Charting human development using a multi-endodermal organ atlas and organoid models

Graphical abstract

Highlights

- Cell atlas of multiple developing human endoderm-derived organs
- Identified organ-specific epithelial stem cell and mesenchymal cell signatures
- Benchmarked intestinal organoid fidelity and maturation using the multi-organ atlas
- Interrogated genetic and culture perturbations of epithelium and mesenchyme development

Authors

Qianhui Yu, Umut Kilik, Emily M. Holloway, ..., Barbara Treutlein, Jason R. Spence, J. Gray Camp

Correspondence

barbara.treutlein@bsse.ethz.ch (B.T.), spencejr@med.umich.edu (J.R.S.), grayson.camp@iob.ch (J.G.C.)

In brief

A reference atlas of multiple human developing endodermal organs of the respiratory and gastrointestinal tract is described and used to provide information regarding cell states, transcription factors, and organ-specific epithelial stem cell and mesenchyme interactions across lineages as well as benchmark stem cell-derived human intestinal organoids under multiple culture conditions.

Yu et al., 2021, Cell 184, 3281–3298

June 10, 2021 © 2021 The Author(s). Published by Elsevier Inc.

https://doi.org/10.1016/j.cell.2021.04.028
Resource

Charting human development using a multi-endodermal organ atlas and organoid models

Qianhui Yu,1,8 Umut Kilik,1,2,8 Emily M. Holloway,3,6 Yu-Hwai Tsai,4 Christoph Harmel,1,2 Angeline Wu,4 Joshua H. Wu,4 Michael Czerwinski,4 Charlie J. Childs,3 Zhisong He,7 Meghan M. Capeling,7 Sha Huang,4 Ian A. Glass,6 Peter D.R. Higgins,4 Barbara Treutlein,5,* Jason R. Spence,3,4,7,* and J. Gray Camp1,2,9,*

1Institute of Molecular and Clinical Ophthalmology Basel, 4031 Basel, Switzerland
2Department of Ophthalmology, University of Basel, 4031 Basel, Switzerland
3Department of Cell and Developmental Biology, University of Michigan Medical School, Ann Arbor, MI 48109, USA
4Department of Internal Medicine, Gastroenterology, University of Michigan Medical School, Ann Arbor, MI 48109, USA
5Department of Biosystems Science and Engineering, ETH Zürich, 4058 Basel, Switzerland
6Department of Pediatrics, Genetic Medicine, University of Washington, Seattle, WA 98195, USA
7Department of Biomedical Engineering, University of Michigan College of Engineering, Ann Arbor, MI 48109, USA
8These authors contributed equally
9Lead contact
*Correspondence: barbara.treutlein@bsse.ethz.ch (B.T.), spencejr@med.umich.edu (J.R.S.), grayson.camp@iob.ch (J.G.C.)
https://doi.org/10.1016/j.cell.2021.04.028

SUMMARY

Organs are composed of diverse cell types that traverse transient states during organogenesis. To interrogate this diversity during human development, we generate a single-cell transcriptome atlas from multiple developing endodermal organs of the respiratory and gastrointestinal tract. We illuminate cell states, transcription factors, and organ-specific epithelial stem cell and mesenchyme interactions across lineages. We implement the atlas as a high-dimensional search space to benchmark human pluripotent stem cell (hPSC)-derived intestinal organoids (HIOs) under multiple culture conditions. We show that HIOs recapitulate reference cell states and use HIOs to reconstruct the molecular dynamics of intestinal epithelium and mesenchyme emergence. We show that the mesenchyme-derived niche cue NRG1 enhances intestinal stem cell maturation in vitro and that the homeobox transcription factor CDX2 is required for regionalization of intestinal epithelium and mesenchyme in humans. This work combines cell atlases and organoid technologies to understand how human organ development is orchestrated.

INTRODUCTION

The human body is composed of an extraordinary diversity of cells that originate from a zygote. Although much about embryogenesis and organogenesis has been revealed using non-human model organisms, differences between human and model organism development highlight the need for human models (Miller et al., 2018; Nikolic et al., 2017). Technologies to measure transcriptomes from single cells have opened up new inroads into understanding developing human organs. Single-cell atlases of developing and mature human organs are revealing new cell types and states (Cao et al., 2020; Han et al., 2020b; Taylor et al., 2019) and providing insight into potential disease mechanisms by identifying cell types that express disease-associated genes (Camp et al., 2019; Cowan et al., 2020; Elmentaite et al., 2020; Fawkner-Corbett et al., 2021; Wu et al., 2018). Human atlases also serve as a gold-standard reference for human cell and tissue engineering, enabling fidelity assessment of in vitro models.

Organoids are three-dimensional multicellular culture systems that recapitulate aspects of human physiology. Organoids are attractive model systems because they can be manipulated genetically, observed in controlled in vitro environments over time, and derived from individuals with diverse genetic backgrounds. There are many protocols to generate organoids from human pluripotent stem cells (hPSCs) that resemble different organ or tissue types (Kechele and Wells, 2019; Rossi et al., 2018), and each system has similarities to and differences from the primary counterpart. Here we use hPSC-derived intestinal organoids (HIOs) as a multilineage model system to study intestinal development (Spence et al., 2011). HIOs are thought to recapitulate the earliest phases of organ specification and complex tissue formation. Generating HIOs relies on directed differentiation through temporal manipulation of key signaling pathways via growth factors and small molecules to mimic intestinal organogenesis (McCracken et al., 2011; Wells and Spence, 2014; Zorn and Wells, 2009). The HIO system has been used over the past years as a model to study human intestinal development (Capeling et al., 2019; Du et al., 2012; Finkbeiner et al., 2015a; Holloway et al., 2020; Kumar et al., 2019; Schlieve et al., 2017; Spence et al., 2011; Tsai et al., 2017; Watson et al., 2014) and...
A developing human multi-organ cell atlas identifies organ-specific features

We established a developing human multi-organ reference cell atlas focusing on endoderm-derived organs by integrating newly generated and published single-cell (sc) transcriptomes from lung, esophagus, liver, stomach, small intestine (duodenum, jejunum, and ileum), and colon (155,232 cells in total) with an age distribution from 7–21 post-conception weeks (PCWs) (Figure 1A; Tables S1 and S2; key resources table; Holloway et al., 2020, 2021; Miller et al., 2020). We integrated all of the developing human scRNA sequencing (scRNA-seq) data using cluster similarity spectrum (CSS) (He et al., 2020) to correct for batch effects. Clustering resolved major cell classes, including epithelial, mesenchymal, immune, endothelial, neuronal, and erythroid populations (Figures 1B and 1C), that could be subdivided into 27 molecularly distinct clusters (Figures 1D and S1A). Some epithelial and mesenchymal clusters were dominated by a specific organ, whereas neural, immune, and endothelial clusters contained cells from multiple organs (Figure S1A). We further quantified the organ specificity of each of these classes by examining the proportion of cells showing transcriptomes similar to one or more other organs (Figure S1B). It showed that epithelial and mesenchymal cells exhibited the strongest organ specificity. We identified transcription factors (TFs) with enriched expression in the epithelium and mesenchymal of each organ (Figure S1C). These results established a higher-order map of cell populations and associated marker genes across 11 regions of 5 developing human organs.

To resolve epithelial heterogeneity, we performed subclustering for each organ separately and annotated cell types based on markers. This analysis identified a total of 39 clusters, recovering previously known epithelial cell types of each organ (Figures 2A–2D; Table S2; Data S1 at https://doi.org/10.17632/x53tts3zfr.1). Next we extracted stem cell clusters from each organ to determine inter-organ similarities and differences (Figure 2E). We identified common genes that associate with stem cells throughout the organ atlas (Table S3; STAR Methods), and these genes were enriched for gene ontologies related to metabolism (glycolysis and gluconeogenesis) and translation initiation. We also observed differences between stem cells among different organs and tissues, indicated by separation of clusters (Becht et al., 2018; Figure 2E). We ordered the stem cells in a pseudospatial trajectory according to regional identity score (Figure 2F; STAR Methods). This revealed distinctions among organs as well as progressive transition from anterior to posterior in the intestine. Organ specificity was highlighted by stem cell markers of the stomach and intestine, LGR5 and OLFM4 (Barker et al., 2007; van der Flier et al., 2009). Although LGR5 is expressed throughout the stem cell populations of the stomach and intestine, OLFM4 is highly enriched in a subpopulation of small intestine stem cells (Figure 2E). Next we identified TFs that vary along this trajectory (Table S3). Some TFs are shared by multiple organs (KLF5 and ZNF706 in the GI tract), whereas others are tissue restricted (ONECUT2, duodenum; SATB2, colon) (Dusing et al., 2010; Munera et al., 2017; Tsai et al., 2017; Figure 2F). Many of these TFs have unexplored roles in specification or maintenance of tissue-specific stemness. These data provide a catalog of molecular profiles that mark cell types along the developing human endoderm-derived epithelium.

We next analyzed neuronal and mesenchymal cell diversity in the developing organs. We cataloged 6 glial and 8 neuronal clusters, most of which were distributed among the organs, and we provide marker genes for diverse cell populations of the developing human enteric nervous system (Data S1). Next we cataloged 21 molecularly distinct mesenchymal clusters (Figure 3A; Table S2; Data S1), identifying known subpopulations, such as vascular smooth muscle cells (VSMCs; cluster 2 [c2], DES+/ RGS5+), interstitial cells of Cajal (ICCs; c4, ANO1+/KIT+), and chondrocyte-like cells (c21, SOX9+/COL9A3+), as well as novel subpopulations, such as GDF10+ (c8, GDF10+/WNT2B+) and SERPINF1-high (c20, SERPINF1+/ABC8+) cells. We find that lung-associated mesenchyme is largely distinct from that of the stomach and intestine, and we identified TFs and signaling molecules that codify lung and GI tract identity (Figures 3B and S1C). We validated the TF NKX2-3 and secreted factor WNT2...
as general markers that distinguish intestine and lung mesenchyme, respectively (Figure 3C; Data S1). Further, the TF SALL1 specifically marks F3+ subepithelial mesenchyme (c15, subepithelial, NRG1+/F3+) in the stomach and intestine, whereas the TF PRRX1 marks chondrocyte-like SERPINF1-high cells and pericytes in the lung (Figures 3B and 3C; Data S1). We also observed signatures that distinguish mesenchyme populations along the GI tract. For example, homeobox genes showed anterior-posterior expression patterns reminiscent of other model systems and human adult tissues (Chang et al., 2002; Young and Deschamps, 2009; Figure 3D). Using receptor-ligand pairing analysis (Efremova et al., 2020), we cataloged interactions between epithelium and mesenchymal populations and identified many pairings that were enriched in particular organs (Figure 3E; Data S1). For example, a LGR5-RSPO3 receptor-ligand pairing was enriched in epithelial stem cell and mesenchyme in the stomach, small intestine, and colon. The multi-organ sc transcriptome reference atlas unveils a rich resource of transcriptional and inter-lineage signaling programs that are specific to cell populations within each organ and may

Figure 2. Epithelial cell heterogeneity and stem cell differences across organs
(A–D) Epithelial cell UMAP, colored by cell type in the developing human (A) lung, (B) esophagus, (C) stomach, and (D) intestine.
(E) Epithelial stem cell UMAP from multiple organs, with cells colored by tissue. The inset shows feature plots of LGR5 and OLFM4 expression.
(F) Heatmap showing expression patterns of TFs across epithelial stem cell bins ordered by pseudospace. The top sidebar shows the tissue proportion within each cell bin (20 cells/bin).
coordinate mesenchymal-epithelial signaling in organ-specific stem cell niches.

Mapping to a multi-organ reference atlas reveals the fidelity of human intestinal organoids

We next sought to use the atlas to understand the fidelity of intestinal organoids. We generated HIOs derived from embryonic stem cells and analyzed cellular heterogeneity 4 and 8 weeks after in vivo transplantation into the kidney capsule of an immunocompromised mouse host (Figures 4A and 4B). Transplanted HIOs (tHIOs) exhibited hallmarks of epithelium maturation similar to that in the developing human intestine, including emergence of stereotypic crypt-villus architecture (Figure 4C; Finkbeiner et al., 2015b; Tsai et al., 2017; Watson et al., 2014), with proliferative (MKI67+) intestinal stem cells (ISC; LGR5+/OLFM4+) localized to crypt domains and differentiated secretory lineages (CHGA+ enteroendocrine, MUC2+ goblet cells) and FABP2+ absorptive enterocytes along the villus epithelium (Figures 4D and 4E; Data S1). We performed scRNA-seq on tHIOs and observed ISCs (c6, LGR5+/OLFM4+), enterocyte precursors (c7, FABP2+/APOA4 low), enterocytes (c8, FABP2+/APOA4+), enteroendocrine cells (c9, CHGA+/ISL1+), M cells (c10, SPIB+/CA7+), and goblet cells (c11, MUC2+/SPINK4+) (Figures 4F and 4G; Data S1). We detected additional epithelial clusters (c12–c15) that lacked the expression of intestinal marker genes such as CDX2 and instead expressed foregut markers such as MUC5AC, SOX2, and FOXJ1 (Data S1). We also identified diverse mesenchymal clusters (c1–c5), including VSMCs (c5, DES+/RGS5+) and CXCL14+ mesenchyme (c4, CXCL14+/NSG1+). We detected additional mesenchymal clusters (c1–c5) with enriched expression of the intestine marker NKX2-3 (Data S1) and showed enriched expression of PRRX1, which is highly...
expressed in the developing lung. This foregut-like off-target mesenchymal subtype accounts for 27.6% (187 of 678) of sequenced tHIO mesenchymal cells.

To provide a quantitative assessment of tHIO transcriptome fidelity, we calculated the distance of each tHIO cell to the developing multi-organ atlas and projected tHIO cells to the reference (Figures 4H, 4I, and S1D). We found that approximately 70% of tHIO epithelial and mesenchymal cells were mapped to the developing human intestine, and these on-target proportions were similar to qualitative assessments based on organ-specific markers. Notably, tHIO stem cells mapped to the intestine and expressed small intestine features, such as OLFM4, CDX2, and PDX1 (Data S1). We compared tHIO stem cells with the GI tract stem cell pseudospace and found that tHIO stem cells are most similar to duodenal stem cells (Figure 4J; Table S2; Data S1). Additionally, we found that genes associated with neonatal and pediatric digestive disorders showed consistent expression patterns between tHIO and developing duodenum (Table S2; Data S1). tHIOs contain diverse epithelial and mesenchymal cells, most of which map to the developing small intestine with high fidelity.

Integrating HIO and primary duodenum data enables tracking of molecular transitions during human ISC development

Next we generated an sc transcriptome reference of the adult duodenum (Figures S2B and S2C; Table S2). We integrated tHIO, developing, and adult intestinal epithelial subtypes to quantify tHIO epithelial cell maturation (Figures 5A and S2A). In an integrated uniform manifold approximation and projection (UMAP), we found that developing duodenum and tHIO stem cell-to-enterocyte differentiation trajectories were distinct from the adult (Figure 5A). Indeed, we found that each tHIO cell type was more similar to the developing counterpart compared with the adult (Figures 5B and S2D–S2F). Notably, we found that tHIO and developing duodenum stem cells are highly similar and molecularly distinct from the adult state (Figures 5C and 5D; Videos S1, S1, and S3). We also compared transcriptomes from adult-derived intestinal enteroids (Fujii et al., 2018) and found that these cell states are more similar to the adult cells (Figures S2G and S2H). These analyses show that tHIO stem cells are more similar to the developing intestine than their adult counterparts.

Next we analyzed sc transcriptomes across an HIO time course from endoderm induction through 30 days of in vitro differentiation to illuminate the early molecular transitions that lead to intestinal stem cell (ISC) specification (Figure 5E). We constructed a force-directed k-nearest neighbor (kNN) graph (Weinreb et al., 2018) to visualize the temporal progression of cell fate acquisition (Figures 5E and 5F; Table S2). Given that there is no human reference for the early differentiation events in HIOs, we compared the in vitro organoid time course with a mouse gastrulation atlas (Pijuan-Sala et al., 2019). This analysis revealed that cell clusters from the HIO early time point (day 0–spheroid, c4 and c13) expressed definitive endoderm (EOMES, SOX17, and FOXA2) and primitive streak markers (MIXL1, GSC, and LHX1) and were mapped predominantly to the anterior primitive streak (Figures S3A–S3D). We identified an HIO epithelial developmental trajectory (c0, c2, c6, c7, c11, c12, and c14) marked by co-expression of CDX2 and CDH1 and mapped predominantly to the mouse gut epithelium (Figures S3A–S3D). We also identified a mesenchymal developmental trajectory (c1, c3, c5, c8, and c10) that showed the highest similarity to mouse mesenchyme/mesoderm populations (Figures S3A–S3D). We noted that, although cluster 9 was CDX2+/CDH1+ and showed comparable similarity to gut epithelium and surface ectoderm, it also exhibited mesoderm features (HAND1+/FOXF1+) (Figures S3A–S3D). We also observed low-abundance neural-like (c15 and c18), endothelial-like (c16), and gut epithelium-surface ectoderm-like (c11) cells (Figures S3A–S3D). The proportion of these cell types decreased after 14 days in culture, and 4-week HIOs were composed predominantly of epithelial and mesenchymal cell populations (Figures 5E and 5F).

Next, we reconstructed an ISC maturation roadmap from pluripotency (Figure 5G). We first selected intestine epithelial cells from the in vitro HIO time course based on CDX2 expression and mapping results to the developing multi-organ reference (Figures 5F and S3C; STAR Methods). The selected in vitro HIO cells were most similar to stem cells of tHIO and primary intestine tissues (Figure S3E). We then combined these cells with stem cells from tHIOs and the developing duodenum. We decomposed the transcriptome of each cell into the early intestinal progenitor (day 0–spheroid) and the most mature developing ISC (19 PCW) component and ordered cells based on increasing developing ISC maturity (Figure S3F). We noted a correspondence between organoid and developing duodenum time points.

Figure 4. hPSC-derived intestinal organoids (HIOs) recapitulate developing small intestine features

(A) Schematic of HIO development.
(B) HIOs 10 weeks after transplantation (HIOs). Scale bar, 1 mm.
(C) IF staining for the epithelial marker Ecad (green) in 10-wks HIOs. Scale bar, 200 μm.
(D) IF staining for 10-week HIO for epithelial lineage markers (FABP2, CHGA, and LCT), proliferation (MKI67), and brush border (SI) counterstained with the epithelial marker (EPCAM, green) and DAPI (blue). Left: schematic of epithelial cell types. Scale bars, 200 μm.
(E) RNA in situ hybridization of stem cell marker genes in HIOs. Scale bars, 200 μm.
(F) UMAP of sc transcriptomes from tHIO colored and numbered by cluster (left) or time point (top right).
(G) Stacked bar plot showing epithelial cell type composition in tHIOs for intestinal cells; color scheme as in (F).
(H) Top: tHIO cells are projected onto the reference atlas embedding and colored by cell fate. Bottom: the proportion of tHIO cells mapping to the epithelium (left) or mesenchyme (right) of each organ.
(I) Boxplots showing CDX2 and NKX2-3 expression of each tHIO cluster. The sidebar shows cell fate as in (H). The stacked bar plot shows the developing organ projection probability.
(J) The top sidebar shows the tissue composition of each cell bin. The bean plot shows the quantile of the pseudospatial score of each tissue among all examined cells.
where data from 47, 59, and 85 post-conception days (i.e., 7, 8, 12 PCWs) tissues showed the most similar ISC identity to day 14 HIO and 4- and 8-week tHIO, respectively (Figures S3F–S3H). We then combined in vitro HIOs, developing and adult duodenum, and classified the ISC state transition process from intestinal epithelial progenitor to mature ISC into 6 phases based on the expression pattern of the ISC development-associated genes.

**Figure 5. Reference tissue and organoid atlases reveal epithelial progenitor and stem cell states during small intestine development**

(A) Integrated UMAP of thIO, developing, and adult duodenum epithelial scRNA-seq datasets, colored by cell type (left) and source (right).

(B) Hierarchical clustering of average transcriptome correlations between cell types of different tissue sources.

(C) Inset of the 3D UMAP, highlighting stem cells colored by cell type (left), source (center), and stem cell score (right).

(D) Boxplots showing the Spearman’s correlation distribution of thIO stem cells compared with developing or adult stem cells (Wilcoxon rank-sum test, nominal ***p < 0.0001).

(E) Schematic of scRNA-seq experiments performed over a time course of in vitro HIO development. SPRING embedding is colored by cell class and shaded by cluster assignment.

(F) HIO time course UMAP colored by time point or marker gene expression. Light purple indicates intestinal epithelial cells.

(G) In vitro HIO, developing, and adult duodenum stem cells ordered by ISC state and colored by phase. Hierarchical clustering of ISC cell bins represents transcriptome distance calculated with ISC development-associated genes. The heatmap shows gene expression changes of signaling pathways during ISC phase transitions.
(8–19 PCWs and adult) are more mature. LGR5 and ASCL2 expression was robustly detected starting at 8 PCWs (phase 4), whereas OLFM4 expression was prominent from 12 PCWs on (phase 5). Although the 12–19 PCW developing duodenum and thIO ISCs co-expressed LGR5 and OLFM4, adult ISCs (phase 6) were dominated by LGR5-low/OLFMA+ populations. Several fibroblast growth factor (FGF) signaling pathway components showed differential expression (DE) between ISC phase 4 and phase 5. For example, although the FGF downstream target gene MYC shows comparable enrichment in phases 4 and 5, another target gene, FOS, is more enriched in phase 5. We also noticed that the Notch signaling pathway target genes NOTCH1, NRARP, and OLFM4 are upregulated in phase 6. We identified genes enriched in each ISC phase as well as differentially expressed genes between developing and adult stem cells (Table S3). To explore ISC maturation regulation, we incorporated primary duodenum ISC expression data and TF binding site predictions (Aibar et al., 2017) to infer genes that are commonly regulated (i.e., regulators). We focused on regulators showing enriched expression in phases 4–6 and found that JUND and MYC are associated with acquisition of LGR5+ ISCs, whereas FOSB and NR1H3 are associated with transition to OLFM4+ ISCs (Figure S3I). These data show that integrating HIO and primary tissue data enables temporal reconstruction of ISC development from early to mature stages.

Integrating HIO and primary duodenum data unveils transient cell states during human intestine mesenchyme development

Mesenchymal-epithelial signaling regulates the emergence of crypt-villus architecture during development and maintains crypt homeostasis after villus formation (Holloway et al., 2021; Kinchen et al., 2018; McCarthy et al., 2020a; Santos et al., 2018; Shoshkes-Carmel et al., 2018; Smillie et al., 2019). To study intestinal mesenchyme development, we first annotated cell clusters in the developing human duodenal mesenchyme (Figure S4A; Data S1). At 7 PCWs, we observed two groups of precursors that showed different expression levels of SFRP1, a WNT pathway antagonist (Data S1). Along development, proportions of the SFRP1-high cell types (c3, c4, and c5) decreased dramatically, whereas the SFRP1-low cell types expanded and diverged into at least 7 subpopulations (c7, c8, c9, c11, c12, c13, and c15), exhibiting differential developmental dynamics and expression patterns of the bone morphogenetic protein (BMP), NOTCH, retinoic acid (RA), WNT, and R-Spondin (RSPO) signaling genes (Figures S4B–S4D; Data S1). Notably, we observed a correlation between ISC state transition and mesenchymal cell type diversification. Although emergence of LGR5+/OLFMA- ISC coincides with expansion of subepithelial mesenchyme (c15), emergence of LGR5+/OLFMA+ ISC coincides with expansion of villus core (c13), GDF10+ (c8), and PITX1+ smooth muscle cell (SMC, c7) (Figures S5G, S4C, S4D, and S4H). Of note, GDF10+ (c8) and subepithelial mesenchyme (c15) expressed the important stem cell niche factors GREM1 (McCarthy et al., 2020b) and NRG1 (Holloway et al., 2021; Jardé et al., 2020), respectively (Data S1).

To understand how mesenchymal cell clusters relate to primary tissues, we integrated thIO and primary developing duodenal mesenchyme (Figure S4E). Subtype counterparts overlap in an integrated UMAP (Figure S4E), and hierarchical clustering further shows that VSMC (c2), ICC (c4), subepithelial (c15), and proliferative mesenchyme (c10) cells in thIO and the developing intestine co-cluster, revealing high transcriptome similarity between these organoid and primary mesenchyme counterparts (Data S1). Fluorescence in situ hybridization validated subepithelial localization of the F3+/DLL1+ mesenchymal subtype (c15) in thIO and the developing duodenum (Data S1). These data showed that thIO recapitulates diverse mesenchymal subtypes observed in developing duodenum.

Next we used the in vitro HIO time course scRNA-seq data to understand the dynamics of intestinal mesenchyme development and interactions with epithelium. We observed a trajectory from mesoderm progenitors marked by CDX2, HAND1, and MKI67 (Mendjan et al., 2014) into multiple intestinal mesenchymal populations, including those resembling cell types most prominent at 7 PCW (M1, M2, and M4), cell types expanded in developing duodenum (M3 and M5), and foregut-associated PRRX1+ mesenchyme (Figures S4F and S4G; Data S1). We combined the in vitro HIO and developing duodenum data to reconstruct mesenchymal-epithelial cell state interaction dynamics over development. We highlighted age-dependent patterns of interacting gene pairs as well as aggregated signaling pathway component expression (Figure S4I; Data S1). We observed increasing epithelium-mesenchyme interaction intensity along development (Figure S4J; Data S1) but also identified early-stage developing duodenum-enriched interactions. For example, the Hedgehog signaling pathway, exemplified by IHH/SHH-PTCH1 epithelial-mesenchymal interactions, becomes active during intestinal specification along the HIO time course and reaches a peak at early time points in the developing duodenum (7–9 PCWs). We also observed diversification of signaling among mesenchymal subtypes. For instance, NRG1-ERBB3 interaction and expression of BMP ligands are enriched in subepithelial mesenchyme, whereas RA ligands are enriched in the villus core, and RSPOs are enriched in GDF10+ mesenchyme and PITX1+ smooth muscle cells (Figure S4J). Combining intestinal organoids and developing duodenum enables detailed characterization of epithelium-mesenchyme developmental dynamics and interactions.

NRG1 promotes intestinal epithelial stem cell maturation in vitro from pluripotency

Given co-occurrence of the transition from LGR5+/ASCL2 to LGR5+/ASCL2+ ISCs and expansion of subepithelial mesenchyme at 8 PCW, we hypothesized an association between subepithelial mesenchyme and ISC maturation. NRG1, a ligand highly enriched in subepithelial mesenchyme, has been reported to be an important stem cell niche factor (Holloway et al., 2021; Jardé et al., 2020). To further investigate the regulatory role of NRG1 in intestine development, we compared epidermal growth factor (EGF), NOG, and RSPO grown in in vitro HI Os (ENR; 100E-0N) with those supplemented with NRG1 instead of EGF (0E-100N). We also generated organoids with no EGF or NRG1 (0E-0N) and with varying concentrations of EGF and NRG1 in combination (1E-100N and 100E-1N; Figures 6A, 6B, and S5A). In epithelium, we noticed a cluster enriched for
Resource
LGR5+/ASCL2+ cells and derived predominantly from organoids grown with high NRG1 concentrations (0E-100N and 1E-100N; Figure 6C). To determine the organ identity of each epithelial cell type, we projected the cells to the developing multi-organ atlas and inspected expression patterns of individual marker genes of each organ. We observed differences in the organ identity of epithelium grown in different media (Figures 6D, 6E, and S5B–S5E). Non-proliferative epithelial cell types of organoids grown without EGF/NRG1 (NOG and RSPO only, 0E-0N) were mostly lung/esophagus-like (92%), expressing SOX2 and basal cell markers (e.g., KRT4 and TP63) or the ciliated cell marker FOXJ1. HIOs grown with NRG1 also had lung/esophagus-like cell types (70%, 0E-100N; 26%, 1E-100N), but more cell types under this condition possessed GI transcriptional features (CDX2+, MUC5AC+, or CLDN18+; 22% [0E-100N] and 60% [1E-100N] are stomach like; 8% [0E-100N] and 14% [1E-100N] are intestine like). In contrast, HIOs grown with high concentrations of EGF had the highest proportion of intestinal epithelial cell types (CDX2+/CDX1+) compared with other conditions (29%, 100E-0N; 36%, 100E-1N) and few lung/esophagus-like cell types (9%, 100E-0N; 7%, 100E-1N) (Figures 6D and 6E). These results suggest that early patterning events are different depending on whether EGF or NRG1 is present in the culture, with EGF acting as a potent driver of the intestinal lineage. On the other hand, although NRG1 conditions appeared to have more off-target epithelial cell types, the intestinal epithelium in these cultures had higher LGR5+/ASCL2+ cells, which was associated with a more mature epithelium. To assess maturity, we computationally extracted intestinal epithelial cells and determined a “maturity score” based on genes associated with stem cell maturation (STAR Methods). Cells grown with a high concentration of NRG1 (1E-100N and 1E-100N) showed higher maturity scores than others and were comparable with the 4-week tHIOs and 8 PCW developing duodenum epithelium (Figure 6F). The ISC markers LGR5 and ASCL2 showed increased expression in NRG1-grown organoids (Figures 6G and S5G–S5I). We extracted intestinal mesenchymal clusters (c1, c2, c3, c7, and c9) from each condition and compared their transcriptome with developing duodenum mesenchymal subtypes. We found that mesenchyme in NRG1-grown HIOs showed higher similarity to the duodenum reference than the EGF-grown counterparts (Figure 6H). The only exception was a cluster (c2) marked by NRG1/DLL1/F3, which showed the highest similarity to the duodenum subepithelial mesenchyme in HIOs grown in high EGF concentrations (Figure 6H). We also noted that EGF-grown HIOs have a higher proportion of NRG1+ subepithelial-like mesenchyme (Figure 6I). The data suggest that there are likely roles for specific EGF family members at different times during development. They suggest that NRG1 facilitates maturation of intestinal epithelial cells and mesenchyme but is less potent than EGF in driving intestinal identity acquisition (Figure 6J).

Next we analyzed HIOs embedded in different 3D matrices to study the effect of extracellular matrix (ECM) on HIO development. We performed scRNA-seq on 4-week HIOs embedded in alginate (Capeling et al., 2019) or Matrigel (Hughes et al., 2010) grown with ENR or EGF-only medium. Using the developing atlas as a reference, we showed that the proportions of intestinal epithelial cells and ISC maturity in alginat-grown and EGF-only HIOs are comparable with or slightly better than those of the Matrigel-embedded, ENR-medium-cultured HIOs of the same experimental batch (Figure S6). These data support previous work demonstrating that alginates is a viable ECM substitute for Matrigel for generating high-fidelity HIOs (Capeling et al., 2019). These data suggest that combinatorial signals derived from the ECM shape intestinal epithelial stem cell fate acquisition and maturation, and HIOs are powerful systems to explore these mechanisms.

**CDX2 deletion leads to loss of intestinal cell fate and gain of foregut features in epithelium and mesenchyme**

We next utilized HIOs to understand the gene-regulatory mechanisms that lead to the specification of the human intestinal epithelium and mesenchymal cell types. We used the HIO time course scRNA-seq data to construct a TF coexpression network...
and identified TFs that distinguish progenitor, epithelial, and mesenchymal cells (Data S1). This analysis revealed differentially expressed TF modules that might drive the diversification of epithelium and mesenchyme from early progenitors. Notably, we find that CDX2, a master regulator of the intestinal epithelium (Gao et al., 2009; Grainger et al., 2010; Kumar et al., 2019; Silberg, 2000; Stringer et al., 2012; Verzi et al., 2010) and regulator of mesoderm cell fate specification (Bernardo et al., 2011; Edgar et al., 2001; Ehrman and Yutzey, 2001; Faas and Isaacs, 2009; Gaunt et al., 2008; Mendjan et al., 2014; van den Akker et al., 2002) is positioned with genes that are coexpressed in early progenitors (Data S1). CDX2 is highly expressed in HIO epithelial cells, and its expression is maintained in the developing small intestine epithelium into adulthood. Interestingly, CDX2 is also expressed in a subset of early mesenchymal progenitors in the HIO (c5 and c9), as well as mesoderm/mesenchyme populations in the mouse gastrulation atlas (Data S1). To further explore the regulatory role of CDX2 on mesoderm patterning, we used time-course HIO scRNA-seq data and TF binding sites to infer regulators in mesenchyme and epithelium, respectively (Data S1). We found that CDX2 is predicted to regulate several HOX genes and intestine-enriched TFs such as HAND1 and NKX2-3 in mesenchyme (Figure S7A). In the epithelium, CDX2 is predicted to regulate the Hedgehog ligand IHH (Figure S7A), whose corresponding receptor PTCH1 is broadly expressed in developing mesenchyme. These results suggest that CDX2 could influence intestinal mesenchymal cell fates via an intrinsic regulatory network in mesoderm and through Hedgehog signaling between epithelium and mesenchyme.

To test the requirement of CDX2 to specify gut epithelium and mesenchymal fate in humans, we generated HIOs derived from hPSCs harboring homozygous CDX2 knockout (KO) and the corresponding control line (Kumar et al., 2019) and compared their sc transcriptomes (Figure 7A). Immunofluorescence (IF) and scRNA-seq data supported loss of CDX2 expression in CDX2 KO organoids (Figures 7B and S7C). Consistent with previous reports (Gao et al., 2009; Kumar et al., 2019; McCracken et al., 2014), CDX2 KO organoids predominantly exhibited a cystic morphology of folded and glandular epithelium (Figures 7B and S7B) and were positive for the foregut markers SOX2 and MUC5AC (Figure 7B). Comparisons of control and CDX2 KO organoids revealed cell type composition differences, with multiple epithelial cell populations abundant in CDX2 KO organoids, such as MUC5AC+ (c7), MUC16+ (c11), and SPPI+ (c10) cells (Figures 7B–7D, S7C, and S7D; Table S2). CDX2 KO organoid epithelial cells mapped predominantly to the developing stomach epithelium (Figure 7E). These data indicate loss of intestine and expansion of stomach identities in CDX2 KO intestinal organoid epithelium.

Next we performed DE analysis of control and CDX2 KO mesenchyme (Figures 7F, S7D, and S7E; Table S2) and examined the expression patterns of DE genes in HIO and reference atlases (Figures 7F–7I and S7F–S7J). Genes with higher expression in CDX2 KO mesenchyme showed enriched expression (e.g., FOXC1, PRRX1, SOX9, and SIX1) in lung subtypes (Figures 7F–7I). SIX1 and PRRX1 are required for lung morphogenesis (El-Hashash et al., 2011; Lu et al., 2013) and vascular development (Ihida-Stansbury et al., 2015; Ihida-Stansbury et al., 2004), respectively. We also observed loss of expression of the ubiquitous intestine mesenchymal TF NKX2-3 in CDX2 KO mesenchyme (Figures 7F and 7I). A kNN (k = 20) search in the developing endoderm cell atlas revealed that CDX2 KO mesenchyme tends to be mapped to the developing lung compared with the control (Figure 7G; Fisher’s exact test, nominal p < 0.0001, odds ratio = 25.23). These data suggest that homozygous CDX2 KO in hPSCs results in loss of intestine identity and gain of lung mesenchymal features in HIOs. We noted that the predicted CDX2 target genes in mesenchyme and epithelium, including HAND1, NKX2-3, and IHH, were downregulated in CDX2 KO organoids. We show that CDX2 is critical for epithelial and mesenchymal cell fate acquisition during human intestine development and provide detailed analyses pinpointing possible mechanisms controlling epithelial and mesenchymal specification and interactions in the developing human intestine.

**DISCUSSION**

Specialized epithelial stem cell populations derived from the endoderm are crucial during development and maintain the capacity to differentiate into epithelial cell types with organ-specific features. We provide a resource of expression features that are specific to stem cells from different developing human organs that can be used to study how human stem cells are established and maintained. In addition, mesenchyme is a major component of stem cell niches, providing structural and biochemical support. Cell type annotation, receptor-ligand pairing, and cell subtype-specific gene expression revealed a rich diversity of mesenchymal populations across each organ and tissue, much of which was previously unexplored. Interactions we identified may coordinate niche-specific physiology within each tissue and lay the groundwork for studies using endoderm-derived human organoid model systems (Dye et al., 2015; McCracken et al., 2014, 2017; Shacham-Silverberg and Wells, 2020; Trisno et al., 2018). There are time points when it is difficult to profile human tissues and capturing changes across human development is challenging. Toward this end, we established a paradigm to compare human organoids with a high-dimensional reference atlas and rigorously assess organoid cell composition over time.

Focusing on HIOs, we identify off-target cells in the culture system, which could affect interpretation of experiments that lack sc resolution. The majority of cells in *in vitro* and transplanted HIOs are specified to intestinal fates, with transplanted HIOs strongly resembling the developing human intestine, and the most mature HIO cells not reaching adult maturation status. This observation provides opportunities to understand the mechanisms that underlie the maturation of human intestinal cell types. Furthermore, we used HIOs to cover early stages of human intestine development and explore molecular transitions up to adulthood. Regulatory, signaling, and metabolic features could be targeted to enhance HIO stem cell maturation *in vitro* by changing the chemical or physical environment through TF overexpression (Miura and Suzuki, 2017), CRISPR inhibition/activation (Campa et al., 2019; Wang et al., 2019), or small-molecule modulation. To highlight this idea, we showed that introduction of NRG1, a growth factor produced by subepithelial intestinal mesenchyme, into HIO culture medium induces ISC maturation *in vitro*. We also observed increased non-intestinal...
Figure 7. CDX2 deletion leads to loss of intestinal cell fate and gain of foregut features in epithelium and mesenchyme

(A) Top: boxplots showing expression of CDX2 in intestinal epithelium and mesenchyme of in vitro HIOs on each day of sampling. Bottom: schematic of CDX2 knockout (KO) effects on HIO cell fate. Epithelial and mesenchymal compartments are colored according to inferred organ identity.

(B) Left: bright-field (BF) images showing the morphology of control (top) and CDX2 KO (bottom) organoids. Dashed curves in the control organoid image show protrusions of crypt-like structures. The dashed square in the magnified control organoid image show tightly packed and polarized epithelial organization. The dashed square in the magnified CDX2 KO organoid image show folded glandular epithelial organization (white arrow). Right: IF staining of organ-enriched epithelial markers (intestine, CDX2; proximal small intestine and stomach antrum, PDX1; foregut, SOX2 and MUC5AC). Scale bars, 200 μm in BF images; 20 μm or 50 μm in IF images.

(C) Integrated UMAP of control and CDX2 KO organoid cells, colored by major cell class, with shade representing cluster assignment (top; epithelium, red; mesenchyme, green; neuron, pink) or condition or CDX2 expression (bottom left). Bottom right: boxplot showing CDX2 expression levels in control and CDX2 KO organoids.

(D) Left bar plot: cell numbers. Right stacked bar plot: condition proportions. Dot plot: cluster marker gene expression.

(E) Top: stacked bar plot showing the projected organ proportion of epithelial cells under each condition (left) or cluster (right). The sidebar shows CDX2 expression. Bottom: boxplot showing log-transformed fold change of control and KO epithelium differentially expressed genes in the developing intestine versus stomach.

(F) Feature plots of differentially expressed TFs between mesenchymal cells of control and CDX2-KO organoids.

(G) Top: UMAP of multi-organ mesenchymal cells, colored by organs. Bottom: stacked bar plot showing the projected organ proportion of mesenchymal cells under each condition.

(H and I) Mesenchyme UMAP colored by CDX2 KO high gene score in each developing organ (H) and expression of differentially expressed TFs between control and CDX2 KO mesenchyme (I).
cell types, indicating that refinement of differentiation approaches is required to efficiently generate mature intestinal organoids. We also note that multiple cell types observed in the developing intestine are underrepresented or absent in HIOs. Alterations of the HIO culture system that enable co-differentiation or co-culture of the endothelium (Holloway et al., 2020), neuron (Workman et al., 2017), immune, or other lineages are possible approaches to achieve a more complete model of human intestine development. Furthermore, maturation may also require perfusable vasculature (Homan et al., 2019; Sidar et al., 2019), ECM alterations (Capeling et al., 2019), interaction with the microbiome (Hill et al., 2017), or other mechanical inductive cues (Poling et al., 2018). Achieving mature intestinal tissue from hPSCs in vitro remains a major challenge, and continued benchmarking against multi-organ reference atlases will be required to quantify the precision of cell state specification.

Currently, very little is known about human intestinal mesenchyme development, but insight into this group of cells is critical for understanding congenital diseases that affect the intestine (Le et al., 2021; Matera et al., 2021; O’Connell et al., 2018; Zhang et al., 2021). It has been shown that subepithelial mesenchyme promotes ISC maturation through epithelial-mesenchymal interactions (Aoki et al., 2016; Kosinski et al., 2010; Szepourginski et al., 2017); however, the signaling, transcriptional networks, and cell-cell communication that regulate human intestinal mesenchyme development and differentiation are only starting to be understood from sc transcriptomic studies (Elmentaitė et al., 2020; Holloway et al., 2021). Notably, cell types within the HIO have in vivo counterparts, and we identify early precursor states in the in vitro HIO. This further highlights the potential of HIOs for interrogating the gene-regulatory logic controlling the emergence of diverse intestinal mesenchyme cell types and for understanding how perturbations of intestinal mesenchyme development may be involved in the developmental process.

The finding that CDX2 is critical for human intestinal epithelial differentiation is consistent with previous reports in mice. Cdx2 deletion in the intestinal epithelium of mice results in homeotic transformation of the intestine into an esophagus-like tube (Gao et al., 2009) or ectopic expression of stomach markers (Grainger et al., 2010), depending on the timing of Cdx2 loss. Consistent with previous studies showing that Cdx2 plays a role in mesoderm posteriorization (Bernardo et al., 2011; Edgar et al., 2001; Ehman and Yutzey, 2001; Faas and Isaacs, 2009; Gaunt et al., 2008; Mendjan et al., 2014; van den Akker et al., 2002), we show that, during HIO development, CDX2 KO mesenchyme loses the intestine identity and acquires a lung-associated mesenchyme identity. Analyses of HIO developmental trajectories suggest that CDX2 is a transient patterning regulator during the early phases of gut-associated mesenchyme development in humans. In addition, the observation that CDX2 KO induced fate transition toward a stomach-like identity in the epithelium but a lung-like identity in mesenchyme indicates that establishment of a regional identity in epithelium and mesenchyme might be decoupled in the absence of CDX2.

Our data and analyses provide a rich resource for understanding human developmental biology and create a framework for integrating complex primary tissue and organoid datasets. We provide highly resolved evidence showing that human organoids can be predictive of human development and that perturbations in organoids can reveal mechanisms of human development. Because of the cellular heterogeneity and presence of off-target cell types in organoids, it will be important to use sc genomic analyses of organoid models to appropriately resolve associated phenotypes. Our data point to the value of high-dimensional human reference cell atlases across development, adulthood, and disease to comprehensively benchmark engineered human tissue culture systems, and such comparisons could be used to modify organoid protocols for improved modeling of development and disease.

**Limitations of study**

Data from diverse genetic backgrounds, time points, and tissues will be required to comprehensively annotate human developmental cell states. Considering the variability of organoids, more organoids derived from different cell lines, batches, and protocols would be necessary to obtain a more robust understanding of HIO technologies. Cell-cell interactions were inferred based on ligand-receptor gene pair expression without comprehensive knowledge of location within the tissue. Therefore, high-resolution spatial transcriptomic or proteomic data will further enhance our understanding of ligand-receptor interactions in developing organs. Finally, many of our analyses were focused on epithelial and mesenchymal cells, and further characterization is required to understand the developmental and lineage dynamics across all abundant and rare cell types and states.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and Code Availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Human samples
  - Human embryonic stem cells
  - Immuno-compromised mice for transplantation
- **METHOD DETAILS**
  - Generation and culture of in vitro human intestinal organoids
  - Human intestinal organoid (HIO) transplantation
  - Chromogenic in Situ Hybridization and Multiplex Fluorescent in Situ Hybridization (FISH)
  - Tissue staining
  - Single-cell dissociation and RNA sequencing library preparation of the developing human tissues
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Alignment of single-cell transcriptomes
  - Analysis of single-cell RNA-seq data
  - Quantification of organ specificity
  - Identification of stem cell clusters, stemness genes shared by multiple organs, and functional enrichment analysis
SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2021.04.028.

ACKNOWLEDGMENTS

We thank the Camp and Treutlein labs for helpful discussions, Marc Perea and Mendurim Rahshiti of IOB-IT for support hosting the GutTubeR data browser, and D-BSSE IT Services for computational support. We thank Judy Opp and the University of Michigan Advanced Genomics Core for expertise regarding operation of the 10X Chromium single-cell capture platform and sequencing. We would also like to thank the University of Washington Laboratory of Developmental Biology staff for supporting this project. J.G.C., J.R.S., and B.T. are supported by grant CZF2019-002440 from the Chan Zuckerberg Initiative. E.M.H. was supported by the Training Program in Organogenesis (NIH-NICHD 5T32 HD007505). Additional support was provided by the University of Michigan Center for Gastrointestinal Research (UMCGR) (NIDDK 5P30DK034933).

AUTHOR CONTRIBUTIONS

U.K., E.M.H., Y.-H.T., M.M.C., and S.H. differentiated organoids used in this study. Y.-H.T. and M.C. performed organoid transplantation experiments. U.K., A.W., Y.-H.T., M.C., and E.M.H. collected, dissociated, and submitted tissue for scRNA-seq. Q.Y. and Z.H. performed scRNA-seq data analyses. J.H.W. and Q.Y. maintained the scRNA-seq data. Y.-H.T., M.M.C., and U.K. performed immunofluorescence staining. E.M.H., Y.-H.T., and C.J.C. performed in situ hybridization and imaging. C.H. developed the GutTubeR data browser. I.A.G. and P.D.R.H. provided critical material for this work. Q.Y., U.K., E.M.H., B.T., J.R.S., and J.G.C. designed the study and wrote the manuscript. All authors read and approved the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 23, 2020
Revised: February 11, 2021
Accepted: April 16, 2021
Published: May 20, 2021

REFERENCES

Albar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulsemans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat. Methods 14, 1083–1086.
Angerer, P., Haghverdi, L., Büttner, M., Theis, F.J., Marr, C., and Buettner, F. (2016). destiny: diffusion maps for large-scale single-cell data in R. Bioinformatics 32, 1241–1243.
Aoki, R., Shoshkes-Carmel, M., Gao, N., Shin, S., May, C.L., Golson, M.L., Zahm, A.M., Ray, M., Wiser, C.L., Wright, C.V., and Kaestner, K.H. (2016). Fox1-expressing mesenchymal cells constitute the intestinal stem cell niche. Cell. Mol. Gastroenterol. Hepatol. 2, 175–188.
Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Hagegebarth, A., Korving, J., Bergthel, H., Peters, P.J., and Clevers, H. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 449, 1003–1007.
Becht, E., McInnes, I., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. 37, 38–44.
Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M., Trotter, M.W., Hemberger, M., Smith, J.C., et al. (2011). BRA-CHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. Cell Stem Cell 9, 144–155.
Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 36, 411–420.
Camp, J.G., Platt, R., and Treutlein, B. (2019). Mapping human cell phenotypes to genotypes with single-cell genomics. Science 365, 1401–1405.
Campa, C.C., Weisbach, N.R., Santinini, A.J., Incarnato, D., and Platt, R.J. (2019). Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. Nat. Methods 16, 887–893.
Cao, J., O’Day, D.R., Pliner, H.A., Kingsley, P.D., Deng, M., Daza, R.M., Zager, M.A., Aldinger, K.A., Blecher-Gonen, R., Zhang, F., et al. (2020). A human cell atlas of fetal gene expression. Science 370, eaba7721.
Capeling, M.M., Czerwinski, M., Huang, S., Tsai, Y.H., Wu, A., Nagy, M.S., Juliar, B., Sundaram, N., Song, Y., Han, W.M., et al. (2019). Nonadaptive Alginates
Hydrogels Support Growth of Pluripotent Stem Cell-Derived Intestinal Organoids. Stem Cell Reports 12, 381–394.

Capeling, M., Huang, S., Mulero-Russe, A., Ciezca, R., Tsai, Y.H., Garcia, A., and Hill, D.R. (2020). Generation of small intestinal organoids for experimental intestinal physiology. Methods Cell Biol. 159, 143–174.

Chang, H.Y., Chi, J.T., Dudoit, S., Bondre, C., van de Rijn, M., Botstein, D., and Brown, P.O. (2002). Diversity, topographic differentiation, and positional memory in human fibroblasts. Proc. Natl. Acad. Sci. USA 99, 12877–12882.

Citri, A., and Yarden, Y. (2006). EGF-ERBB signaling: towards the systems level. Nat. Rev. Mol. Cell Biol. 7, 505–516.

Cowan, C.S., Renner, M., De Gennaro, M., Gross-Scherf, B., Goldblum, D., Hou, Y., Munz, M., Rodrigues, T.M., Krol, J., Szirka, T., et al. (2020). Cell Types of the Human Retina and Its Organoids at Single-Cell Resolution. Cell 172, 1623–1640.e34.

Du, A., McCracken, K.W., Walp, E.R., Terry, N.A., Klein, T.J., Han, A., Wells, J.M., and May, C.L. (2012). Arx is required for normal enteroendocrine cell development in mice and humans. Dev. Biol. 365, 175–188.

Dusing, M.R., Maer, E.A., Aronow, B.J., and Wignaton, D.A. (2010). One-cut knockout mice fail to thrive during early postnatal period and have altered patterns of gene expression in small intestine. Physiol. Genomics 42, 115–125.

Dye, B.R., Hill, D.R., Ferguson, M.A., Tsai, Y.H., Nagy, M.S., Dyal, R., Wells, J.M., Mayshe, C.N., Nattiv, R., Klein, O.D., et al. (2015). In vitro generation of human pluripotent stem cell derived lung organoids. eLife 4, e05098.

Edgar, L.G., Carr, S., Wang, H., and Wood, W.B. (2001). Zygotic expression of the caudal homolog pat-1 is required for posterior patterning in Caenorhabditis elegans embryogenesis. Dev. Biol. 229, 71–88.

Efremova, M., Vento-Tormo, M., Reichmann, S.A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat. Protoc. 15, 1484–1506.

Ehrman, L.A., and Yutzey, K.E. (2001). Anterior expression of the caudal homologue CdxB activates a posterior genetic program in avian embryos. Dev. Dyn. 221, 412–421.

El-Hashash, A.H., Al Alam, D., Turcategel, G., Rogers, O., Li, X., Bellusci, S., and Warburton, D. (2011). Six1 transcription factor is critical for coordination of epithelial, mesenchymal and morphogenetic processes in the mammalian lung. Dev. Biol. 353, 242–258.

Elmentaite, R., Ross, A.D.B., Roberts, K., James, K.R., Ortmann, D., Gomes, T., Nakay, K., Tuck, L., Pritchard, S., Bayraktar, O.A., et al. (2020). Single-Cell Sequencing of Developing Human Gut Reveals Transcriptional Links to Childhood Crohn’s Disease. Dev. Cell 55, 771–783.e5.

Faas, L., and Isaacs, H.V. (2009). Overlapping functions of Cdx1, Cdx2, and Cdx4 in the development of the amphibian Xenopus tropicalis. Devon. Dyn. 238, 835–852.

Fawknor-Corbett, D., Antanaviciute, A., Parikh, K., Jagielowicz, M., Gerós, A.S., Gupta, T., Ashley, N., Khamis, D., Fowler, D., Morrissey, E., et al. (2021). Spatiotemporal analysis of human intestinal development at single-cell resolution. Cell 184, 810–826.e23.

Finkbeiner, S.R., Zeng, X.L., Utama, B., Atmar, R.L., Shroyer, N.F., and Estes, M.C. (2010). ER stress causes rapid loss of intestinal epithelial stemness and epithelial-mesenchymal signaling by Cdx2. Dev. Cell 23, 787–793.e6.

Fujii, M., Matano, M., Toshimitsu, K., Takano, A., Mikami, Y., Nishikori, S., Sugimoto, S., and Sato, T. (2018). Human Intestinal Organoids Maintain Self-Renewal Capacity and Cellular Diversity in Niche-Inspired Culture Condition. Cell Stem Cell 23, 382–397.e3.

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

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

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

Han, L., Chaturvedi, P., Kishimoto, K., Koike, H., Nasr, T., Iwasawa, K., Giesbrecht, K., Witcher, P.C., Eicher, A., Haines, L., et al. (2020a). Single cell transcriptomics identifies a signaling network coordinating endoderm and mesoderm diversification during foregut organogenesis. Nat. Commun. 11, 4158.

Han, X., Zhou, Z., Fei, L., Sun, H., Wang, R., Chen, Y., Chen, H., Wang, J., Tang, H., Ge, W., et al. (2020b). Construction of a human cell landscape at single-cell level. Nature 581, 303–309.

He, Z., Brazovskaja, A., Ebert, S., Camp, J.G., and Treutlein, B. (2020). CSS-cluster similarity spectrum integration of single-cell genomics data. Genome Biol. 21, 257.

Heijmans, J., van Lidth de Jeude, J.F., Koek, B.K., Rosekrans, S.L., Wielenga, I.M., van de Wetering, M., Ferrante, M., Lee, A.S., Onderwater, J.J., Paton, J.C., et al. (2013). ER stress causes rapid loss of intestinal epithelial stemness through activation of the unfolded protein response. Cell Rep. 3, 1128–1139.

Hill, D.R., Huang, S., Nagy, M.S., Yadagiri, V.K., Fields, C., Mukherjee, D., Bons, B., Dedhia, P.H., Chin, A.M., Tsai, Y.H., et al. (2017). Bacterial colonization stimulates a complex physiological response in the immature human intestinal epithelium. eLife 6, e29132.

Holloway, E.M., Wu, J.H., Czerwiniski, M., Sweet, C.W., Wu, A., Tsai, Y.H., Huang, S., Stoddard, A.E., Capeling, M.M., Glass, I., and Spence, J.R. (2020). Differentiation of Human Intestinal Organoids with Endogenous Vascular Endothelial Cells. Dev. Cell 54, 516–528.e7.

Homan, K.A., Gupta, N., Kroll, K.T., Kolesky, D.B., Skylar-Scott, M., Miyoshi, T., Mau, D., Valerius, M.T., Ferrante, T., Bonventre, J.V., et al. (2019). Flow-enhanced vasculization and maturation of kidney organoids in vitro. Nat. Methods 16, 255–262.

Hu, H., Miao, Y.R., Jia, L.H., Yu, Q.Y., Zhang, Q., and Guo, A.Y. (2019). AnimalFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. Nucleic Acids Res. 47 (D1), D33–D38.

Hughes, C.S., Postovit, L.M., and Lajoie, G.A. (2010). Matrigel: a complex protein mixture required for optimal growth of cell culture. Proteomics 10, 1886–1890.

Ihida-Stansbury, K., McKeen, D.M., Gebb, S.A., Martin, J.F., Stevens, T., Nenomoff, R., Akeson, V., Vaughn, J., and Jones, P.L. (2004). Paired-related homeobox gene Px1 is required for pulmonary vascular development. Circ. Res. 94, 1507–1514.

Ihida-Stansbury, K., Ames, J., Chokshi, M., Aidi, N., Sanyal, S., Kawabata, K.C., Levental, I., Sundararaghavan, H.G., Burdick, J.A., Janmey, P., et al. (2015). Role played by Px1-dependent extracellular matrix properties in vascular smooth muscle development in embryonic lungs. Pulm. Circ. 5, 382–397.

Jardé, T., Chen, W.H., Rossello, F.J., Kaur Kahlon, T., Theocharous, M., Kurian Arackal, T., Flores, T., Giraud, M., Richards, E., Chan, E., et al. (2020). Mesenchymal Niche-Derived Neuregulin-1 Drives Intestinal Stem Cell Proliferation and Regeneration of Damaged Epithelium. Cell Stem Cell 27, 646–662.e7.
Josephson, R., Sykes, G., Liu, Y., Ording, C., Xu, W., Zeng, X., Shin, S., Loring, J., Maitra, A., Rao, M.S., and Auerbach, J.M. (2006). A molecular scheme for improved characterization of human embryonic stem cell lines. BMC Biol. 4, 28.

Kant, S., Boyle, M.J., He, Z., Santel, M., Weigert, A., Sanchis-Calleja, F., Guijarro, P., Sidow, L., Fleck, J.S., Han, D., et al. (2019). Organoid single-cell genomic atlas uncovers human-specific features of brain development. Nature 574, 418–422.

Kechele, D.O., and Wells, J.M. (2019). Recent advances in deriving human endodermal tissues from pluripotent stem cells.Curr. Opin. Cell Biol. 61, 92–100.

Kinchin, J., Chen, H.H., Parikh, K., Antanaviciute, A., Jagielowicz, M., Fawkner-Corbett, D., Ashley, N., Cubitt, L., Mellado-Gomez, E., Attar, M., et al. (2018). Structural Remodeling of the Human Colonic Mesenchyme in Inflammatory Bowel Disease. Cell 175, 372–386.e17.

Kosinski, C., Stange, D.E., Xu, C., Chan, A.S., Ho, C., Yuen, S.T., Mifflin, R.C., Powell, D.W., Clevers, H., Leung, S.Y., and Chen, X. (2010). Indian hedgehog regulates intestinal stem cell fate through epithelial–mesenchymal interactions during development. Gastroenterology 139, 893–903.

Kumar, N., Tsai, Y.H., Chen, L., Zhou, A., Banerjee, K.K., Saxena, M., Huang, S., Toke, N.H., Xing, J., Shivdasani, R.A., et al. (2019). The lineage-specific transcription factor CDX2 navigates dynamic chromatin to control distinct stages of intestine development. Development 146, dev172198.

Le, T.L., Galmiche, L., Levy, J., Suseemarat, P., Hellemackers, D.M., Morarak, K., Boismoreau, F., Theunissen, T.E., Lefebvre, M., Pelet, A., et al. (2021). Dysregulation of the NRG1-ERBB pathway causes a developmental disorder with gastrointestinal dysmotility in humans. J. Clin. Invest. 131, e145837.

Lemmon, M.A., Schlessinger, J., and Ferguson, K.M. (2014). The EGFR family: not so prototypical receptor tyrosine kinases. Cold Spring Harb. Perspect. Biol. 6, a020768.

Leslie, J.L., Huang, S., Opp, J.S., Nagy, M.S., Kobayashi, M., Young, V.B., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics 27, 1739–1740.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P., and Mesirov, J.P. (2011). An integrated encyclopedia of DNA elements in the human genome. Nature 477, 97–101.

Liu, K., Reddy, R., Berika, M., Warburton, D., and El-Hashash, A.H. (2013). Abrogation of Eya1/Six1 disrupts the saccular phase of lung morphogenesis and causes remodeling. Dev. Biol. 382, 110–123.

Matsuo, Y., Nishizaki, C., and Drexler, H.G. (1999). Efficient DNA fingerprinting of neoplasms by polymerase chain reaction using a single specific primer. Lab. Invest. 79, 1147–1155.

McCracken, K.W., Aihara, E., Martin, B., Crawford, C.M., Broda, T., Treguir, J., Zhang, X., Shannon, J.M., Montrose, M.H., and Wells, J.M. (2017). Wnt/β-catenin promotes gastric fundus specification in mice and humans. Nature 541, 182–187.

Mendjan, S., Mascetti, V.L., Ortmann, D., Ortiz, M., Karjoukoskaro, D.W., Ng, Y., Moreau, T., and Pedersen, R.A. (2014). NANOG and CDX2 pattern distinct subtypes of human mesoderm during exit from pluripotency. Cell Stem Cell 15, 310–325.

Miller, A.J., Hill, D.R., Nagy, M.S., Aoki, Y., Dye, B.R., Chin, A.M., Huang, S., Zhu, F., White, E.S., Lama, V., and Spence, J.R. (2018). In Vitro Induction and In Vivo Engraftment of Lung Bud Tip Progenitor Cells Derived from Human Pluripotent Stem Cells. Stem Cell Reports 10, 101–119.

Miller, A.J., Dye, B.R., Ferrer-Torres, D., Hill, D.R., Overeem, A.W., Shea, L.D., and Spence, J.R. (2019). Generation of lung organoids from human pluripotent stem cells in vitro. Nat. Protoc. 14, 518–540.

Moreau, M.A., Walker, T., Glass, I.A., Treutlein, B., et al. (2020). In Vitro and In Vivo Development of the Human Airway at Single-Cell Resolution. Dev. Cell 53, 117–128.e6.

Mura, S., and Suzuki, A. (2017). Generation of Mouse and Human Organoid-Forming Intestinal Progenitor Cells by Direct Lineage Reprogramming. Cell Stem Cell 21, 456–471.e6.

Nakanishi, S., Kato, T., Yoshida, S., Nomura, T., and Ito, S. (2010). Establishment of a murine trophoblast genetics screen. Mol. Cell. Biol. 30, 1059–1071.

Nitzan, M., Kairaikos, N., Friedman, N., and Rajewsky, N. (2019). Gene expression cartography. Nature 576, 132–137.

O’Connell, A.E., Zhou, F., Shah, M.S., Murphy, Q., Rickner, H., Kelsen, J., Boyle, J., Doyle, J.J., Gangwani, B., Thigharaja, J.R., et al. (2018). Neonatal-Onset Chronic Diarrhea Caused by Homozygous Nonsense WNT2B Mutations. Am. J. Hum. Genet. 103, 131–137.

Ootani, A., Li, X., Sangiorgi, E., Ho, Q.T., Ueno, H., Toda, S., Sugihara, H., Fujimoto, K., Weissman, I.L., Capecchi, M.R., and Kuo, C.J. (2009). Sustained expression cartography. Nature 460, 430–436.

Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Cellerio-Nieto, F.J., Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. Nature 566, 490–495.

Poling, H.M., Wu, D., Brown, N., Baker, M., Hausfeld, T.A., Huyhn, N., Chaffron, S., Dunn, J.C.Y., Hogan, S.P., Wells, J.M., et al. (2018). Mechanically induced development and maturation of human intestinal organoids in vivo. Nat. Biomed. Eng. 2, 429–439.

Rodansky, E.S., Johnson, L.A., Huang, S., Spence, J.R., and Higgins, P.D. (2015). Intestinal organoids: a model of intestinal fibrosis for evaluating anti-fibrotic drugs. Exp. Mol. Pathol. 98, 346–351.

Rossi, G., Manfrin, A., and Lutolf, M.P. (2018). Progress and potential in organoid research. Nat. Rev. Genet. 19, 671–687.

Santos, A.J.M., Lo, Y.H., Mah, A.T., and Kuo, C.J. (2018). The Intestinal Stem Cell Niche: Homeostasis and Adaptations. Trends Cell Biol. 28, 1062–1078.

Schließ, R., Fowler, K.L., Thornton, M., Huang, S., Hajiali, I., Hou, X., Grubbs, B., Spence, J.R., and Gkirskeft, T.C. (2017). Neural Crest Cell Implantation Restores Enteric Nervous System Function and Alters the Gastrointestinal Transcriptome in Human Tissue-Engineered Small Intestine. Stem Cell Reports 9, 883–896.
## STAR METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Antibodies**      |        |            |
| Rabbit anti-Ki67    | Thermo Scientific | Cat# RM-9106-S1 ; RRID: AB_149792 |
| Goat anti-LCT       | Santa Cruz Biotechnology | Cat# sc-240614, RRID: AB_10917595 |
| Goat anti-Sucrase-Isomaltase | Santa Cruz Biotechnology | Cat# sc-27603, RRID: AB_2188721 |
| Goat anti-Chr-A     | Santa Cruz Biotechnology | Cat# sc-1488, RRID: AB_2276319 |
| Rabbit anti-FABP2   | Abcam | Cat# ab89195, RRID: AB_2041190 |
| Rabbit anti-EPCAM   | Sigma-Aldrich | Cat# HP026761, RRID: AB_1848198 |
| Goat anti-Vimentin  | R&D Systems | Cat# AF2105, RRID: AB_355153 |
| Mouse anti-Mucin 5AC| Abcam | Cat# ab79082, RRID: AB_1603327 |
| Goat anti-DPP4      | R&D Systems | Cat# AF954, RRID: AB_355739 |
| Mouse anti-CDX2     | BioGenex | Cat# AM392, RRID: AB_2650531 |
| Rabbit anti-ASBT (anti-SLC10A2) | Sigma-Aldrich | Cat# HP004795, RRID: AB_1856953 |
| Rabbit anti-Mucin2  | Santa Cruz Biotechnology | Cat# sc-15334, RRID: AB_2146667 |
| Goat anti-SOX9      | R&D Systems | Cat# AF3075, RRID: AB_2194160 |
| Rabbit anti-SM22    | Abcam | Cat# ab14106, RRID: AB_443021 |
| Rabbit anti-PDX1    | Epitomics, Inc | Cat# 3470-1, RRID: AB_10703013 |
| Goat anti-E-Cadherin| R&D Systems | Cat# AF748, RRID: AB_355568 |
| Mouse anti-E-Cadherin| BD Transduction Laboratories | Cat# 610181, RRID: AB_397580 |
| Goat anti-Sox2      | Santa Cruz Biotechnology | Cat# sc-17320, RRID: AB_2286684 |
| Rabbit anti-PDGFRa  | Cell Signaling Technology | Cat#3164S, RRID: AB_2162351 |
| Mouse anti-SMA      | Sigma-Aldrich | Cat# C6198, RRID: AB_476856 |
| Donkey anti-goat 488| Jackson Immuno | Cat# 705-545-147, RRID: AB_2336933 |
| Donkey anti-goat 647| Jackson Immuno | Cat# 705-605-147, RRID: AB_2340437 |
| Donkey anti-goat Cy3| Jackson Immuno | Cat# 705-165-147, RRID: AB_2307351 |
| Donkey anti-mouse 488| Jackson Immuno | Cat# 715-545-150, RRID: AB_2340846 |
| Donkey anti-mouse 647| Jackson Immuno | Cat# 715-606-151, RRID: AB_2340866 |
| Donkey anti-mouse Cy3| Jackson Immuno | Cat# 715-165-150, RRID: AB_2340813 |
| Donkey anti-rabbit 488| Jackson Immuno | Cat# 711-545-152, RRID: AB_2313584 |
| Donkey anti-rabbit 647| Jackson Immuno | Cat# 711-605-152, RRID: AB_2492288 |
| Donkey anti-rabbit Cy3| Jackson Immuno | Cat# 711-165-152, RRID: AB_2307443 |
| **Biological samples** |        |            |
| Human fetal tissue samples, ages 7-21 post-conceptional weeks | University of Washington Laboratory of Developmental Biology | N/A |
| Human adult duodenum samples, age 45 and 50 years | University of Michigan Hospital | N/A |
| **Chemicals, peptides, and recombinant proteins** |        |            |
| Activin A            | R&D    | Cat#338-AC |
| EGF                  | R&D    | Cat#236-EG |
| Noggin-Fc            | Purified from conditioned media; Heijmans et al., 2013 | N/A |
| R-Spondin 1          | Conditioned media; Ootani et al., 2009 | N/A |
| FGF4                 | Purified in house; Sugawara et al., 2014 | N/A |
| CHIR99021            | ApexBio | Cat#A3011 |
| B27 supplement       | Life Technologies | Cat#1750404 |
| HEPES                | Life Technologies | Cat#15630080 |

(Continued on next page)
Continued

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| GlutaMAX            | Life Technologies | Cat#35050061 |
| Red Blood Cell Lysis Buffer | Roche | Cat#11814389001 |
| NRG1                | R&D    | Cat#5898-NR-050 |
| Alginic acid sodium salt, low viscosity | Alfa Aesar | Cat#11420828 |

Critical commercial assays

- RNAscope Multiplex Fluorescent Reagent Kit v2: ACD, Cat#323100
- RNAscope 2.5 HD Reagent Kit - Brown: ACD, Cat#322300
- Neural Tissue Dissociation Kit (P): Miltenyi, Cat#130-092-628

Deposited data

- Raw and analyzed data: This work, E-MTAB-10187, E-MTAB-10268
- Raw and analyzed data: Miller et al., 2020, E-MTAB-8221
- Raw and analyzed data: Holloway et al., 2021, E-MTAB-9489
- Raw and analyzed data: Holloway et al., 2020, E-MTAB-9228, E-MTAB-9363
- Raw and analyzed data: Pijuan-Sala et al., 2019, E-MTAB-6967
- Raw and analyzed data: Fuji et al., 2018, GSE119969

Experimental models: Cell lines

- H9-WT: WiCell, Cat#WA09; RRID:CVCL_9773
- H9-CDX2-CON: Derived from H9; Kumar et al., 2019, N/A
- H9-CDX2-KO: Derived from H9; Kumar et al., 2019, N/A

Oligonucleotides

- RNAscope Probe Hs-LGR5: ACD, Cat#311021
- RNAscope Probe Hs-OLFM4: ACD, Cat#311041
- RNAscope Probe Hs-DLL1: ACD, Cat#532631
- RNAscope Probe Hs-F3-C2: ACD, Cat#407611-C2
- RNAscope Probe Hs-WNT2: ACD, Cat#584071
- RNAscope Probe Hs-SALL1: ACD, Cat#514331
- RNAscope Probe Hs-SP5: ACD, Cat#406541
- RNAscope Probe Hs-NKX2.3: ACD, Cat#581651

Software and algorithms

- R (version 3.6.0): N/A, https://www.r-project.org
- Cell Ranger: N/A, https://github.com/10XGenomics/cellranger
- Seurat (version 3.1): Butler et al., 2018, https://github.com/satijalab/seurat
- simspec: He et al., 2020, https://github.com/quadbiolab/simspec
- SPRING: Weinreb et al., 2018, https://github.com/AllonKleinLab/SPRING
- e1071: N/A, https://github.com/rajewsky-lab/e1071
- presto: N/A, https://github.com/immunogenomics/presto
- uwot: N/A, https://github.com/lmelville/uwot
- RANN: N/A, https://github.com/jefferslab/RANN
- CellPhoneDB (version 2.0): Efremova et al., 2020, https://github.com/Teichlab/cellphonedb
- destiny: Angerer et al., 2016, https://github.com/theislab/destiny
- splines: N/A, https://github.com/cran/splines
- Quadprog: N/A, https://github.com/cran/quadprog
- novoSpaRc: Nitzan et al., 2019, https://github.com/rajewsky-lab/novosparc
- igraph: N/A, https://github.com/igraph/igraph

(Continued on next page)
RESOURCE AVAILABILITY

Lead contact
Further information and request for resources and reagents should be directed to and will be fulfilled by the Lead Contact, J. Gray Camp (grayson.camp@iob.ch).

Materials availability
This study did not generate any unique reagents.

Data and Code Availability
All code used for single-cell analysis and data presentation is available via GitHub at https://github.com/Camp-Lab/human_endoderm_atlas. The accession numbers for the mRNA sequencing data reported in this paper are ArrayExpress: E-MTAB-10187 and E-MTAB-10268. Data S1 and supporting raw data analysis have been deposited to Mendeley Data: https://doi.org/10.17632/x53tts3zfr.1. The expression and epithelium-mesenchyme interaction data could be explored via the interactive web browser GutTubeR (http://guttuber.iob.ch/app/). The code for the Shiny app is available via GitHub at https://github.com/Camp-Lab/GutTubeR/. Please contact the authors for inquiries.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human samples
Normal, de-identified human adult intestinal biopsies were collected from a 45-year-old male and a 50-year-old female with approval from the University of Michigan Institutional Review Boards (IRB). Biopsy specimens were stored on ice in a sterile saline solution prior to single-cell dissociation. Normal, de-identified developing human tissues were obtained from the University of Washington Laboratory of Developmental Biology, and all work was approved by the University of Washington and the University of Michigan IRB. Tissue was shipped overnight in Belzer-UW Cold Storage Solution (ThermoFisher, NC0952695) with cold packs.

A list of tissue specimens with detailed sample information can be found in Table S1. We collected different organs within comparable age range. For certain time points, we collected multiple organs from the same specimen, which enables rigorous quantification of organ-specific features due to the mitigation of confounding inter-specimen and technical variables. We collected tissues based on tissue quality, age, location and availability, and did not perform sample size estimation prior to study. To alleviate the influences of sex on results of the developing atlas, we excluded genes located on sex chromosomes before gene expression normalization and following analysis.

Human embryonic stem cells
Human ESC line H9 (NIH registry #0062, RRID: CVCL_9773) was obtained from the WiCell Research Institute and CDX2-control and CDX2- knockout lines used in this study were generated from the hESC H9 line (Kumar et al., 2019). All experiments using human ES cells were approved by the University of Michigan Pluripotent Stem Cell Research Oversight Committee. H9 cells were authenticated through Short Tandem Repeat (STR) DNA profiling (Matsuo et al., 1999) at the University of Michigan DNA Sequencing Core and showed an STR profile identical to the STR characteristics published by Josephson et al. (2006). The H9 cell line was negative for Mycoplasma. Stem cells were maintained in mTeSR1 medium (STEMCELL Technologies, Vancouver, Canada) with Matrigel (BD Biosciences, San Jose, CA). hESCs were passaged and differentiated into human intestinal organoid tissue as previously described (McCracken et al., 2011; Spence et al., 2011).
Immunocompromised mice for transplantation

Immunocompromised NOD-SCID IL2Rg null (NSG) mice for organoid transplantation were purchased from Jackson Laboratory (strain no. 0005557). Animal experiments were conducted in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The University of Michigan Animal Care and Use Committee approved all animal research performed in this study (IACUC; protocol # PRO00006609).

METHOD DETAILS

Generation and culture of in vitro human intestinal organoids

hESCs were differentiated into human intestinal organoids (HIOs) based on the previously described protocol (Capeling et al., 2020; Spence et al., 2011; Tsai et al., 2017). Briefly, hESCs were patterned into definitive endoderm (DE) by culturing in the presence of Activin A (100 ng/mL) in RPMI-1640 media for 3 days with increasing concentrations of HyClone FBS respectively (0%, 0.2%, 2%). Midgut/hindgut patterning was carried out in 2% HyClone FBS containing RPMI-1640 supplemented with FGF4 (500 ng/mL) (Sugawara et al., 2014) and CHIR99021 (2 μM) with daily media changes. The spheroids were collected at the end of day 5 of midgut/hindgut patterning except for CDX2 organoids which the spheroids generated from day 4-6 pooled. The collected spheroids were embedded in matrigel (8 mg/mL, Corning, 354234), cultured in ENR media (mini gut basal media, supplemented with EGF (RD Systems, 236-EG-01M, 100 ng/mL), Noggin-Fc (100 ng/mL) (purified from conditioned media; (Heijmans et al., 2013)), and R-Spondin 1 (5% conditioned medium) (Ootani et al., 2009). The media was changed every 4-5 days. The HIOs were passaged when they outgrew matrigel and/or accumulated excessive internal debris (Capeling et al., 2020). The mini gut basal media is composed of the following components: Advanced DMEM/F-12 (Life Technologies, 12634), 1x B27 supplement (Life Technologies, 17504044), 2 mM L-Glutamine (Life Technologies, 25030), 15 mM HEPES (Life Technologies, 15630080). All media used in the differentiation process contain 1x Penicillin-Streptomycin (Life Technologies, 15140).

For some experiments, varying concentrations and combinations of EGF (R&D, 236-EG) and/or NRG1 (R&D, 5898-NR-050) were tested instead of 100 ng/mL EGF in standard ENR media. The following conditions were tested: 100 ng/mL EGF without NRG1 [100E-0N], 100 ng/mL EGF and 1 ng/mL NRG1 [100E-1N], without EGF and 100 ng/mL NRG1 [0E-100N], 1 ng/mL EGF and 100 ng/mL NRG1 [1E-100N] and without either EGF and NRG1 [0E-0N]. HIOs for this experiment were collected and analyzed after 40 days of culture. Additionally, some experimental conditions contained organoids that were grown in ENR for the first 72 hours to pattern duodenal identity (Munera et al., 2017) and changed to Minigut basal media containing only EGF (100 ng/mL). Lastly, for some experiments, HIOs were grown in nonadhesive alginate hydrogel (1%) as previously described (Capeling et al., 2019).

Human intestinal organoid (HIO) transplantation

HIOs were grown in matrigel and ENR media for 4 weeks prior to transplantation. An exception to this were the HIOs processed for FISH experiments, which were grown in ENR media for the first 72 hours to pattern duodenal identity (Munera et al., 2017), after which media was changed to only contain EGF (100ng/mL) in mini gut basal media. On the day of the transplant, HIOs were mechanically dissociated from matrigel. Then, HIOs were implanted under the kidney capsules of immunocompromised NOD-SCID IL2Rg null (NSG) mice (Jackson Laboratory strain no. 0005557) as previously described (Finkbeiner et al., 2015b; Watson et al., 2014). In summary, mice were anesthetized using 2% isoflurane. A left-flank incision was used to expose the kidney after shaving and sterilization. The spheroids generated from day 4-6 pooled. The collected spheroids were embedded in matrigel (8 mg/mL, Corning, 354234), cultured in ENR media (mini gut basal media, supplemented with EGF (RD Systems, 236-EG-01M, 100 ng/mL), Noggin-Fc (100 ng/mL) (purified from conditioned media; (Heijmans et al., 2013)), and R-Spondin 1 (5% conditioned medium) (Ootani et al., 2009). The media was changed every 4-5 days. The HIOs were passaged when they outgrew matrigel and/or accumulated excessive internal debris (Capeling et al., 2020). The mini gut basal media is composed of the following components: Advanced DMEM/F-12 (Life Technologies, 12634), 1x B27 supplement (Life Technologies, 17504044), 2 mM L-Glutamine (Life Technologies, 25030), 15 mM HEPES (Life Technologies, 15630080). All media used in the differentiation process contain 1x Penicillin-Streptomycin (Life Technologies, 15140).

For some experiments, varying concentrations and combinations of EGF (R&D, 236-EG) and/or NRG1 (R&D, 5898-NR-050) were tested instead of 100 ng/mL EGF in standard ENR media. The following conditions were tested: 100 ng/mL EGF without NRG1 [100E-0N], 100 ng/mL EGF and 1 ng/mL NRG1 [100E-1N], without EGF and 100 ng/mL NRG1 [0E-100N], 1 ng/mL EGF and 100 ng/mL NRG1 [1E-100N] and without either EGF and NRG1 [0E-0N]. HIOs for this experiment were collected and analyzed after 40 days of culture. Additionally, some experimental conditions contained organoids that were grown in ENR for the first 72 hours to pattern duodenal identity (Munera et al., 2017) and changed to Minigut basal media containing only EGF (100 ng/mL). Lastly, for some experiments, HIOs were grown in nonadhesive alginate hydrogel (1%) as previously described (Capeling et al., 2019).

Chromogenic in Situ Hybridization and Multiplex Fluorescent in Situ Hybridization (FISH)

Paraffin blocks were sectioned to generate 5 μm-thick sections. Sections were mounted to SuperFrost Plus Slides (Thermo Scientific, 10149870) and used within one week for optimal results. The assay was carried out in RNase-free conditions by treating all materials with RNase-away (Molecular Bioproducts Inc., 7005-11) prior to use. Slides were stored at room temperature in a sealed slide box with silica desiccant packets. Slides were baked for 1 hour in a 60° C dry oven a day prior to starting the procedure. Chromogenic in situ hybridization (ISH) and multiplex fluorescent in situ hybridization (FISH) protocols were performed according to the manufacturer’s instructions (ISH–ACD; RNAscope 2.5 Assay BROWN 322310-QCK Rev B; FISH–ACD; RNAscope Multiplex Fluorescent v2 manual protocol, 323100-USM), under standard antigen retrieval conditions (15 min) and optimized protease treatment conditions for each tissue (intestine–30 min, tHIO–20 min). FISH samples were imaged using Nikon A1 confocal and images were assembled using Photoshop CC. Imaging parameters were kept consistent for all images within the same experiment and any post-imaging manipulations (i.e., brightness, contrast, LUTs) were performed equally on all images from a single experiment.

Tissue staining

Immunofluorescence stainings were conducted as previously described (Spence et al., 2009). Briefly, tissues were fixed in either 4% PFA overnight or 10% neutral buffered formalin (NBF) for approximately 24 hours at room temperature. The following day, tissues were washed in UltraPure Distilled Water (Invitrogen, 10977-015) for 3 changes for a total of 2 hours. Tissue was gradually dehydrated in a methanol series (25%, 50%, 75%, 100%). Tissue was stored long-term in 100% Methanol at 4°C. Prior to paraffin processing,
tissue was equilibrated in 100% ethanol for an hour followed by 70% ethanol. Tissue was paraffin perfused using an automated tissue processor (Leica ASP300) with 1-hour solution changes overnight. Paraffin processed tissue was embedded and 5-7 μm sections were cut.

Paraffin sections were first deparaffinized in Histoclear and re-hydrated. Antigen retrieval was performed by steaming slides in a sodium citrate buffer for 20 minutes. Slides underwent a blocking step using the appropriate serum (5% normal donkey serum in PBS + 0.5% Triton-X) for 1 hour at room temperature. Primary antibodies were diluted in blocking solution (1:500) and slides were incubated with antibodies overnight at 4°C. The following day, slides were washed and incubated with appropriate secondary antibodies (1:500) diluted in a blocking buffer for 1 hour at room temperature together with DAPI staining (1 μg/mL). Slides were washed and mounted using Prolong Gold (Thermo Fisher, P10144). A list of antibodies can be found in the Key Resources Table.

H&E staining was performed with Harris Modified Hematoxylin (Fisher Scientific, SH26-500D) and Shandon Eosin Y (Thermo Scientific, 6766007) based on the manufacturer’s instructions.

**Single-cell dissociation and RNA sequencing library preparation of the developing human tissues**

Developing human tissue and organoid dissociations for scRNA-seq were performed according to a previously published study (Miller et al., 2019). Initially, all tubes and pipette tips were pre-washed with 1% BSA in HBSS with Mg²⁺ and Ca²⁺ to prevent adhesion of cells to the plastic. Organoids were mechanically removed from matrigel droplets. Next, organoids and primary tissues minced into small fragments, using a scalpel in a Petri dish filled with ice-cold 1X HBSS. Then, tissue and organoid fragments were transferred into a 15 mL conical tube. Enzymatic dissociation started in the Neural Tissue Dissociation Kit (Miltenyi Biotec, 130-092-628), and all incubation and centrifugation steps were carried out in a refrigerated centrifuge pre-chilled to 10°C for the developing tissue samples, whereas, 37°C for organoid samples unless otherwise stated. Based on the manufacturer’s instructions, the fragments of tissue were treated for 15 minutes with Mix 1. Mix 2 was added to the digestion, and tissue was incubated for 10-minute increments until digestion was complete. After each 10-minute incubation, tissue was agitated using a P1000, and tissue digestion was visually assessed under a stereomicroscope. This process continued until the tissue was fully digested. Cells were filtered through a 70 μm filter coated with 1% BSA in 1X HBSS, spun down at 500 g for 5 minutes at 10°C, and resuspended in 500μl 1X HBSS. 1 mL red blood cell lysis buffer (Roche, 11814389001) was then added to the tube and the cell mixture was placed on a rocker for 15 minutes at 4°C only for the primary tissue samples.

Cells were spun down (500 g for 5 minutes at 10°C) and washed twice by suspension in 2 mL of HBSS + 1% BSA followed by centrifugation. Two exceptions were definitive endoderm and spheroid (60 HIOs) samples that were dissociated using TrypLE Express (GIBCO, 12605-010) at 37°C. Placement of the cell suspension was changed every 15 minutes. The pH of the cell suspension was monitored in 15 mL tubes filled with ice-cold 1X HBSS for 15 minutes. The pH was then adjusted to 7.4 using NaHCO₃. Cells were counted using a hemocytometer and kept on ice. Single-cell droplets were immediately prepared on the 10x Chromium according to manufacturer instructions by The Advanced Genomics Core at the University of Michigan, with a target of capturing 5,000-10,000 cells. Single-cell libraries were prepared using the 10x Chromium Single Cell 3’ v2 and Next GEM Single Cell 3’ v3.1 kits according to manufacturer instructions.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

**Alignment of single-cell transcriptomes**

We used Cell Ranger (10x Genomics) to demultiplex base call files to FASTQ files and align reads. Default alignment parameters were used to align reads to the human reference genome provided by Cell Ranger. In vitro organoid data (except for CDX2 control and knockout data) and human primary tissue data were mapped to hg19. CDX2 control and knockout line-derived organoid data were mapped to hg38. Transplanted organoid data were mapped to the human-mouse dual genome (hg19 and mm10). We only used human cells classified by Cell Ranger for downstream analysis.

**Analysis of single-cell RNA-seq data**

Cell Ranger output of scRNA-seq data of developing human duodenum and early small intestine (ArrayExpress: E-MTAB-9489, E-MTAB-9363), lung (ArrayExpress: E-MTAB-8821), days 0 and 3 in vitro matrigel-embedded ENR media cultured human intestinal organoid data (ArrayExpress: E-MTAB-9228) and mouse gastrulation atlas data (ArrayExpress: E-MTAB-6967) were retrieved from the ArrayExpress database (Holloway et al., 2020, 2021; Miller et al., 2020; Pijuan-Sala et al., 2019). scRNA-seq data of adult ileum fresh crypt and adult stem cell-derived organoids were retrieved from GEO with the accession number GSE19969 (Fujii et al., 2018).

In vitro HIO time course, CDX2 knockout and control HIOs, tHIO, developing multi-organ atlas, in vitro HIOs grown with NRG1, EGF or neither (referred as “NRG1 or EGF grown HIOs” hereafter), 4-week in vitro HIOs embedded in alginate or matrigel, grown with ENR or EGF-only media (“alginate or matrigel embedded HIOs”), adult duodenum data, and adult ileum fresh crypt and adult stem cell-derived enteroid data (“adult enteroid”) were analyzed individually. Seurat (v3.1) package (Butler et al., 2018) was applied to the scRNaseq data for preprocessing. Generally, cells with less than 1,000 or more than 20,000 detected genes, as well as those with mitochondrial transcript proportion higher than 10% (all except adult duodenum samples) or 50% (adult duodenum samples) were excluded. For the developing human multi-organ cell atlas data, CDX2 knockout and control HIOs, NRG1 or EGF grown HIOs, alginate or matrigel embedded HIOs and adult enteroid datasets, ribosomal genes, mitochondrial genes, and genes located on sex chromosomes were removed. After log-normalization, 2,000 or 3,000 highly variable genes were identified using the default vst method. The normalized expression levels were then z-transformed, followed by principal component analysis (PCA) for dimension reduction. Uniform manifold approximation and projection (UMAP) (Becht et al., 2018) and Louvain clustering was applied to the
top 10 principal components (PCs) of the esophagus and stomach epithelium, and top 20 PCs of the alginate or matrigel embedded HIO dataset to visualize and understand the cellular heterogeneity.

For data other than the stomach and esophagus epithelium, and the alginate or matrigel embedded HIO dataset presented in this manuscript, Cluster Similarity Spectrum (CSS) was calculated to integrate data of different samples before UMAP calculation and clustering as described in (He et al., 2020). In brief, cells from each sample were subsetted, and Louvain clustering (with resolution 0.6), implemented in Seurat, was applied based on the pre-calculated top 20 PCs. The average expression of the pre-defined highly variable genes was calculated for each cluster in each sample. Afterward, Spearman’s correlation coefficient was calculated between every cell and every cluster in each sample, resulting in a correlation vector for each cell. For each cell, its correlations with different clusters of each sample were z-transformed (similarity spectrum). Its z-transformed similarities to clusters of different samples were then concatenated as the final CSS representation. We note that we obtained similar integration results using MNN implemented in Seurat.

For the time course in vitro HIO dataset, its SPRING-based cell embedding for visualization was generated as previously described (Kanton et al., 2019). In brief, single cells from the same cluster of the same sample with a similar transcriptome were grouped into pseudo-cells. The correlation distance between CSS of each pair of pseudo-cells was calculated and a k-nearest neighbor (kNN) network (k = 20) was then calculated with the constraint to only consider pseudo-cells from the same or nearby stages when screening for nearest neighbors. The kNN network was visualized using SPRING (Weinreb et al., 2018). Coordinates of single cells were predicted from pseudo-cells based on CSS using support vector regression (SVR) implemented in the e1071 package. Each SVR model was trained for one dimension of coordinates. Such coordinates were further refined by pushing each cell to its nearest pseudo-cell with the smallest correlation distance of CSS to be 80% closer.

Cluster annotation was done based on expression patterns of canonical cell type markers found in literature, together with cluster markers identified using the presto package. For analysis of thIO, epithelium of developing lung, stomach, intestine and duodenum, duodenum mesenchyme, and adult duodenum, clusters with similar transcriptome and shared annotation were merged. Clusters with marker expression of multiple cell class and without specifically enriched positive markers were considered as doublet clusters and removed from the downstream analysis. During epithelial cellular heterogeneity analysis of each organ, we examined the expression of canonical tissue markers and removed the potentially mis-dissected cells from nearby tissues on the cluster level based on the expression patterns of those markers. Specifically, we used CDX2 to mark the intestine; co-expression of GATA4 and SOX2 to mark the stomach; expression of PTF1A, PROX1, NKX6-1 and ONECUT1 to mark pancreas; SOX2 and NKX2-1 to mark lung; expression of SOX2 and depletion of GATA4 to mark esophagus.

Identification of cluster and cell type markers were identified by calculating the area under the receiver operator curve (AUC) implemented in the presto package. We used AUC > 0.6, log-transformed expression level fold change > 0.25 as the significance cutoff.

Quantification of organ specificity
K-nearest neighboring cells (k = 100) were obtained for each cell of the developing human multi-organ atlas by the nn2 function implemented in the RANN package, based on the Euclidean distance in CSS space. For each cell, we calculated the proportions of each organ among its neighboring cells. We also calculated the proportion of each organ in the whole dataset. Our rationale is, for a cell, if the proportion of an organ estimated from its neighboring population exceeds that estimated from the whole dataset, and if this organ is different from the organ identity of the cell, i.e., this is a non-self-organ, this cell tends to share transcriptome similarity with that non-self-organ. The larger number of non-self-organs fitting this situation indicates this cell intermix with more organs on the transcriptome level, i.e., this cell shows lower organ specificity. Therefore, we could use this number to quantify the organ specificity of each cell. The distribution of this number among a group of cells could indicate the organ specificity of that cell population.

Identification of stem cell clusters, stemness genes shared by multiple organs, and functional enrichment analysis
We performed sub-clustering in epithelial cells of each organ, and annotated cell types according to canonical cell type markers and de novo identified cluster markers. We selected bud tip progenitors and basal cells from the lung, basal cell progenitor from the esophagus, antrum and corpus stem cells from the stomach and intestine stem cells from different regions of the small intestine as the stem cell clusters for downstream analysis. In the downstream analysis, we only considered samples that are at the age of 11, 14, 15, 17, 18, 19 post-conception weeks to ensure the age ranges of different tissues are comparable.

To identify core features of stem cells that are shared by multiple developing organs, we first identified differentially expressed (DE) genes between selected stem cells mentioned above and other cells in the age controlled multi-organ epithelium data. We used AUC > 0.6, log-transformed expression level fold change > 0.25, expressed in over 25% cells of the tested cluster as the significance cutoff. To identify the stem cells enriched genes shared by multiple organs, we did stepwise filtering. First, we calculated the non-proliferative cell type average expression level and expressed proportion for each gene, and selected only DE genes that are expressed in more than 25% cells of each stem cell cluster. Next, we constructed an artificial pattern with “1” in all selected stem cell clusters and “0” in other cell types and correlated the expression pattern of the selected genes with the artificial pattern using Spearman correlation coefficients (SCC). Then, we permuted the pattern and calculated the correlation 100 times. Only those with real SCC > 0.5 and real SCC > permuted SCC in all permutations would be further selected. Among the remaining genes, we defined the top 100 genes ranked by real SCC as the core stemness genes.
To perform KEGG pathway and GO biological process (BP) enrichment analysis for the core stemness genes, we downloaded KEGG and BP annotation from MSigDB (v7.1) (Liberzon et al., 2011; Subramanian et al., 2005). We examined enrichment by one-sided Fisher’s exact test. For GO BP enrichment, we set Bonferroni corrected p < 0.01 and more than 10 core stemness genes overlapped with the term as significance cutoff. For KEGG pathway enrichment, we set Benjamini-Hochberg (BH)-corrected p < 0.05 and more than 2 core stemness genes overlapped as significance cutoff.

Quantification of pseudospace transition and identification of genes associated with tissue identity in stem cells

We characterized the pseudospace transition of epithelial stem cells from various tissues with a transcriptome-deconvolution-based pseudo-spatial score. We considered the reconstructed pseudospace transition to reflect two layers of difference: 1) difference between digestive tract organs and the lung; 2) difference among digestive tract organs along the anterior-to-posterior axis. Lung bud tip progenitor and colon stem cells were thus considered as the two endpoints of the pseudospace. Intermediates were considered as mixtures of different proportions of the endpoints. We quantified this transition using transcriptome deconvolution. To do this, we first identified differentially expressed (DE) genes among the lung, esophagus, stomach, small intestine, and colon, or among duodenum, jejunum, and ileum, or between stomach corpus and antrum. We used AUC > 0.6, log-transformed expression level fold change > 0.25, expressed in over 25% cells of the tested cluster and Wilcoxon rank-sum test Benjamini-Hochberg-corrected P value < 0.01 as the significance cutoff. We took the union set of top-50 DE genes (ranked by AUC) among organs, among small intestine regions, and between stomach regions as feature genes. We then calculated the average expression level of feature genes in lung basal cell and colon epithelial stem cells. Finally, we used the obtained average expression matrix as a reference transcriptome, and performed transcriptome deconvolution (Treutlein et al., 2014) for each epithelial stem cell using quadratic programming via the quadprog package in R. The estimated fraction of colon identity was used as the pseudo-spatial score. We used the same method, reference pattern and feature genes to estimate the regional identity of thIO stem cells.

In addition to the DE genes identified above, we took a second approach to identify molecular features associated with tissue identity specification in stem cells. Epithelial stem cells were ordered according to the ascending order of colon identity fraction. Every 20 cells were binned. We calculated the bin average expression levels for each gene. Next, we applied the ‘ns’ function (implemented in the splines R package) to construct a natural spline linear regression model (df = 5) for each gene, using pseudospace orders of bins as variables and bin average gene expression levels as a response. Then we used an F-test to compare the residual of variations that could not be explained by the natural spline regression model to that of a constant model. Genes with significant expression changes along pseudospace were defined as F-test BH-corrected p < 0.01 and log-transformed fold change of maximum versus minimum bin average expression levels > 1. We combined the DE gene lists from the two approaches and presented the transcription factors annotated in HumanTFDB of animalTFDB 3.0 (Hu et al., 2019).

Inference of interaction between epithelium and mesenchyme in each organ

We applied CellPhoneDB (v2.0) (Efremova et al., 2020) to the normalized expression levels of epithelium and mesenchyme of each developing organ with default parameters to investigate the between cell type communications. We took lung, esophagus, stomach, small intestine and colon samples that are at the age of 11, 14, 15, 17, 18, 19 post-conception weeks to ensure the age ranges of different organs are comparable. In mesenchyme, we randomly took at most 8,000 cells from each organ, and only kept cell types with more than 95 cells after sub-sampling. In epithelium, we randomly took at most 3,000 cells from each organ, and only kept cell types with more than 30 (in the stomach) or 50 (all except stomach) cells after sub-sampling. We applied CellPhoneDB to each organ separately. We used the default significance cutoff by CellPhoneDB. We focused on simple ligand-receptor pairs that were significant in at least epithelial-mesenchymal cell type pairs of one organ. We presented the reported mean values of each significant pair with heatmap, with interaction direction classified.

Projecting organoid cells to the developing human multi-organ atlas

First, we calculated the representations of in vitro HIO and thIO cells which were compatible with the CSS-represented developing human multi-organ atlas. In brief, we calculated the Spearman’s correlation coefficient between the in vitro HIO and thIO cells and every cluster of each developing human sample using the highly variable genes defined in the developing human atlas. Afterward, correlations of each HIO cell with different clusters of each developing sample were z-transformed. The z-transformed similarities to clusters of different samples were then concatenated as the final Reference Similarity Spectrum to the developing human atlas representation (RSS in short), which share the same dimensionality and definitions as the CSS representation used in the atlas.

To allow the projection of the thIO cells to the UMAP embedding of developing human multi-organ atlas, we first built a UMAP model based on the CSS of the human developing atlas using the ‘umap’ function (with ‘ret_model’ parameter as TRUE) implemented in the uwot package. Based on the RSS of thIO cells and the pre-trained UMAP model of the developing human endoderm cell atlas data, we used the ‘umap_transform’ function in the uwot package to project thIO cells to the UMAP embedding of the developing human multi-organ atlas.

To predict the cell type and cell state identities of cells in both in vitro HIO, thIO samples, as well as the CDX2-KO and control organoids, we identified the k-nearest (k = 20) developing human cell neighbors of each in vitro HIO or thIO cell, by calculating...
Euclidean distance between RSS of organoid cells and CSS matrices of the developing human multi-organ cell atlas using the ‘nn2’ function in the RANN package. The most frequent developing human cell identity among the nearest neighbors was defined as the mapped human in vivo identity of each organoid cell.

10-fold cross-validation of projection-based organ identity inference
To estimate the accuracy of the projection-based organ identity inference as described above, we performed 10-fold cross-validation on the multi-organ developing atlas. We randomly split the whole dataset into 10 groups, with roughly equal numbers of cells in each group. Each time we pooled 9 groups of cells as a training set to build the prediction model and applied the model to the remaining group used as the test set. This process was repeated ten times. For cells of each organ, the number of cells with different predicted organ identities were counted and compared. This was done for different cell classes separately.

Resolving intestinal region specificity of epithelial stem cells in developing human and tHIOs
To avoid confounding effects of individual variations, we only selected developing human specimens with at least three intestine regions collected, i.e., the day 80, 101 and 132 old specimens. For each of the cell types of interest, we first identified genes showing differential expression (DE) levels between the small intestine and colon, or between different small intestine regions. Specifically, to identify DE genes between the small intestine and colon, we compared each of the collected small intestine tissues to the colon for each specimen separately. In each comparison, both groups were required to have at least 10 cells, otherwise, the comparison was skipped. We defined genes with BH-adjusted Wilcoxon rank-sum test P-value < 0.01 and log-transformed expression fold change > 0.25 as positive markers. We took an equal number of top positive markers (ranked by AUC) from each of the two groups in comparison and defined those as DE genes between a specific small intestine region and colon in one specific specimen. Genes defined as DE genes in a comparison between any of the small intestine regions and colon, in any of the selected specimens were defined as DE genes between small intestine and colon. Similarly, to identify DE genes among different small intestine regions, we did a pairwise comparison between small intestine regions in each specimen and took their union set. The only difference is that positive markers in comparisons among the small intestine regions were defined as log-transformed expression fold change > 0.1 and BH-adjusted Wilcoxon rank-sum test P-value < 0.01. DE genes identified in any of the comparisons, which were also detected in tHIO cells, were defined as intestinal region-dependent genes.

We then calculated the average expression levels of intestine region-dependent genes in tHIO in each selected cell type, and compared it to the cell type counterpart of each developing intestine region of each specimen by calculating Spearman’s correlation coefficients.

Transcriptome similarity-based mapping between datasets
To map the in vitro HIO, CDX2-KO and control organoid cells to the mouse gastrulation cell types, we first identified 3,000 highly variable genes in the mouse gastrulation atlas (Pijuan-Sala et al., 2019) using the default vst function implemented in the Seurat package. We retrieved human-mouse orthologous gene information from Ensembl (version 92). Highly variable genes with one-to-one human-mouse orthologs that were detected in the query data of human organoids were used as the feature genes. We used the average expression levels of the feature genes in each annotated mouse cell type as the reference and calculated Spearman correlation coefficients between organoid cells and mouse cell type. Organoid cells were mapped to the mouse cell type with the highest similarity.

To map mesenchymal cells of tHIO to the developing human duodenum mesenchymal subtypes, we first identified developing duodenum mesenchymal subtype markers as feature genes. We used average expression levels of feature genes in each duodenum mesenchymal subtype as the reference and calculated Spearman correlation coefficients between expression profiles of organoid mesenchymal cells and developing duodenum mesenchymal subtypes. An organoid mesenchymal cell was mapped to the duodenum mesenchymal subtype with the highest similarity.

To identify the precursors or counterparts to the 7 PCW developing duodenum non-proliferative mesenchymal subtypes in the in vitro HIOs, we first extracted non-epithelial cell clusters from the time course in vitro HIO data, did sub-clustering and refined HIO mesenchymal cell cluster annotation based on sub-clustering results. We defined two NKX2-3+/MKI67− clusters that are enriched for 4-week cells as intestinal differentiated mesenchymal cell clusters. We did further sub-clustering on these two clusters and obtained five sub-clusters (M1−M5). We identified the top 3000 highly variable genes in the duodenum mesenchyme dataset using the default vst method in Seurat as feature genes. Cell-cycle-related genes were excluded from the feature genes. Next, we calculated the average expression levels of the identified feature genes in the five organoid differentiated mesenchymal sub-clusters, as well as the 7 PCW duodenum mesenchymal subtypes. We combined CD81-high, GDF10+ and PITX1+ SMC into one cell type (CD81-high like) in PCW 7 considering these three cell types co-clustered in the individual analysis of this sample. Spearman correlation coefficients between the average expression profiles of the organoid differentiated mesenchymal sub-clusters and the 7 PCW duodenum mesenchymal subtypes were calculated. An organoid mesenchymal sub-cluster was mapped to one or several duodenum mesenchymal subtypes with relatively higher similarity.

To map non-proliferative mesenchymal cells in CDX2-KO and control organoid to the in vitro HIO non-proliferative mesenchymal subtypes, we identified markers of non-proliferative mesenchymal subtypes in 4-week in vitro HIOs as feature genes. We used average expression levels of feature genes in each of the selected subtypes in 4-week in vitro HIO as the reference, and calculated
Spearman correlation coefficients between CDX2-KO/control organoid non-proliferative mesenchymal cells and the reference. Selected CDX2-KO/control cells were mapped to the 4-week in vitro HIO mesenchymal subtype with the highest similarity. We annotated mesenchymal clusters (c1, 3, 8, 10) observed in the time course HIO data based on enrichment of 4-week in vitro HIO mesenchyme subtypes, and further mapped the CDX2-KO and control cells to time course in vitro HIO mesenchymal cells accordingly.

**Identification of in vitro HIO epithelial cells that have acquired intestinal identity**

To identify the in vitro HIO cells that have acquired intestinal identity, we took several filtering steps. Here, we only considered cells of day 0 (spheroid) to day 30 HIO, excluding cells of definitive endoderm. First, we calculated the Pearson’s correlation coefficients between HIO cells and cell types of mouse gastrulation (MG) stage atlas (Pijuan-Sala et al., 2019) using the 2,000 highly variable genes defined in the MG data, and kept the HIO clusters with a minimum of 80% of the cells being most similar to the gut. Within these HIO clusters, HIO cells not mapping to the gut were also excluded. Then we sub-clustered each sample separately. For samples of days 0, 3 or 7, we excluded clusters showing depleted expression of CDX2. For samples of day 14 or later, we inferred the developing organ identity based on the kNN method as described for each HIO cell. We choose day 14 as a boundary age because from day 14 on, we observed clusters enriched in cells mapping to the developing intestine. HIO samples younger than day 14 were considered as too young to use the developing human multi-organ cell atlas to determine organ identity. For samples of day 14 or later, cell clusters with more than 50% cells that were not mapping to the developing intestine, as well as any cells not mapping to the developing intestine, were excluded. Finally, we examined the HIO clusters defined for the whole in vitro HIO dataset again, and further excluded the clusters with more than 75% cells excluded in the previous steps.

**Identification of differentially expressed genes during maturation of developing human duodenum epithelial stem cells (pseudotime-dependent genes of developing duodenum stem cells)**

To study the maturation of developing human duodenum epithelial stem cells, we ordered these cells in a pseudo-temporal order to generate a pseudo time course. We applied DiffusionMap (implemented in destiny package, with default setting except for k = 20) (Angerer et al., 2016) to CSS of developing duodenum stem cells, and used the rank of the first diffusion component as pseudotime.

To identify genes showing differential expression levels along the constructed pseudotime course, we grouped every 50 cells into a bin based on their estimated pseudotime, and calculated the average gene expression level of every gene. Next, we applied the ‘ns’ function (implemented in the splines R package) to construct a natural spline linear regression model (df = 5) for each gene, using pseudotime orders of bins as the independent variable and average gene expression levels of bins as the response. Then we used an F-test to compare the residual of variations that could not be explained by the natural spline regression model to that of a constant model. Genes with significant expression changes along pseudotime were defined as genes with F-test BH-corrected p < 0.01 and log-transformed fold change of maximum versus minimum bin average expression levels > 0.5.

**Quantification of stem cell maturity and identification of stem-cell-phase-enriched genes**

We quantified stem cell maturity using a transcriptome deconvolution approach. Here we selected only epithelial cells of in vitro HIOs that have acquired intestine identity, as well as stem cells in the HIO and the developing duodenum for this analysis. We first removed the mitochondrial genes and genes located on the sex chromosomes and log-normalized the count matrix. Cells in the HIO spheroid were considered to have the lowest maturity, while the epithelial stem cells in the most mature (19 post-conception weeks, 19 PCW) developing duodenum were considered to have the highest maturity. The maturation signatures of other intermediates were considered as the mixtures with different proportions of the least and the most mature stem cells, which can be quantified using transcriptome deconvolution. To do this, we first applied hierarchical clustering to genes with significant expression changes during maturation of developing human duodenum epithelial stem cells which were identified as described above and selected the gene clusters showing a monotonous increase or decrease along the trajectory as feature genes for deconvolution. We then calculated the average expression levels of feature genes in HIO c12 (the day 0 epithelium primed cluster) and stem cells of 19 PCW (the latest time point of the developing duodenum sample in the dataset), separately. Finally, we used the obtained average expression matrix as the reference transcriptome, and performed transcriptome deconvolution for each selected cell using quadratic programming via the quadprog package in R. The estimated fraction of 19 PCW intestinal stem cell identity was used as the proxy of intestinal stem cell (ISC) maturity.

We defined ISC state as the fraction of the most mature developing (19 PCW / 132 post-conception days) stem cell identity estimated by quadratic programming in organoids and developing duodenum, and pseudo-temporal orders of stem cells in adult duodenum. To identify the genes associated with ISC development, we ordered in vitro HIOs intestinal epithelial cells and the developing duodenum stem cells in the ascending order of the estimated ISC maturity and grouped every 50 cells into a cell bin. We used ISC maturity orders of bins as the independent variable and average gene expression levels of bins as the response, identified genes showing significant variable expression levels along the trajectory with the spline-based F-test method described above.

We applied the same DiffusionMap-based method as described above to resolve the stem cell heterogeneity of the adult duodenum, and grouped every 20 adult stem cells into a cell bin. We then performed hierarchical clustering on the ISC bins of in vitro HIOs, developing duodenum and adult duodenum, based on Spearman correlation distances across the ISC development associated genes, to classify the cell bins into 6 phases. We then identified ISC phase-enriched genes using the presto package. We used AUC > 0.6, log-transformed expression level fold change > 0.25, expressed in more than 20% cells of the testing cell type,
Benjamini-Hochberg (BH) adjusted P value of Wilcoxon rank-sum test < 0.01 and expressed cell proportion difference between testing cell type and others > 20% as the significance cutoff. We excluded genes that are significantly enriched in enterocyte of developing or adult duodenum, which were AUC > 0.7, log-transformed expression level fold change between enterocyte over stem cells > 0.2, expressed in more than 25% of enterocytes, BH-adjusted P value of Wilcoxon rank-sum test < 0.01 and expressed cell proportion in enterocyte is higher than that of stem cells for more than 20%.

For the comparison of intestinal epithelial stem cell maturity among HIOs cultured in different conditions, we first integrated epithelial cells of *in vitro* HIOs cultured in different conditions (4-week or older) with MNN (Haghverdi et al., 2018), applied Louvain clustering to the top 20 MNN components and projected cells into 2D-embedding using UMAP. To choose intestinal epithelial cells, we did several filtering steps. We only kept cells being projected to the intestine and belonging to the clusters defined in the integrated dataset and individual datasets that have more than half of the cells being projected to the developing intestine. We excluded an enterocyte cluster which is APOA4+. Finally, we used the same quadratic programming-based method as described to estimate their ISC maturity, with the same reference pattern and feature genes for deconvolution being used.

We also used additional approaches to identify genes related to stem cell maturation from spheroids to mature developing duodenum stem cells, we combined two different approaches. In the first approach, we first calculated Spearman’s correlation coefficients between ISC identity fraction and gene expression level, and then took the top 100 genes showing the strongest positive and negative correlation, respectively. In the second approach, we first ordered cells included in the analysis based on their ISC identity in ascending order. Next, we identified genes with expression associated with the ISC identity using the same method to identify pseudotime-dependent genes of developing duodenum stem cells as described above. We then classified genes into gene modules with hierarchical clustering on their expression patterns across bins, took the cluster that turned on at the late HIO stage and kept stable expression throughout the developing duodenum stage and selected 100 genes showing the highest correlation with ISC identity fraction in this cluster. The union of genes identified with the two approaches was considered as genes related to stem cell maturation.

For the identification of DE genes between the developing and adult duodenum stem cells, we selected the developing stem cells with 100% ISC identity fraction and compared them with the adult duodenum stem cells using the presto R package. Genes with significant differential expression levels between developing and adult stem cells were defined as AUC > 0.6, log-transformed expression fold change > 0.25 and expressed in over 25% cells of the group being tested. Stem cell scores of tHIO and duodenum epithelial cells were calculated as cumulative z-transformed expressions of the union set of stem cell markers identified in the developing or adult tissue.

**Pseudo-spatial reconstruction of scRNA-seq data**

To reconstruct the pseudo-spatial distribution of the developing duodenum epithelial and mesenchymal cells, we applied novoSpaRc (Nitzan et al., 2019) to the corresponding scRNA-seq datasets. For the developing duodenum, we used the 19 PCW tissue sample. We first obtained the top 100 cell type markers (ranked by AUC, with presto R package) for each cell type in epithelium and mesenchyme, separately. We used the normalized expression levels of the identified markers as well as 22 reference genes as the inputs for novoSpaRc. The chosen reference genes are APOA4, BMP3, CDH1, CHGA, COL1A2, DLL1, EGF, ECAD, EBB3, F3, GPX3, LGR5, MKI67, MUC2, NPY, NRG1, PDGFRα, RSP02, RSP03, SLC2A2, TAGLN and WNT2B. We manually curated binarized expression patterns of selected reference genes, according to previously reported *in situ* hybridizations (RNA-scope) experiments in the developing human intestine (Holloway et al., 2020, 2021) and mouse enterocyte zonation atlas (Moor et al., 2018). Binarized expression patterns of reference genes were used to guide novoSpaRc inference. Reference intestine morphology with epithelium and mesenchyme was provided to novoSpaRc. novoSpaRc was run with the author’s suggested parameters (https://github.com/rajewsky-lab/novosparc).

**Reconstruction of regulons in duodenum intestinal stem cells, HIO mesenchymal cells and epithelial cells**

To reconstruct regulons in HIO epithelial cells, we applied pySCENIC (Aibar et al., 2017) on the count data of epithelial cell clusters (c0, 2, 6, 12, 7, 9, 11, 14) with default parameters. Due to the stochasticity of regulon reconstruction by pySCENIC, we repeated it 10 times and only kept the TF-target reported at least twice.

To reconstruct regulons in HIO mesenchymal cells, we applied pySCENIC on the count data of epithelial cell clusters (c1, 3, 5, 8, 10) with default parameters. We repeated it 10 times and only kept the TF-target reported at least twice. To identify regulons associated with intestine mesenchyme development, we first identified markers of the HIO mesenchymal clusters using presto. We defined cluster markers as AUC > 0.55, logFC of expression levels > 0.1, expressed cell proportion in testing cluster > 20%, difference of expressed cell proportion in testing cluster compared to others > 10, BH-adjusted P value of Wilcoxon rank-sum test < 0.05. We defined intestine mesenchyme development associated regulations as the TF-target pairs where both the genes are enriched in the mesoderm progenitor cluster (CDX2+/HAND1+/MKI67+) or NKX2-3+ clusters. To visualize the correlation network of intestine mesenchyme development associated regulons, we got the regulon activity by summing up the normalized expression levels of the master TF and its predicted target genes in each intestinal mesenchymal cell of *in vitro* HIOs. We calculated the pairwise Pearson correlation coefficients among activities of different regulons and kept the regulons with PCC > 0.4 with more than 4 other regulons for network presentation. We visualized the regulon correlation network using the igraph package in R.
To reconstruct regulons in intestinal stem cells (ISCs) of developing and adult duodenum, we applied pySCENIC on the count data of developing and adult duodenum ISC clusters with default parameters. We repeated it 10 times and only kept the TF-target pairs reported at least twice. We defined the TF–target pairs that are both ISC phase 4, 5 or 6 enriched genes as the phase 4, 5, or 6 ISC associated regulons.

Construction of transcription factor (TF) co-expression network
We used HIO clusters 0, 1, 2, 3, 5, 7, 8, 9, 10, 12, 14 to construct the TF co-expression network associated with the development of intestinal epithelium and mesenchyme. The Human TF list was downloaded from Human TFDB (Hu et al., 2019). With differential expression analysis done by the presto R package, we identified cluster enriched TFs with the criteria AUC > 0.6, log-transformed fold change > 0.1 and expressed in more than 10 cells. We calculated Pearson’s correlation coefficients of cluster enriched TF expression patterns across clusters. Only correlation > 0.8 was defined as a significant correlation, and only genes with no less than 10 expression-correlated genes were kept. We then used the igraph for network visualization.

Curation of GI disease-associated genes
The list of disease genes was drawn from the NIH Genetics Home Reference (https://medlineplus.gov/genetics/) and The Genetic and Rare Diseases Information Center (GARD) (https://rarediseases.info.nih.gov). We included all genes they listed as being associated with a GI disease from these two databases. When there are no genes reported to be associated with a GI disease listed in these resources, we manually searched for at least one human study reported on the particular GI-disease. The curated disease-associated genes are listed in Table S2. Median expression levels of NPSR1 and NLRC4 in each tHIO cell type were 0, so these two genes are not presented in the heatmap of Figure S5. All other genes were listed in Table S2.
Figure S1. Characterization of the developing human multi-organ cell atlas, related to Figure 1
A) The dendrogram shows cluster similarity (Pearson) based on highly variable genes, with selected cell class markers, top cluster marker genes, and a sidebar showing the proportion of cells of each organ per cluster. B) UMAP embedding of the developing human reference atlas colored by organ intermixing index. This index represents the number of organs with shared transcriptome features of each cell, excluding the organ of the examined cell. See also STAR methods. Right, stacked bar plots show the distribution of the organ intermixing index of each cell class. C) Dot plots show the expression of organ enriched transcription factors (TF) in the mesenchyme (left) or epithelium (right). D) Bar plots show 10-fold cross-validation of organ identity inference based on projection to developing multi-organ atlas. See also STAR Methods. Plot titles indicate real organ identity. Bars indicate predicted organ identity.
Figure S2. Developing human duodenum and THIO epithelial stem cells are distinct from the adult state, related to Figure 5
A) UMAP embedding of the developing duodenum and B) adult duodenum epithelial cells colored and numbered by annotated cell type. C) Dot plot shows the average gene expression levels (color) and expressed proportion (size) of cell type markers in the adult duodenum. D) Left, volcano plot shows genes with higher expression in the developing (left, purple) or adult (right, red) intestinal stem cells (ISCs). x axis presents log-transformed expression fold change in adult versus developing duodenum; the y axis presents the absolute difference between the area under receiver operator (auROC, AUC) and 0.5, which was used to quantify

(legend continued on next page)
the effectiveness of using a gene to classify two groups. Right, boxplots (right) show expression distributions of differentially expressed genes in adult, developing, and THIO stem cells. E) Heatmap shows the relative cluster average expression of intestinal epithelial cell type marker genes in THIO, developing, and adult intestine integrated data. F) Feature plots show cell type marker expressions in the THIO, developing and adult intestine integrated UMAP embedding. G) UMAP embedding of THIO, developing duodenum, adult small intestine and adult stem cell-derived enteroids of EGF and SB202190 (SB, enteroid 2) or IGF1 and FGF2 treated ileal enteroids (IF, enteroid 1) (Fuji et al., 2018), with cells colored by cell type (left) or source (right). H) Dendrogram showing transcriptome distance between cell types of different sources.
Figure S3. Mouse gastrulation map enables assessment of cell composition of in vitro HIO time course data, related to Figure 5

A) Heatmap showing relative expression levels of top 50 cluster markers across each HIO time course cluster. B) HIO cells were compared to a single-cell reference atlas covering multiple stages of mouse gastrulation (MG) (Pijuan-Sala et al., 2019). SPRING plot of HIOs colored by maximum Spearman correlation coefficient (SCC) to the MG reference cluster. Inset shows MG clusters with the highest similarity in the HIO. C) Feature plots of definitive endoderm (FOXA2 and SOX17), primitive streak (MIXL1), the gastrointestinal epithelium (CDH1 and GPX2), and mesoderm or mesenchyme (BMP4, FOXF1 and PDGFRα) markers in HIO time course and MG data. D) Heatmap shows the transcriptome similarity (Spearman’s correlation coefficients, SCCs) between HIO and MG clusters using highly variable genes in the MG dataset that have expressed human one-to-one orthologs. E) Hierarchical clustering of average transcriptome Pearson’s correlations between the developing and adult intestine and thIO cell type counterparts together with HIO epithelial clusters committed to intestine identity, using cell-type markers identified in either developing or adult duodenum. F) In vitro HIO, thIO, and primary developing duodenum stem cells were ordered by ISC maturity and colored by time point and tissue source. ISC maturity is defined as the fraction of the most mature developing (19 PCW / 132 post-conception days) stem cell identity estimated by quadratic programming. See also STAR methods. G) Distributions of ISC maturity estimated by quadratic programming for each sample ordered by the median. Age of thIO = in vitro period (4 weeks) + post-transplantation period, so as to be comparable with age of HIO, i.e., 4- and 8- week thIO are day 56 and day 84. Boxplots colored by time point and tissue source. H) Heatmap shows molecular transitions during an increase of ISC maturity. Cells are ordered according to ISC maturity. Every 100 cells were grouped into a bin, and the top sidebars show the most frequent time point, average LGR5, and OLFM4 expression within each cell bin. Selected transcription factors (TFs, circles), ligand, and receptor genes are marked next to the heatmap. I) Top, heatmap showing aggregated normalized expression patterns of state 4, 5, or 6 ISC associated regulons across ISC cell bins (column). Bottom, heatmap showing normalized expression patterns of LGR5, OLFM4 and their respective predicted regulator. Top sidebar colored by cell state bins. See also STAR Methods.
Figure S4. Integrating HIO and developing duodenum data unveils transient cell states during human intestine mesenchyme development, related to Figures 4 and 5

A) UMAP embedding of mesenchymal cells from developing human early small intestine and duodenum, with cells colored by type (top left) or age (bottom left). Cell type annotations are shown top right. B) Feature plots show the aggregated expression of ligands of each signaling pathway in developing duodenum mesenchyme. C) Time point distribution of mesenchymal cells on the UMAP embedding, with cells colored by type. D) Stacked bar plot shows the proportion of mesenchymal subtypes at different time points in the developing duodenum, with color scheme consistent with panel A. E) Integrated UMAP embedding of developing duodenum and 4-week THIO intestinal mesenchymal cells. Left, developing duodenum cells in the integrated embedding colored by subtype. Right, THIO cells colored by the most similar developing duodenum subtype: gray cells represent developing duodenum. F) In vitro HIO time course bi-potent cluster (c9 of HIO time course in Figure 5E) and clusters in the mesenchymal trajectory (c5, c3, c10, c1 and c8) were extracted (top left, clusters from original SPRING embedding are highlighted) and projected to a new UMAP embedding to visualize mesenchymal cell developmental trajectories. Cells were sub-clustered and annotated or colored by age (top right). DE: definitive endoderm. G) Relative transcriptome similarity between differentiated intestinal mesenchymal clusters of HIOs (M1-M5) and mesenchymal subtypes observed in 7 PCW developing duodenum. H) Schematic suggests developmental relationships between duodenum mesenchymal subtypes based on transcriptome similarity and coordinated changes in intestinal epithelial stem cell states. Dashed arrows indicate lower probability. I) Top, schematic shows selected HIO mesenchymal cells as precursors to subepithelial mesenchyme, villus core, GDF10+ mesenchyme and smooth muscle cells. Dashed lines represent separation of in vitro HIO and developing duodenum time points. Bottom, dots represent the interaction intensity between intestinal stem cells and specified mesenchymal subtypes, or signaling pathway component expression of specific cell types at each time point. Lines represent interpolated spline curves. Interaction intensity is quantified as the averaged expression of a ligand-receptor gene pair in intestinal epithelial stem cells and each of the specified mesenchymal subtypes. Pathway signaling component expression is calculated as the sum of normalized expression levels of related genes in each cell type.
Figure S5. Supplemental analysis of NRG1 and EGF effects on HIO development. related to Figure 6
A) Brightfield images of HIOs grown in varying concentrations of EGF and/or NRG1 after 40 days of in vitro culture. All culture conditions also contained NOG (100 ng/mL) and RSPO (5% conditioned media). Scale bars represent 2 mm. B) Heatmap shows cluster average expression of marker genes (row) for each organoid cluster (column) with selected genes shown. Same cell type color as in Figure 7B. The dendrogram shows transcriptome distance between clusters calculated using cluster markers. C) (Top) Stacked bar plot showing the probability of projection to each developing organ of the reference atlas for epithelial cells of day 40 in vitro organoids grown in different media. (Bottom) Schematic indicating influence of organoid media component on the distribution of cell fate probability. -, + and ++ represent 0, 1ng/mL and 100ng/mL, respectively. D) Feature plots showing expression of multiple cell type markers in the UMAP embedding of organoids grown in different media conditions. E) Feature plots showing expression of organoid media components in the multi-organ atlas. F) Schematic indicating inferred pathways associated with LGR5 and ASCL2 upregulation in NRG1 treated organoids. Pathway proposed through the integration of expression, regulon, ligand-receptor and KEGG annotation. G) UMAP of mesenchymal cells from organoids grown in different culture conditions, with cells colored by cluster or condition. H) Feature plots showing aggregated expression levels of organ markers in organoid mesenchymal cells. I) Feature plots showing expression of the organ or cell type markers of mesenchyme.
Figure S6. Integrated analysis of culture conditions on HIO epithelial cell states in vitro, related to Figure 6
A) Histological analysis of alginate grown HIOs after 28 days of in vitro culture. (top) H&E (bottom) Immunofluorescent staining for ECAD (white), PDGFRA (green), and aSMA (red) counterstained with DAPI (blue). B-D) UMAP embedding of day 28 in vitro organoids grown in different conditions, with cells colored by condition (B), cell class and cluster annotation (C), or projected developing organ (D). E) Schematic summary of culture conditions of analyzed in vitro HIOs. F) Stacked bar plot showing the probability of projection to the multi-organ atlas for epithelial cells of organoids grown in different conditions. G) UMAP embedding of organoid epithelial cells grown in different culture conditions, with cells colored by projected organ identity (top), or culture condition (bottom). H) Bean plot showing the distribution of ISC maturity score of intestinal organoid epithelial cells grown in different culture conditions, 4 wks thIO, and primary developing intestine (8 PCW, 10 PCW).
Figure S7. Cell composition and differential gene expression analysis of control and CDX2-KO ESC-derived organoids, related to Figure 7
A) Schematic indicates a network of CDX2 predicted target subset based on HIO time course data. See also STAR Methods. B) Bright-field and immunofluorescence images of control and CDX2-KO hPSC-derived organoids. C) Heatmap shows relative cluster average expression levels of marker genes (row) of each cluster (column) of the control and CDX2-KO integrated datasets. Same cluster index as in Figure 7 C. D) Feature plots of cell cluster marker genes on the integrated UMAP embedding shown in Figure 7 C. E) Volcano plot shows differentially expressed genes between CTHRC1+ cells of CDX2-KO organoids and...
SLC2A3+ cells of control organoids. Dark gray indicates genes with significant differences between groups. Yellow and blue indicate the top 10 positive markers of each group. F) Left, UMAP of mouse gastrulation (MG) atlas (Pijuan-Sala et al., 2019), with mesenchymal cell types mapped to integrated control and CDX2-KO datasets highlighted. Right, stacked bar plot shows the proportion of mesenchymal cells mapped to MG cell types. G) MG atlas feature plots of TFs showing differential expression levels between control and CDX2-KO HIO mesenchyme. H) UMAP embedding of in vitro HIO time course mesenchymal progenitor/mesenchymal cells (c1,3,5,8,9,10 from Figure 5E) with cells colored by cell type assignment. Same UMAP embedding as shown in Figure S4F. I) Feature plots of control and CDX2-KO mesenchyme differentially expressed transcription factors (TFs) in the HIO time course mesoderm/mesenchyme cells. J) UMAP embedding of integrated control and CDX2-KO dataset mesenchymal cells colored by most similar HIO mesenchymal subtypes. Expression patterns of CDX2 predicted targets in K) epithelium and L) mesenchyme of CDX2 control and KO data.