changes accounted for the largest variation in gene expression, with mRNA profiles diverging by region after E12, these findings agree with observations that the intestinal lining is undifferentiated at E12 until villus primordia first appear at approximately E14 [Walton et al. 2012] and mature thereafter [Fig. 1B]. K-means clustering of genes differentially expressed at E16 \( q < 0.05, \) fold change \( \geq 4 \), reads per kilobase per million mapped reads [RPKM] \( >1 \) yielded groups with expression restricted to the Eso/FS, the HS, the Int, or two of these prospective epithelia [Supplemental Fig. S1D]. These region-specific genes represent the purpose of digestive tract patterning and reflect the outcomes of spatio-temporal chromatin organization.

To determine the corresponding chromatin states, we first used the assay for transposase-accessible chromatin [ATAC] with sequencing [ATAC-seq] [Buenrostro et al. 2015] on Eso/FS, HS, and Int epithelia at postnatal day 1 [P1] [Supplemental Table S2]. Replicate samples were highly concordant, regional differences in open chromatin were readily evident [Supplemental Fig. S2A,B], and different regions with expression restricted to the Eso/FS, the HS, or Int were identified genomic sites where chromatin access differed by region [Supplemental Fig. S2C]. At sites located \( >2 \) kb away from transcription start sites [TSSs], we thus detected candidate enhancers unique to each organ and sites shared among two or all three tissues [Fig. 1C]. Areas selectively accessible in P1 intestine showed active histone marks and RNA polymerase II [Pol II] binding only in the adult intestine [Fig. 1C; Supplemental Fig. S2D], attesting that they are region-specific cis elements, and GREAT [Genomic Regions Enrichment of Annotations Tool] analysis [McLean et al. 2010] verified that genes \( <50 \) kb from differential ATAC sites serve region-specific roles [Supplemental Fig. S2E]. Tissue-restricted chromatin access of some cis-regulatory domains was readily apparent at known tissue-specific loci [Fig. 1C; Supplemental Fig. S2F,G], and binding and expression target analysis [BETA] [Wang et al. 2013] revealed association of the Eso/FS and Int site clusters with genes highly enriched in the E16 Eso/FS and E16 Int, respectively [Fig. 1D]. Because enhancers can act at distances \( >50 \) kb and on loci other than the nearest gene [Shlyueva et al. 2014; Long et al. 2016], such associations speak to likely activities of ATAC+ sites in aggregate and not the established function of any given site. Nevertheless, for simplicity, we refer to this collection of putative regulatory sites as “enhancers.”

Open chromatin dynamics identify a likely basis for cell fate plasticity in the developing intestine

To study trajectories of chromatin access that culminate in the neonatal tissues, we used ATAC-seq to map open chromatin in EPCAM+ cells purified from each regional endoderm at E12, E14, and E16. Figure 2A displays the normalized data from these series [quantified in Supplemental Fig. S3A,B] for every region-specific cluster. Although each P1 profile was strongest and restricted in the corresponding E16 epithelium, ATAC signatures in the FG were apparent in the respective regions by E14, whereas intestine-specific sites opened later [Fig. 2A]. This differential trajectory is compatible with rostral-to-caudal progression of digestive tract differentiation, as judged by morphology in diverse vertebrate species [Wells and Mleton 1999; Grapin-Botton 2005]. Notably, FG-restricted profiles were evident at equal measure in all regions, including the Int, at E12 and E14; some FG sites even showed access in Int cells until E16. In contrast, a global view of Int
enhancers revealed a lack of open chromatin at any time in FG-derived endoderm [Fig. 2A]. These findings were evident in independent replicates of Int endoderm and do not reflect tissue contamination because the pylorus is an unmistakable landmark at the FG–midgut boundary. Nevertheless, we also examined Int endoderm at E11, the earliest time that the Int and HS are visibly distinct and EPCAM is restricted to the endoderm. FG enhancers were as open in E11 Int endoderm as at E12 (Fig. 2A). As open enhancer chromatin reflects the potential to express nearby genes, these findings imply that before developing the ability to express Int genes, midgut endoderm may...
harbor the potential to activate FG-specific genes and, by extension, for Eso/FS or HS differentiation. In contrast, FG endoderm appears to lack the reciprocal capacity for intestinal differentiation, at least by E12.

To ascertain whether open chromatin in Int bears other enhancer features, we analyzed data on the active histone mark H3K27ac in E12 and E14 EPCAM+ Int endoderm and adult epithelium (Kazakevych et al. 2017). Int-restricted enhancers carried H3K27ac in E14 and adult intestinal cells, as expected, so did many FG-specific sites in E12 and E14 intestinal endoderm (Fig. 2B). Both open chromatin and H3K27ac at FG enhancers were abrogated in P1 or adult Int cells (Fig. 2A,B), and to impute their cis-regulatory activity, we examined RNA levels of region-specific genes [Supplemental Fig. S1D], which the corresponding regional enhancers control (Fig. 1D). Median mRNA levels of FG genes were fourfold to >100-fold higher than those of Int genes in E16 FG endoderm but were comparable in E12 cells [Fig. 2C; Supplemental Fig. S3C]. Eso/FS-specific Sox15 (Sulahian et al. 2015), for example, shows selectively high mRNA in the rostral FG by E16 but even higher levels in Int than in Eso/FS at E12 [Fig. 2C].

Like many such loci, Sox15 and HS-restricted Gkn3 are open and marked with H3K27ac in E12 Int endoderm and lose these features over time, even as the same sites gain enhancer features in the Eso/FS (Fig. 2D). These observations together reveal that thousands of FG enhancers are measurably active in undifferentiated Int endoderm and later decommissioned, concomitant with activation of Int-specific enhancers and determination of Int cell fate (Fig. 2E).

**The timing of CDX2 loss strongly influences homeotic intestinal transformation**

Enhancer states represent the activities of sequence-specific TFs and recruited coregulators (Long et al. 2016). Indeed, motifs for known tissue-restricted TFs were the most enriched among the respective P1 tissue-specific enhancers [Fig. 3A]. These TF transcripts were present at low levels in each E12 endoderm, over time, their levels increased in the cognate endoderm and declined reciprocally in heterologous epithelia [Fig. 3B]. Cdx2, however, stands out because it is restricted to the intestine, as reported previously (Beck et al. 1995; Silberg et al. 2000), and peak levels were achieved by E12, much before regional endoderm specification. Conditional loss of Cdx2 at different developmental stages has varied consequences (Gao et al. 2009). To determine its tissue requirements in relation to chromatin and TF dynamics, we deleted Cdx2 using ShhCre, which acts in the luminal gut endoderm by approximately E10 (Echelard et al. 1993). Jejunal villi were stunted and, in lieu of crypt-villus units, the ileum had a stratified lining of the Eso/FS type [Fig. 3C, red arrows]. Jejunal glands expressed the gastric parietal cell marker ATP4B at the base [Fig. 3C, white arrowheads], and, in place of scattered periodic acid Schiff (PAS)-stained goblet cells, duodenal and jejunal villi showed numerous gastric foveolar cells [recognized by a mauve apical rim; black arrows] extensively replace intestinal goblet cells [recognized by a purple cytoplasmic blob; black arrowheads]. These contrasting features of wild-type and mutant jejunal villi are highlighted in the dashed boxes and magnified at the right. In addition, ATP4B+ gastric parietal cells appear in the mutant jejunal [white arrowheads] but not duodenal crypts. The mutant ileum loses crypt–villus structure, including PAS+ goblet cells, and is lined instead by TRP63+ stratified squamous mucosa [red arrows]. Bars, 25 µm.

**Figure 3.** ShhCre-mediated Cdx2 deletion results in anterior homeotic transformation. [(A)] The highest enriched sequences within region-specific enhancer clusters represent canonical TF motifs. [(B)] Trajectories of Eso/FS- and Int-specific TF transcripts in E12, E14, and E16 Eso/FS [left] and Int [right] endoderm. Log2 transformed RPKM values, plotted over time, show increases and decreases in the different regions. [(C)] Intestinal regions (from the duodenum to the ileum) stained with PAS or with ATP4B or TRP63 antibody in E18 wild-type and ShhCre;Cdx2CreERT2 embryos. In the mutant duodenum and jejenum, gastric foveolar cells [recognized by a mauve apical rim; black arrows] extensively replace intestinal goblet cells [recognized by a purple cytoplasmic blob; black arrowheads]. These contrasting features of wild-type and mutant jejunal villi are highlighted in the dashed boxes and magnified at the right. In addition, ATP4B+ gastric parietal cells appear in the mutant jejunal [white arrowheads] but not duodenal crypts. The mutant ileum loses crypt–villus structure, including PAS+ goblet cells, and is lined instead by TRP63+ stratified squamous mucosa [red arrows]. Bars, 25 µm.
foveolar cells, where PAS classically stains the cellular apex (Fig. 3C, black arrowheads and arrows, respectively). Notably, these ectopic regional features were not intermingled: Squamous marker TRP63 was expressed in the Cdx2−/− ileum but not the duodenum or jejunum, and, conversely, ATP4B was present in the jejunum but not the ileum (Fig. 3C). Thus, ShhCremediated Cdx2 loss results in regionally confined alterations of endodermal fate determination, distinct from those reported when Cdx2 is deleted earlier using Foxa3Cremediated Cdx2 protein (Fig. 4A) were depleted. ShhCremediated loss of Cdx2 yielded mRNA profiles that were the most distinct from wild-type E16 intestine and most similar to FG derivatives (Fig. 4B). Considering region-specific RNA clusters (Supplemental Fig. S1D), early Cdx2 loss markedly attenuated Int-specific genes and produced substantial overlap with FG-restricted transcripts (Fig. 4C; Supplemental Fig. S4B). In contrast, deletion at E13 decreased Int transcripts modestly, and fewer FG genes were concomitantly activated, whereas deletion at E15 resulted in virtually no FG-specific gene expression (Fig. 4C), similar to adult Cdx2−/− jejunum (Verzi et al. 2013). Moreover, after CRE activation

Figure 4. Time-sensitive activation of FG-specific transcriptional programs in the absence of Cdx2. (A) Schema of Cdx2 deletions mediated by either ShhCremediated Cdx2 deletion at approximately E10 or treatment of Villin-CreER-T2 dams with tamoxifen at E13 or E15. Embryos deleted before E14 were assessed at E16, and those deleted at E15 were assessed at E18. Cdx2 immunostains verify protein loss. (B) Unsupervised hierarchical clustering (Pearson coefficients) for the 2604 genes expressed differently (greater than fourfold, q < 0.05) in E16 wild-type tissues (Supplemental Fig. S1D) and E16 Cdx2−/− Int after deletion at approximately E10 or E13. Transcriptomes in early, but not late, Cdx2-deleted Int endoderm resemble those in FG epithelia. (C) mRNA differences in Cdx2−/− Int after deletion at approximately E10 (n = 3), E13, or E15 in adult duodenum and ileum (n = 2 each). Clusters are those with differential expression in wild-type E16 EPCAM+ Eso/FS, HS, and Int cells (from Supplemental Fig. S1D). For each FG cluster, colored violin plots represent the distributions of mRNA levels (colors represent the region, and white dots represent the median) in wild-type tissues, and gray violins show the distributions after Cdx2 deletion (knockout) at different times. Violin plots for Int genes are shown in Supplemental Figure S3B. The Mann-Whitney U-test was applied to determine the significance between Cdx2 knockout at different times. (n.s.) Not significant; [*] P < 0.05; [**] P < 0.0001. (D–F) IGV tracks of RNA-seq data showing native (wild-type FS, HS, and Int) and ectopic FG gene expression in Cdx2−/− intestines after deletion at E10, E13, or E15. Numbers represent the scale for normalized read counts.
in adult mice, duodenal and ileal Cdx2−/− villus cells did not show significant levels of FG transcripts (Fig. 4C). Thus, CDX2 depletion markedly affects both FG and intestinal gene activity, with the largest effect on both programs occurring after early loss of CDX2 and progressively less disruption upon later losses.

Several features of FG gene activity in developing Cdx2−/− intestines merit note. First, at any time of CDX2 loss, not all FG-specific genes were activated (Supplemental Fig. S4C). Expression of canonical Eso/F5 and HS genes explains the FG morphology when Cdx2 was deleted by E10 (Figs. 3C, 4D; Supplemental Fig. S4D), but other wild-type FG transcripts were never detected in the Cdx2−/− midgut (Fig. 4E). Second, FG-specific mRNA levels in mutant intestines approached and often exceeded the native levels in the wild-type Eso or stomach (e.g., see Fig. 4D,F; Supplemental Fig. S4D), indicating robust ectopic gene activity. Third, transcripts expressed in both the wild-type Eso and stomach ("overlapping") accounted for a substantial fraction of aberrant gene activity (Fig. 4C,D, Supplemental Fig. S4C), in line with the abundance of open chromatin at corresponding loci in the wild-type E11–E14 midgut (Fig. 2A). Fourth, whereas Cdx2 deletion at E13 activated fewer FG genes and with less strength than deletion by ShhCre, some genes (especially overlapping and HS) were activated only after E13 deletion (Fig. 4F; Supplemental Figs. S4C,E). Finally, although global mRNA profiles after Cdx2 deletion at E15 suggested almost no FG gene activity, ~20% of genes were activated greater than fourfold (Supplemental Fig. S4F).

These findings collectively demonstrate that midgut endoderm is initially biased toward FG differentiation and that CDX2 helps override this latent potential. Although midgut competence subsequently narrows, deletion of Cdx2 as late as E13 increases intestinal mRNA levels of 53% of gastric genes by greater than twofold and 43% of gastric genes by greater than fourfold. By E15, however, the window for FG gene activity is closed and, except at a few residual enhancers, no longer requires CDX2 to prevent it. Of note, mRNA (Fig. 1B) and open chromatin (Fig. 2A) trajectories in wild-type embryos reveal a crucial transition at approximately E14, and Cdx2 deletion through E13 reflects the same window of time, when endoderm is yet undetermined and FG genes can be stably activated in the midgut.

Cdx2 loss at different times produces distinct tissue defects and open chromatin profiles

Foxa3Cre-driven loss did not induce gastric morphology or ectopic stomach transcripts (Gao et al. 2009), whereas deletion at E13 induced HS cells only in part of the proximal duodenum and scant expression of gastric genes (Grainger et al. 2010). ShhCre-mediated Cdx2 deletion likely occurs between these times and, in addition to ileal squamous differentiation, induces overt gastric features in the proximal and mid-intestine and robustly activates FG genes. The lower levels of FG-restricted transcripts when Cdx2 was deleted at E13 or E15 predicted that anteriorizing homeotic transformation may be reduced or absent at the later times. Indeed, following Cdx2 deletion at E13 or E15, PAS staining throughout the intestine showed features of intestinal goblet—not gastric foveolar—cells, and the ileum did not carry a stratified epithelium but a typical columnar villous type (Fig. 5A). Thus, strikingly different phenotypes ensue when Cdx2 is deleted at different times (Supplemental Fig. S4G), mirroring the inherent early accessibility and subsequent closure of FG enhancers.

The Cdx2 sequence motif is enriched only at intestinespecific and not at FG-specific enhancers (Fig. 3A). Nevertheless, CDX2’s potent role in preventing FG gene activity could reflect repressive action from binding at FG enhancers or an indirect effect exerted by other means. To distinguish among these possibilities, we used CDX2 antibody for ChIP [chromatin immunoprecipitation] on EPCAM* wild-type midgut endodermal cells, where it appeared at Int-specific enhancers by E14, and binding strengthened over time. In contrast, FG-specific enhancers globally showed negligible binding at E14 or E13, the earliest times that ChIP was feasible (Fig. 5B; Supplemental Fig. S5A), and inspection of individual FG sites failed to reveal appreciable CDX2 signals (Fig. 5C). This absence did not reflect ChIP insensitivity because binding at E13 and E14 was evident at other H3K27ac-marked enhancers with open chromatin (Supplemental Fig. S5B). Rather, the minimal overlap of E13 binding sites with regions of open FG-specific chromatin (Fig. 5D) implies that CDX2 precludes FG gene activation indirectly. In contrast, the sites and chronology of CDX2 binding at E14 and E17 (Fig. 5B) correlate with, and may underlie, intestine-specific gene activity.

To ascertain how FG- and Int-specific areas of open chromatin respond to loss of Cdx2, we deleted Cdx2 at various times and assessed Int endoderm by ATAC-seq. As with mRNA profiling, when Cdx2 deletion occurred at E10 or E13, we studied EPCAM+ cells at E16, and, when Cdx2 deletion occurred at E15, we examined epithelial cells 3 d later. We also examined adult Cdx2−/− epithelium that we harvested from the ileum, the region of highest native expression (Silberg et al. 2000) and the most severely affected after embryonic (Fig. 3C) or adult (Verzi et al. 2010) Cdx2 loss. ShhCre-driven loss of Cdx2 and VillinCre-mediated loss at E13 caused a marked persistence of open chromatin at FG-specific enhancers, especially at "overlapping" and HS sites (Fig. 6A; Supplemental Fig. S5C,D). Among genes up-regulated after E10 deletion (Fig. 4C), 31% of TSSs lie <25 kb and 46% lie <50 kb from these areas of aberrantly open chromatin (Fig. 6B), and BETA analysis (Wang et al. 2013) confirmed strong statistical association (Supplemental Fig. S5E). Also in response to early Cdx2 deletion, intestinal enhancers lacked open chromatin, indicating failure of intestinal gene activation. In contrast, deletion at E15 or in adults diminished intestinal ATAC signals to a lesser degree and failed to hold chromatin open at FG-specific enhancers (Fig. 6A; Supplemental Fig. S5C). These data identify a basis in chromatin for TF-dependent fate determination across thousands of region-specific cis elements (Fig. 6C). They also demonstrate Cdx2’s dual role in de-commissioning FG enhancers without direct binding
and in establishing intestine-specific enhancers where CDX2 occupancy is pervasive.

Discussion

This study of mouse endoderm across developmental space and time reveals substantial chromatin, transcriptional, and cell fate plasticity foreshadowing the commitment to a unique epithelial identity. We report two core findings. First, before the midgut endoderm activates an intestine-specific transcriptional program and adopts villous morphology, enhancers that control FG-specific genes are broadly accessible to the same degree as in FG endoderm. At least after E11, the reverse is not true: Intestinal enhancers were inaccessible in multiple EPCAM+ cell isolates from the FS (future squamous) or HS (future glandular stomach). Beyond having accessible chromatin and the active histone mark H3K27ac, FG-specific genes near these cis elements are expressed in early midgut endoderm to levels at least the same as—and often higher than—the same transcripts in early FG endoderm. Together, these findings indicate that early midgut endoderm enhancers carry the potential to activate FG genes, and the tissue can therefore differentiate into stratified or glandular gastric epithelia. Thus, tissue determination reflects the imposition of intestinal—at the expense of FG—identity. Our observations mirror those showing broad priming of enhancers in human ESC-derived endoderm before differentiation of the pancreatic lineage [Wang et al. 2015]. It is unclear why many FG genes express at even higher levels in early midgut than in FG endoderm [Fig. 2C]. At a minimum, our findings imply that crucial FG TFs are present in midgut endoderm and can engage accessible FG enhancers to drive transcription of FG genes.

A second core finding reveals the finite window in which CDX2 acts to suppress alternative tissue identity. Depletion of CDX2 at approximately E10 induces Eso/FS differentiation in the distal midgut, glandular HS differentiation in the proximal midgut, and failure to activate Int-specific genes. By E15 and in adult mice, acute CDX2 loss diminishes Int-specific transcripts but barely affects FG genes, whereas loss at E13 has an intermediate effect: activation of many FG-specific genes without induction of overt FG histology. Importantly, the window in which CDX2 is required to set chromatin marking and accessibility patterns (thus overcoming FG gene activity)—before but not after E14—coincides with that of initial epithelial divergence [Fig. 1B] and FG enhancer decommissioning in the wild-type midgut [Fig. 2A]. Thus, CDX2 loss before E14 stabilizes intrinsically accessible FG enhancers, which are subsequently unavailable to stabilize. These findings reveal the gene regulatory logic of rostro-caudal gut tube patterning and a chromatin basis for determination of one epithelial fate at the expense of others. Of note, our study examined only antero-posterior patterning of the luminal gut; development of other endoderm derivatives—the liver, pancreas, and lungs—may use the same or different regulatory logic.

The fact that FG enhancers open earlier than intestinal enhancers is consistent with observations that vertebrate gut endoderm differentiates in an anterior-to-posterior wave [Wells and Melton 1999; Grapin-Botton and Melton 2000]. Moreover, judging by both profiles of open chromatin and the outcomes of Cdx2 deletion at different times, anterior [FG] potential predominates throughout the early
tubal endoderm and is superseded gradually by posterior (intestinal) fate in the midgut (Fig. 6D). Consequently, in the absence of factors that impose intestinal fate, distal regions may remain susceptible to anterior homeotic transformation for a longer time than proximal regions, which explains nicely why squamous differentiation is confined to the ileum when Cdx2 is deleted at or before approximately E10. We note, however, that when Cdx2 is deleted by ShhCre, prominent HS differentiation ensues in the jejunum (Fig. 3C), a homeotic transformation not seen with earlier deletion using Foxa3-Cre (Gao et al. 2009). To explain this difference, we postulate that Eso/FS and HS potential sweep through midgut endoderm in successive waves, each countered by CDX2-imposed intestinal identity. In terms of chromatin, Eso/FS-specific enhancers still dominate in the prospective ileum at approximately E10 (ShhCre), hence favoring squamous cell differentiation, while HS-specific enhancers have supplanted them in the future jejunum and promote gastric glandular differentiation. Meanwhile, Int-specific enhancers may replace both Eso/FS and HS enhancers early in the anterior midgut, so the prospective duodenum manifests intestinal potential when Cdx2 is deleted using either Foxa3-Cre or ShhCre. Because our molecular analyses probed the whole intestine, they lack the spatial resolution to confirm the details of this model (Fig. 6D), a limitation that also applies to other aspects of the study. For example, chromatin at thousands of “overlapping” and HS-specific enhancers was open in wild-type midgut endoderm until approximately E14, but we did not detect access at the global pool of FS-only enhancers. Moreover, depletion of CDX2 as early as E10 stabilized only “overlapping” and HS-unique enhancers and activated nearby genes but did not open FS-only sites, at least as detected in whole Int endoderm. Possibly, these enhancers were transiently open in the midgut before E11, or the signal was obscured because it was confined to small portions of the posterior wild-type and Cdx2−/− midgut; alternatively, different patterning mechanisms may apply early and late in endoderm development. Distinguishing among these possibilities will require fine resolution of chromatin states in various regions of the E9–E11 midgut using gene-linked cell sorting or single-cell approaches to examine tiny cell populations before anatomic landmarks become clear.

Together with the tissue defects that result from CDX2 loss, the enhancer accessibility dynamics indicate that without intestinal TF activity, FG fates represent the
Materials and methods

Experimental mice and treatments

Wild-type CD1 mice were purchased from Charles River Laboratories. Cdx2<sup>Fli/Fli</sup> [Verzelli et al. 2010], Shh<sup>Cre<sup>-</sup></sup> [Harfe et al. 2004], and Villin-Cre<sup>ER−T2</sup> [el Marjou et al. 2004] mice were reported previously. The morning of a vaginal plug was regarded as E0. Embryos from crosses between Shh<sup>Cre<sup>-</sup></sup>, Cdx2<sup>Fli/Fli</sup>, and Cdx2<sup>ER−T2</sup> females were used to isolate intestinal RNA or chromatin from adult gastric organoids, which, for the most part, retain stomach identity [Simmini et al. 2014]. Thus, Cdx2 can impose intestinal identity only during the brief window in development when enhancer dynamics permit choices among digestive endoderm fates. Such functions are not restricted to Cdx2 and the intestine. Sox2 deficiency, in particular, results in a posterior homeotic transformation of the FS, with reduced keratinization and a columnar HS-type epithelium [Que et al. 2007], whereas continuous ectopic Sox2 expression in the midgut after E8.5 induces anterior homeosis with marked stomach differentiation [Raghoebir et al. 2012]. Sox2 and Cdx2 first appear at E7.75 in the most anterior and posterior endodermal cells, respectively, and their domains extend progressively toward the FG–midgut boundary, where the two TFs are coexpressed briefly before becoming mutually exclusive near the gastric pylorus by E9.5 [Sherwood et al. 2009]. These temporal waves of expression—likely controlled by retinoic acid, BMP, and Wnt signaling posteriorly and by inhibition of these pathways anteriorly [Wells and Spence 2014, Davenport et al. 2016]—suggest how TF expression and chromatin dynamics together specify alternative endodermal fates. We propose further that initially accessible heterologous enhancers get decommissioned indirectly through mutual cross-repression by region-restricted TFs such as Sox2 and Cdx2.

In conclusion, our findings offer a basis in chromatin for classical and recent observations on endoderm plasticity and determination. The repertoire of available enhancers, likely established by positional cues, dictates the limited spectrum of possible cell fates. As late as mid-gestation, the basis for cell plasticity is that different TFs can reinforce different sets of accessible tissue-specific enhancers and allow the remaining open sites to close. Irreversible commitment to one adult cell fate (determination) follows when specific TFs have reinforced the correct complement of cis elements, and the enhancers that enable alternative cell identities are no longer available. Similar enhancer logic and plasticity likely underlie the differentiation of other organs.

Materials and methods

Tissue collection and cell isolation

Embryos were collected in ice-cold phosphate-buffered saline [PBS], and the digestive tract was collected in ice-cold Hank’s balanced salt solution [HBSS; Thermo Fisher]. The rostral half of E12 and E14 stomachs was considered the FS, and the caudal half was considered as the HS. In E16 embryos and neonates, the visible squamo–columnar junction was applied to separate the FS from the HS, discarding the interim zone. Intestines were harvested from 1 mm distal to the pylorus until [but excluding] the cecum. Tissues were digested with 0.25% Trypsin-EDTA [Thermo Fisher] for 30 min at 37°C. Released material was passed through 40-µm filters [Corning] to obtain single cells. Neonatal intestines were washed with ice-cold PBS and incubated in 5 µM EDTA solution in PBS at 4°C with gentle shaking for 30 min to release villus epithelium. Suspensions were passed through 70-µm filters [Corning] to retain villi, which were washed in cold PBS and dissociated into single cells by incubation in 4x TrypLE solution [Thermo Fisher] for 30 min at 37°C. After neutralization with 5% FBS, cells were strained over 40-µm filters, pelleted at 1200 g for 5 min, resuspended in FACs buffer (PBS, 5 mM EDTA, 2% FBS), and stained with APC-conjugated EPCAM antibody (1:100; Biolegend) for 1 h at 4°C in FACs buffer. EPCAM<sup>+</sup> cells were sorted on an Aria II SORP flow cytometer [Becton–Dickinson]. In some experiments, single cells were incubated with anti-CD326 [EPCAM] antibody conjugated to magnetic microbeads [Miltenyi Biotec, 130-105-958] for 40 min on ice and then collected over MS Columns [Miltenyi Biotec, 130-043-201] in a magnetic field. The purity of EPCAM<sup>+</sup> cells isolated by magnetic beads or flow cytometry was comparable. All replicates in the study refer to biological replicates collected on different days. In rare experiments lacking a biological replicate, data from other samples or time points support the conclusions.

RNA-seq, ATAC-seq, and ChIP-seq (ChIP combined with high-throughput sequencing)

Total RNA was isolated from FACs-sorted epithelial cells using TRIzol [Thermo Fisher] and RNeasy microkit [Qiagen]. Libraries were prepared using the SMARTer-Seq version 4 low-input mRNA library kit [Clonetech]. Five-thousand to 30,000 cells were used for ATAC-seq as described previously [Buenrostro et al. 2015], with minor modifications. Cells were lysed in ice-cold buffer [10 mM Tris at pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.1% [v/v] Igepal CA-630], and nuclei were precipitated by centrifugation at 500 g for 10 min at 4°C then treated with Nextera Tn5 transposase [Illumina, FC-121-1030] for 30 min at 37°C. Transposed chromatin was purified with QIAquick PCR purification kits [Qiagen] and amplified with High-Fidelity 2X PCR Master Mix [New England Biolabs] using a universal forward primer and unique reverse primers in a 50-µL reaction for five cycles. Five microliters of the reaction mix was amplified by quantitative PCR for 20 cycles to determine the optimum number of addition cycles, defined as one-third the number needed to achieve maximum fluorescence intensity. Primer dimers were excluded from these PCR-amplified libraries using Agencourt Ampure XP beads [Beckman Coulter].
For fetal ChIP-seq, intestines were dissected from up to 36 embryos in ice-cold HBSS and disrupted by vigorous pipetting followed by cross-linking in 1% formaldehyde for 20 min at room temperature, incubation for 10 min on ice in chromatin preparation buffer (ChIP-IT high-sensitivity kit, Active Motif, 53040) with protease inhibitors [Roche, 11873580001] and phenyl methyl sulfonyl fluoride (PMSF), and centrifugation at 1250 g for 5 min at 4°C. The pellet was resuspended in 0.5 mL of ChIP buffer (ChIP-IT high-sensitivity kit, Active Motif) and incubated for 10 min on ice. Lysates were sonicated in an E220 focused ultrasonicator (Covaris). CDX2 antibody-bound (Cell Signaling Technology) protein A and protein G magnetic Dynabeads (1:1 ratio; Thermo Fisher) were incubated overnight at 4°C with 2 μg of chromatin in 300 μL of ChIP buffer containing protease inhibitors and two carriers: 20 μL of recombinant human histone H2B solution [New England Biolabs, M2505S] and 1 μL of carrier RNA from the RNaseasy microkit [Qiagen, 74004]. Beads were washed five times with LiCl-RIPA buffer (10 mM Tris-HCl at pH 8, 500 mM LiCl, 1 mM EDTA, 1% NP-40) and once in 10 mM Tris-EDTA (pH 8). Beads resuspended in 100 μL of immunoprecipitation elution buffer containing 25 μg of RNase A were incubated for 30 min at 37°C followed by 1 μL of 20 mg/mL proteinase K (Active Motif), and cross-links were reversed overnight at 65°C. DNA was purified using MinElute PCR purification kits [Qiagen]. ChIP-seq for adult villi with CDX2 [Cell Signaling Technology, 12306] or H3K4me1 [Diagenode, C15410194] antibody was as described previously [Saxena et al. 2017]. Libraries were prepared using ThruPLEX kits [Rubicon, R400427].

All libraries were sequenced on a NextSeq 500 instrument [Illumina] to generate 75-base-pair [bp] single-end reads.

**Histochemistry and immunohistochemistry (IHC)**

Intestines were fixed overnight in 4% paraformaldehyde, washed in PBS, dehydrated in an ethanol series, and embedded in paraffin to prepare 5-μm-thick sections. Slides were incubated in 0.5% periodic acid and stained with Schiff’s reagent [Alfa Aesar, J612171], and alkaline phosphatase [ALP] activity was measured using AP staining kit II [Stemgent]. For IHC, antigens were retrieved in 10 mM sodium citrate buffer (pH 6), and sections were then incubated overnight at 4°C with antibody against CDX2 (Cell Signaling Technology, 12306) or H3K4me1 [Diagenode, C15410194] antibody was as described previously [Saxena et al. 2017]. Libraries were prepared using ThruPLEX kits [Rubicon, R400427]. All libraries were sequenced on a NextSeq 500 instrument [Illumina] to generate 75-base-pair [bp] single-end reads.

**RNA-seq analysis**

Reads were aligned to the mouse reference genome Mm9 [NCBI build 37] using STAR aligner version 2.5.3a [Dobin et al. 2013], and raw read counts from HTSeq [Anders et al. 2015] were used to estimate transcript levels. Unsupervised PCA was performed on log-transformed normalized counts using R [https://www.r-project.org]. We then used DESeq2 [Love et al. 2014] to normalize data using size factors or rlog transformation from E16 wild-type F5, HS, and Int [±2.5 samples each] and identify significant (q < 0.05) differential expression. RPMK values were calculated using Cufflinks [Trapnell et al. 2012]. Pearson correlation coefficients of normalized count values from biological replicates [n ≥ 2 samples each] were calculated using the corrpack package in R. Data tracks to visualize on the Integrative Genomics Viewer [IGV] version 2.3 [Robinson et al. 2011] were generated using bamCoverage modules in deepTools2 [Ramirez et al. 2016]. We determined optimal cluster numbers for K-means clustering of rlog transformed counts using the clusGap function in R, derived Gene Ontologies in DAVID [Huang et al. 2009], and generated violin plots using the ggplot2 package [Wickham 2009] in R studio. We used GENE-E software [Broad Institute] to plot log transformed values in heat maps.

To determine genes associated with P1 Eso/FS- and Int-specific enhancer clusters (Fig. 1) and up-regulated genes associated with retained FG enhancers [Supplemental Fig. SSE], we used BETA [Wang et al. 2013]. Enriched expression of gene sets in the FS or Int was determined using gene set enrichment analysis [GSEA] [Subramanian et al. 2005], and normalized enrichment scores were derived using 1000 permutations of random gene sets of the same size. RNA-seq data from wild-type and Cdx2−/− (deleted at E10) intestines were associated with a merged set of HS and “overlapping” ATAC+ sites persisting in E10 deleted Cdx2−/− intestines.

**ATAC-seq and ChIP-seq analyses**

For ChIP-seq, 75-bp single-end reads were aligned to mouse reference genome Mm9 [NCBI build 37] using Bowtie2 [Langmead and Salzberg 2012]. Following routine quality control, we reassessed all ATAC-seq data and aligned reads using the ENCODE pipeline [Supplemental Tables S1, S2]. Nonunique and duplicate reads were removed using SAMtools version 1.3.1 [Li et al. 2009]. We used individual or merged raw signals from replicate ChIP-seq and ATAC-seq files to visualize in IGV tracks and then created bigwig files using the bamCoverage module in deepTools2. Data were quantile-normalized using Haystack version 0.4 [Pinello et al. 2014] followed by deepTools2 to plot heat maps. ATAC-seq signals were computed using SitePro in Cistrome tools [Liu et al. 2011], and intensity values were used to plot cumulative signals using R. Differential ATAC-seq signals between tissue pairs were identified using DiffReps version 1.55.4 [Shen et al. 2013] with z-score cutoff of 5, window size of 500 bp, step size of 100 bp, and P-value of 0.05. Enhancers were defined by excluding promoters (±1 kb from a TSS). We then used BEDTools [Quinlan and Hall 2010] to determine the overlap between the FS and HS in the DiffReps output. These overlapping regions were subtracted from the individual FS and HS sets and assigned to a distinct “Overlap” set, while others were considered specific to one region or the other. Similarly, sites common or unique to FS versus Int and HS versus Int pairs were determined by intersecting with the set of Int-specific regions. To identify the shared regions, peaks were called using MACS2 [Zhang et al. 2008] for each tissue and intersected using BEDTools, excluding promoters.

To determine correlations between samples, normalized read counts from replicate samples were collated over merged called peaks (generated by MACS2) from individual bam files using BEDTools. To determine cross-correlations, read counts were collated over peak files merged from each sample and plotted in R. Functions of genes within prescribed distances from ATAC sites were identified using GREAT version 3.0 [McLean et al. 2010], and motif enrichment was identified using HOMER version 4.7.2 [Heinz et al. 2010] based on cumulative binomial distributions against the genome background. ChIP-seq and ATAC-seq read counts ±250 bp from peak summits were calculated from normalized data and are represented as log transformed values in violin plots. Differences among groups were determined using the Kolmogorov-Smirnov test. As P-values were generally significant (<0.001 to <0.01) owing to large input sample sizes, we inferred similarities and differences using the D statistic, the maximum distance between two cumulative distribution functions (e.g., see Supplemental Figs. S3B, S5C).
All data have been deposited in the Gene Expression Omnibus, series GSE115541.

Acknowledgments

We thank S. Robine for generously sharing Villin-CreERT2 mice, and Natalia Soshnikova for helpful discussion. This work was supported by National Institutes of Health awards R01DK08 2889 and R01DK081113 to R.A.S., R01CA190558 and U01DK1 03141 to M.P.V., F32DK115080 to M.S., K01DK113067 to U.J., and P50CA127003.

Author contributions: A.C., K.K.B., N.K., M.P.V., and R.A.S. conceived and designed the studies. K.K.B., A.C., N.K., L.C., M.S., N.H.T., and U.J. acquired data. K.K.B., M.S., N.K.O., S.M., and M.P.V. performed computational analyses. M.P.V. and R.A.S. supervised the study. M.S. and R.A.S. drafted the manuscript, with input from all authors.

References

Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166–169.

Beck F, Erler T, Russell A, James R. 1995. Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. Dev Dyn 204: 219–227.

Beck F, Chawengsaksophak K, Waring P, Playford RJ, Furness JB. 1999. Reprogramming of intestinal differentiation and inter-calary regeneration in Cdx2 mutant mice. Proc Natl Acad Sci 96: 7318–7323.

Buennonstro JD, Wu B, Chang HY, Greenleaf WJ. 2015. ATAC-seq: a method for assaying chromatin accessibility genome-wide. Curr Protoc Mol Biol 109: 21.29.21–29.

Davenport C, Dieckmann U, Budd I, Detering N, Nauisk O. 2016. Anterior-posterior patterning of definitive endoderm generated from human embryonic stem cells depends on the differential signaling of retinoic acid, Wnt- and BMP-signaling. Stem Cells 34: 2635–2647.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29: 15–21.

Duluc I, Freund JN, Leberquier C, Keding M. 1994. Fetal endoderm primarily holds the temporal and positional information required for mammalian intestinal development. J Cell Biol 126: 211–221.

Echelard Y, Epstein DJ, St-Jacques B, Shen L, Mohler J, McMahon JA, McMahon AP. 1993. Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. Cell 75: 1417–1430.

el Marouf J, Janssen KP, Chang BH, Li M, Hindle V, Chan L, Louvand DC, Chambon P, Metzger D, Robine S. 2004. Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. Genesis 39: 186–193.

Gao N, White P, Kaestner KH. 2009. Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2. Dev Cell 16: 588–599.

Grainger S, Savory JG, Lohnes D. 2010. Cdx2 regulates patterning of the intestinal epithelium. Dev Biol 339: 155–165.

Grapin-Botton A. 2005. Antero-posterior patterning of the vertebrate digestive tract. Int J Dev Biol 49: 335–347.

Harfe BD, Scherz PJ, Nissim S, Tian H, McMahon AP, Tabin CJ. 2004. Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. Cell 118: 517–528.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38: 576–589.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protoc 4: 44–57.

James R, Erler T, Kazenwadel J. 1994. Structure of the murine homeobox gene cdx-2. Expression in embryonic and adult intestinal epithelium. J Biol Chem 269: 15229–15237.

Kazakevych J, Sayola S, Messner B, Kriemke C, Soshnikova N. 2017. Dynamic changes in chromatin states during specification and differentiation of adult intestinal stem cells. Nucleic Acids Res 45: 5770–5784.

Kong J, Crissey MA, Funakoshi S, Kreindler JL, Lynch JP. 2011. Ectopic Cdx2 expression in murine esophagus models an intermediate stage in the emergence of Barrett’ s esophagus. PLoS One 6: e18280.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9: 357–359.

Lara-Astasio D, Weiner A, Lorenzo-Vivas E, Zaretksy I, Jaitin DA, David E, Keren-Shaul H, Mildner A, Winter D, Jung S, et al. 2014. Immunogenetics. Chromatin state dynamics during blood formation. Science 345: 943–949.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Liu T, Ortiz JA, Taing L, Meyer CA, Lee B, Zhang Y, Shin H, Wong SS, Ma J, Lei Y, et al. 2011. Cistrome: an integrative platform for transcriptional regulation studies. Genome Biol 12: R83.

Long HK, Prescott SL, Wysocka J. 2016. Ever-changing landscapes: transcriptional enhancers in development and evolution. Cell 167: 1170–1187.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15: 550.

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 28: 495–501.

Mizuno T, Yasugi S. 1990. Susceptibility of epithelia to directive signals during mammalian development. Cell Diff Dev 31: 151–159.

Mutoh H, Hakamata Y, Sato K, Eda A, Yanaka I, Honda S, Osawa H, Kaneko Y, Sugano K. 2002. Conversion of gastric mucosa to intestinal metaplasia in Cdx2-expressing transgenic mice. Biochem Biophys Res Commun 294: 470–479.

Nord AS, Blow MJ, Attanasio A, Akymaeva JA, Holt A, Hosseini R, Phouanenavong S, Plajzer-Frick I, Shoukry M, Afzal V, et al. 2013. Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. Cell 155: 1521–1531.

Pinello L, Xu J, Orkin SH, Yuan GC. 2014. Analysis of chromatin-state plasticity identifies cell-type-specific regulators of morphogenesis and cytodifferentiation. Cell Diff Dev 31: 151–159.

Que J, Okubo T, Goldenring JR, Nam KT, Kurotani R, Morrissey EE, Tian LY, Vegas P, Pevny LH, Hogan BL. 2007. Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. Development 134: 2521–2531.

Chromatin-based endoderm plasticity

Data availability

All data have been deposited in the Gene Expression Omnibus, series GSE115541.
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842.

Raghoeber L, Bakker ER, Mills IC, Swagemakers S, Kapen MB, Munck AB, Driegen S, Meijer D, Grosfeld F, Tibboel D, et al. 2012. SOX2 redirects the developmental fate of the intestinal epithelium toward a premature gastric phenotype. J Mol Cell Biol 4:377–385.

Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. 2016. DeepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44:W160–W165.

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. Nat Biotechnol 29:24–26.

Saxena M, Roman AKS, O’Neill NK, Sulahian R, JadHAV U, Shivdasani RA. 2017. Transcription factor-dependent ‘anti-repressive’ mammalian enhancers exclude H3K27me3 from extended genomic domains. Genes Dev 31:2391–2404.

Shen L, Shao NY, Liu X, Maze I, Feng J, Nestler EJ. 2013. diffReps: detecting differential chromatin modification sites from ChiP-seq data with biological replicates. PLoS One 8:e65598.

Sherwood RJ, Jitianu C, Cleaver O, Shaywitz D, Lemenzo JO, Chen AE, Golub TR, Melton DA. 2007. Prospective isolation and global gene expression analysis of definitive and visceral endoderm. Dev Biol 304:541–555.

Sherwood RJ, Chen TY, Melton DA. 2009. Transcriptional dynamics of endodermal organ formation. Dev Dyn 238:29–42.

Shlyueva D, Stampfel G, Stark A. 2014. Transcriptional enhancers: from properties to genome-wide predictions. Nat Rev Genet 15:272–286.

Silberg DG, Swain GP, Suh ER, Traber PG. 2000. Cdx1 and cdx2 expression during intestinal development. Gastroenterology 119:961–971.

Silberg DG, Sullivan J, Kang E, Swain GP, Moffett J, Sund NJ, Sackett SD, Kaestner KH. 2002. Cdx2 ectopic expression induces gastric intestinal metaplasia in transgenic mice. Gastroenterology 122:689–696.

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

Slack JM. 1991. From egg to embryo: regional specification in early development. Cambridge University Press, New York.

Spence JR, Lauf R, Shroyer NF. 2011. Vertebrate intestinal endoderm development. Dev Dyn 240:501–520.

Stergachis AB, Nep S, Reynolds A, Humbert R, Miller B, Paige SL, Vernet B, Cheng JB, Thurman RE, Sandstrom R, et al. 2013. Developmental fate and cellular maturity encoded in human regulatory DNA landscapes. Cell 154:888–903.

Stringer EJ, Duluc I, Saandi T, Davidson I, Bialecka M, Sato T, Barker N, Clevers H, Fritchard CA, Winton DJ, et al. 2012. Cdx2 determines the fate of postnatal intestinal endoderm. Development 139:465–474.

Subramanian A, Tamayo P, Mootha VK, Mukhejee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci 102:15545–15550.

Sulahian R, Chen J, Arany Z, JadHAV U, Peng S, Rustgi AK, Bass AJ, Srivastava A, Hornick JL, Shivdasani RA. 2015. SOX15 governs transcription in human stratified epithelia and a subset of esophageal adenocarcinomas. Cell Mol Gastroenterol Hepatol 1:598–609 e596.

Trappell C, Roberts A, Goff L, Pertia G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7:562–578.

Verzi MP, Shin H, He HH, Sulahian R, Meyer CA, Montgomery RK, Fleet JC, Brown M, Liu XS, Shivdasani RA. 2010. Differentiation-specific histone modifications reveal dynamic chromatin interactions and partners for the intestinal transcription factor CDX2. Dev Cell 19:713–726.

Verzi MP, Shin H, Ho LL, Liu XS, Shivdasani RA. 2011. Essential and redundant functions of caudal family proteins in activating adult intestinal genes. Mol Cell Biol 31:2026–2039.

Verzi MP, Shin H, San Roman AK, Liu XS, Shivdasani RA. 2013. Intestinal master transcription factor CDX2 controls chromatin access for partner transcription factor binding. Mol Cell Biol 33:281–292.

Walton KD, Kolterud A, Czerwinski MJ, Bell MJ, Prakash A, Kushwaha J, Grosse AS, Schnell S, Gumucio DL. 2012. Hedgehog-responsive mesenchymal clusters direct patterning and emergence of intestinal villi. Proc Natl Acad Sci 109:15817–15822.

Wang S, Sun H, Ma J, Zang C, Wang C, Wang J, Tang Q, Meyer CA, Zhang Y, Liu XS. 2013. Target analysis by integration of transcriptome and ChIP-seq data with BETA. Nat Protoc 8:2502–2515.

Wang A, Yue F, Li Y, Xie R, Harper T, Patel NA, Muth K, Palmer J, Qiu Y, Wang J, et al. 2015. Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. Cell Stem Cell 16:386–399.

Wells JM, Melton DA. 1999. Vertebrate endoderm development. Annu Rev Cell Dev Biol 15:393–410.

Wells JM, Melton DA. 2000. Early mouse endoderm is patterned by soluble factors from adjacent germ layers. Development 127:1563–1572.

Wells JM, Spence JR. 2014. How to make an intestine. Development 141:752–760.

Wickham H. 2009. ggplot2: elegant graphics for data analysis. Springer-Verlag, New York.

Yasugi S, Mizuno T. 2008. Molecular analysis of endoderm regionalization. Dev Growth Differ 50:879–896.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based analysis of ChIP-seq [MACS]. Genome Biol 9:R137.

Zorn AM, Wells JM. 2009. Vertebrate endoderm development and organ formation. Annu Rev Cell Dev Biol 25:221–251.