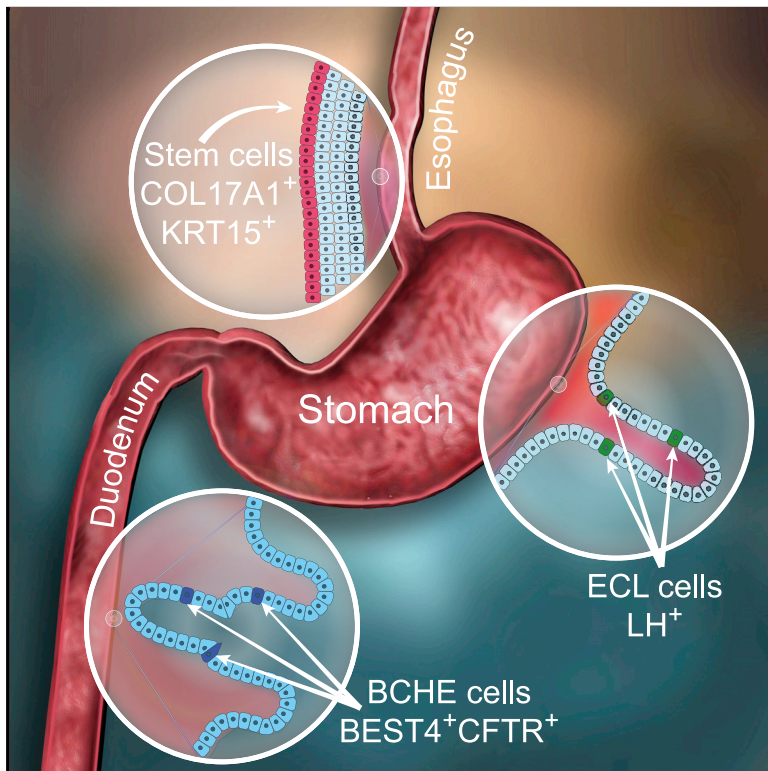**Article**

Human gastrointestinal epithelia of the esophagus, stomach, and duodenum resolved at single-cell resolution

Graphical Abstract

Authors

Georg A. Busslinger, Bas L.A. Weusten, Auke Bogte, Harry Begthel, Lodewijk A.A. Brosens, Hans Clevers

Correspondence

h.clevers@hubrecht.eu

In brief

Busslinger et al. characterize the human epithelia of the esophagus, stomach, and duodenum by single-cell analysis to define the expression signatures of all known and rare uncharacterized cell types. Moreover, they define the expression patterns of transporter genes along the upper gastrointestinal tract.

Highlights

- Expression of COL17A1 and KRT15 identifies esophageal stem/progenitor cells

- Enterochromaffin-like (ECL) cells express the luteinizing hormone (LH)

- Expression of BEST4 and CFTR identifies a rare duodenal cell type called BCHE cells

- Expression patterns of gastric cell types show differences between human and mouse

CellPress

Article

Human gastrointestinal epithelia of the esophagus, stomach, and duodenum resolved at single-cell resolution

Georg A. Busslinger,[1,2,5,6] Bas L.A. Weusten,[3] Auke Bogte,[3] Harry Begthel,[1] Lodewijk A.A. Brosens,[4] and Hans Clevers[1,2,7,*]

[1]Hubrecht Institute and Oncode Institute, Royal Netherlands Academy of Arts and Sciences (KNAW), Utrecht, the Netherlands
[2]Princess Maxima Center for Pediatric Oncology, Utrecht, the Netherlands
[3]Department of Gastroenterology and Hepatology, UMC Utrecht, University of Utrecht, Utrecht, the Netherlands
[4]Department of Pathology, UMC Utrecht, University of Utrecht, Utrecht, the Netherlands
[5]Present address: Research Center for Molecular Medicine (CeMM) of the Austrian Academy of Sciences, Vienna, Austria
[6]Present address: Division of Gastroenterology and Hepatology, Department of Internal Medicine III, Medical University of Vienna, Vienna, Austria
[7]Lead contact
*Correspondence: h.clevers@hubrecht.eu
https://doi.org/10.1016/j.celrep.2021.108819

SUMMARY

The upper gastrointestinal tract, consisting of the esophagus, stomach, and duodenum, controls food transport, digestion, nutrient uptake, and hormone production. By single-cell analysis of healthy epithelia of these human organs, we molecularly define their distinct cell types. We identify a quiescent COL17A1^{high} KRT15^{high} stem/progenitor cell population in the most basal cell layer of the esophagus and detect substantial gene expression differences between identical cell types of the human and mouse stomach. Selective expression of BEST4, CFTR, guanylin, and uroguanylin identifies a rare duodenal cell type, referred to as BCHE cell, which likely mediates high-volume fluid secretion because of continual activation of the CFTR channel by guanylin/uroguanylin-mediated autocrine signaling. Serotonin-producing enterochromaffin cells in the antral stomach significantly differ in gene expression from duodenal enterochromaffin cells. We, furthermore, discover that the histamine-producing enterochromaffin-like cells in the oxyntic stomach express the luteinizing hormone, yet another member of the enteroendocrine hormone family.

INTRODUCTION

The main functions of the gastrointestinal (GI) tract are food digestion, nutrient uptake, microbe defense, and hormone production. The GI tract consists of multiple organs that act together in a coordinated manner and that are lined by specialized epithelia (Thompson et al., 2018). A thick, stratified squamous epithelium protects the esophagus from abrasion by the incoming raw food (Zhang et al., 2017). In contrast, the stomach, small intestine, and colon consist of single-layered epithelia, which are involved in the secretion of hydrochloric acid and pre-digestive enzymes and in nutrient absorption (Clevers, 2013; Gehart and Clevers, 2019; Kim and Shivdasani, 2016; Willet and Mills, 2016). The secretory and absorptive capacities of gastrointestinal epithelia are largely determined by transmembrane channels and transporters, most of which belong to the family of solute carrier proteins (Hediger et al., 2013). The gastrointestinal epithelia also contain dispersed and rare (1%) enteroendocrine cells (EECs), which together, constitute the largest endocrine system of humans. These cells sense the intestinal content and release distinct hormones to regulate the gastrointestinal activity, systemic metabolism, and food intake

(Gribble and Reimann, 2017). An important function of the stomach is the secretion of gastric acid, which not only initiates the digestive process but also functions as a first line of defense against food-borne microbes (Hunt et al., 2015). Malfunctioning of gastric acid secretion results in colonization with undesired microbes and contributes to tumor formation (Hunt et al., 2015). Excessive water secretion or impaired absorption causes secretory diarrhea (Thiagarajah et al., 2015), whereas insufficient hydration leads to obstruction of the GI tract, as manifested in patients with cystic fibrosis (van der Doef et al., 2011).

Advances in single-cell RNA sequencing (scRNA-seq) facilitate analysis of the distinct expression patterns of individual cell types constituting the GI epithelia. Recent scRNA-seq studies analyzed the mouse small intestine (Haber et al., 2017), human fetal digestive tract (Gao et al., 2018), human colon in different disease stages (Parikh et al., 2019; Smillie et al., 2019), and the human ileum, colon, and rectum (Wang et al., 2020). However, systematic scRNA-seq analysis of the human upper gastrointestinal (UGI) organs has not yet, to our knowledge, been performed.

Here, we analyzed human biopsies from the healthy esophagus, three different gastric regions, and the duodenum, by
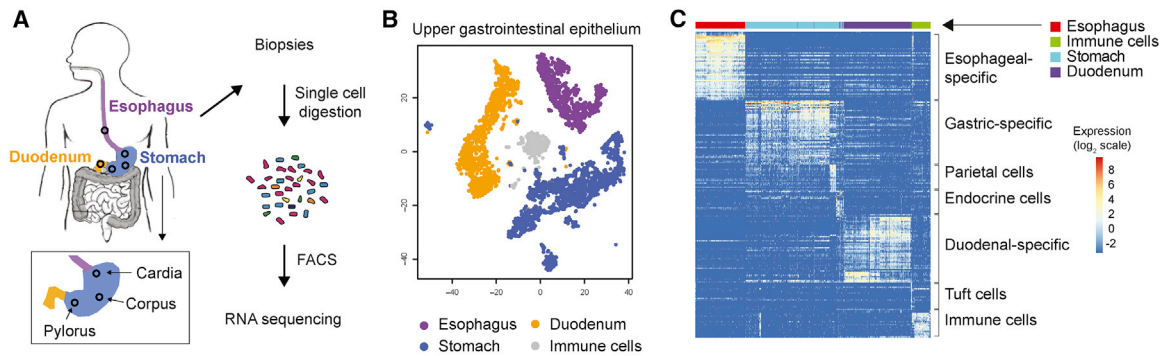
Figure 1. Analysis of the human upper gastrointestinal epithelium by scRNA-seq

(A) Experimental setup. A schematic diagram of the upper gastrointestinal tract is shown, together with the anatomical locations (circles) at which biopsies were taken. Single cells of digested biopsies were isolated by fluorescence-activated cell sorting (FACS), followed by scRNA-seq analysis.

(B) t-SNE map of all combined scRNA-seq data displaying separate cell clusters for the esophagus (purple), stomach (blue), duodenum (orange), and immune cells (gray).

(C) Heatmap displaying the expression of the most differentially expressed genes, clustered according to their expression in the three UGI organs and parietal, endocrine, tuft, and immune cells.

single-cell analysis, which identified rare cell types and provided insights into the cellular composition and gene expression patterns along the UGI epithelium under normal homeostasis.

RESULTS

Biopsies of healthy esophageal squamous, gastric glandular and duodenal crypt, and villus epithelia were collected from patients who were under endoscopic surveillance for Barrett's esophagus (Table S1; Method details). Epithelial cells were isolated and processed for scRNA-seq using the sorting and robot-assisted transcriptome sequencing (SORT-seq) protocol (Muraro et al., 2016), followed by data analysis with the RaceID algorithm (Grün et al., 2015; Figures 1A and S1; Method details). For each UGI organ, we separately analyzed biopsies from at least three patients (Figure S1), and the number of high-quality cells analyzed after bioinformatic filtering are shown for each biopsy and organ in Table S2 (see Method details). Because most cells were assigned to organ-specific clusters with unique gene expression profiles (Figures 1B and 1C), we focused our initial analysis on the cellular composition of each organ separately.

Identification of COL17A1^{high} KRT15^{high} stem/progenitor cells in the basal cell layer of the human esophagus

Esophageal cells were largely grouped together (Figures 2A and S1A; Tables S3 and S4), except for one small cluster of immune cells with dendritic cell characteristics (Figure 2B). The esophageal cells constituted a continuum of cells that could be categorized as basal cells or early, mid, and late suprabasal cells, based on the expression of known marker proteins, such as the keratins KRT4, KRT13, and KRT14 (Figures 2A–2C). Unexpectedly, KRT5 exhibited a broad expression pattern in the esophageal epithelium with highest levels in early suprabasal cells as shown by scRNA-seq (Figure 2C) and immunohistochemical analysis (Figure 2D), whereas its expression is strictly confined to basal cells in the skin (Lersch and Fuchs, 1988). Notably, we detected two sub-

populations within the basal layers, a proliferative one, as indicated by expression of the cell-cycle-regulated genes MKI67 and TOP2A, and a quiescent one with high expression of ZFP36L2 (Figure 2C), a gene associated with cell cycle arrest (Galloway et al., 2016). The quiescent basal cells were further enriched for expression of genes, including COL17A1, COL7A1, and DST (Figure 2C). COL17A1 and DST are components of the hemidesmosome, which is required for the attachment of epidermal cells to the basal lamina in the skin (Künzli et al., 2016; Natsuga et al., 2019) and which is known to interact with integrins and laminins that were also expressed by the same quiescent basal cell population (ITGA3, ITGB4, and LAMB3; Figure S2A). Recently, Col17a1 expression was associated with stem cell fitness in the mouse skin epithelium (Liu et al., 2019), and Krt15 expression was reported to mark long-lived esophageal progenitors of the mouse (Giroux et al., 2017). Importantly, both KRT15 and COL17A1 were most highly expressed in the quiescent basal cell population in our data (Figure 2C), which suggests that these cells may function as stem/progenitor cells in the human esophagus. Immunohistochemical analyses confirmed strong expression of COL17A1 and KRT15 in the same cells of the most basal cell layer of the esophagus (Figures 2D and S2B). As revealed by co-staining, the proliferation marker MKI67 was expressed in only 4.7% and 4.2% of the COL17A1^{high} and KRT15^{high} basal cells, respectively (Figures 2D and 2E). Hence, COL17A1 and KRT15 are predominantly expressed in quiescent esophageal stem/progenitor cells, whereas the presence of a few proliferating COL17A1^{high} KRT15^{high} cells is consistent with the reported asymmetric division of human esophageal stem cells giving rise to differentiating keratinocyte (Seery and Watt, 2000). The quiescent basal cells were further characterized by high expression of the NOTCH ligands, DLL1 and JAG2, whereas the NOTCH3 receptor was expressed in the adjacent proliferating basal and suprabasal cells (Figure 2F). Moreover, WNT4, WNT5B, WNT10A, and DKK3, encoding a WNT antagonist, were preferentially expressed in the quiescent basal cells, whereas WNT5A expression was detected in differentiated
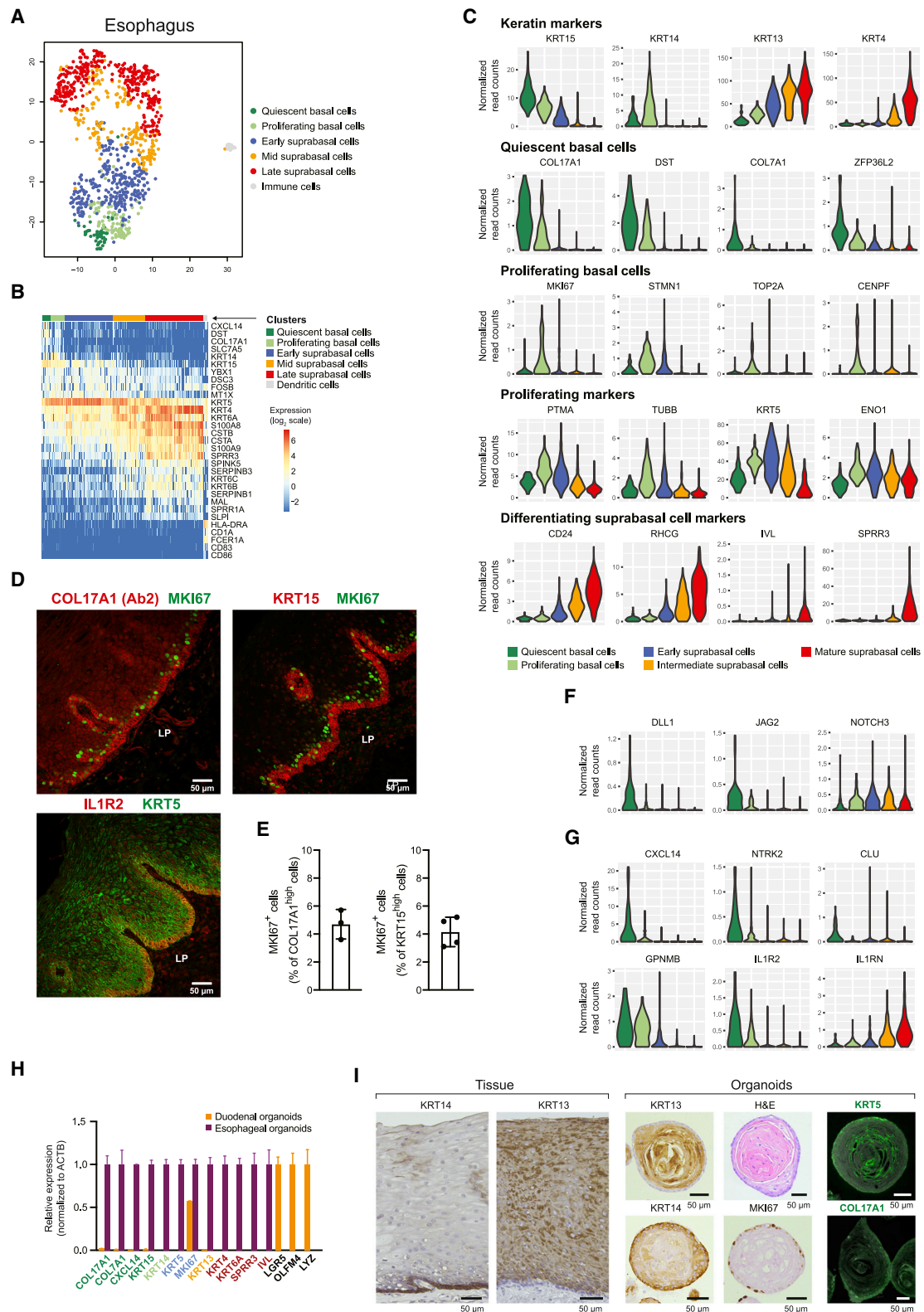
Figure 2. Identification of quiescent KRT15high COL17A1high stem/progenitor cells in the basal cell layer of the human esophagus

(A) t-SNE map displaying the scRNA-seq data of the esophagus (Tables S3 and S4). The distinct cell types are indicated in different colors.

(B) Heatmap of the scRNA-seq expression data of selected genes from the six clusters corresponding to the different esophageal cell types.

(legend continued on next page)

keratinocytes (Figure S2C). The quiescent basal cells were also enriched for expression of the transcription factor SOX6 (Figure S2C) and clusterin (CLU; Figure 2G), a marker of quiescent stem cells in the mouse intestine (Ayyaz et al., 2019). Together, these data suggest that the COL17A1^{high} KRT15^{high} cells constitute a quiescent stem/progenitor cell population in the most basal layer of the human esophagus. To further investigate this, we used the StemID program (Grün et al., 2016) to calculate a lineage tree for the esophageal cell clusters, which revealed a linear progression from the quiescent stem cells to the terminally differentiated keratinocytes (Figure S2E).

The quiescent basal cells also expressed decorin (DCN; Figure S2D), a secreted proteoglycan, and the neurotrophic receptor tyrosine kinase 2 (NRTK2; Figure 2G), which is activated by the brain-derived neurotrophic factor and neurotrophin 4 (Cocco et al., 2018). These cells were also enriched for expression of CXCL14, GPNMB, and the interleukin-1 receptor 2 (IL1R2) (Figure 2G), which was confirmed by immunohistochemical analysis (Figure 2D). CXCL14, a chemoattractant for immature monocytes, and GPNMB are likely involved in macrophage recruitment and clearance of cellular debris (Lu et al., 2016; van der Lienden et al., 2018). IL1R2 functions as a secreted decoy receptor for IL-1 to antagonize IL-1-induced inflammation (Peters et al., 2013). Of note, an inverse correlation was found between the expression of IL1R2 and IL1RN, another secreted IL-1 receptor antagonist, within the esophageal epithelium (Figure 2G). Hence, IL-1-induced inflammation may be restrained by different antagonists in the basal and suprabasal cell layers.

We also established organoid cultures from healthy esophageal biopsies and confirmed the esophageal origin of those organoids by qRT-PCR analysis of esophageal-specific gene expression (Figure 2H). The esophageal organoids, which could be maintained *in vitro* for up to 2 months (data not shown), recapitulated the general architecture of the esophageal epithelium, as shown by immunohistochemical staining of KRT13, KRT14, KRT5, COL17A1, and MKI67 (Figure 2I).

In summary, these analyses identified and characterized the COL17A1^{high} KRT15^{high} basal cells as stem/progenitor cells of the human esophageal epithelium.

Gene expression differences in gastric cell types between human and mouse

The human stomach is subdivided into the cardia, corpus, and pylorus regions. Its single-layered epithelium is composed of glandular units, which are referred to as oxyntic or antral glands depending on their anatomic location in either the cardia and corpus or the pylorus region, respectively (Willet and Mills, 2016; Figure S3A). These gastric glands are compartmentalized into subdomains consisting of the surface pit, isthmus, and neck and base regions (Figure S3A). Our human gastric single-cell atlas, derived from biopsies of the cardia, corpus, and pylorus regions (Figures S1B–S1F), identified all known cell types, including the mucus-secreting pit cells (GKN1 and GKN2), acid-secreting parietal cells (ATP4A and ATP4B), proliferating isthmus cells (MKI67 and STMN1), neck cells (MUC6), chief cells (LIPF and PGA5), endocrine cells (GAST, GHRL, SST, and HDC), and rare tuft cells (TRPM5 and SH2D6), based on expression of the respective signature genes (Figures 3A, 3B, and 3D; Tables S5 and S6). *In situ* RNA hybridization on histological resection specimens validated the specific expression of *SLC5A5*, *GHRL*, *MUC6*, and *LIPF* in pit, endocrine, neck, and chief cells, respectively (Figures 3C and S3B).

The cells of the oxyntic glands from the cardia and corpus regions clustered together, whereas cells of the antral glands from the pylorus region largely formed a separate cluster (Figure 3A). The main cell types of the oxyntic glands compared with the antral glands were the chief cells (13.4% [oxyntic] versus 1% [antral]), prezymogenic cells (16% versus 1%), and parietal cells (11.2% versus 0.9%), whereas the main cells types of the antral glands were the neck cells (1.7% [oxyntic] versus 11.7% [antral]), isthmus cells (0.5% versus 21.5%), as well as immature (3.2% versus 18.2%) and mature (2% versus 7.2%) pit cells (Table S5). The REG3A$^+$ cells expressing the antimicrobial genes REG3A, LTF, and LCN2 (Figure S3C; Table S5) were present only in the oxyntic glands (Figure 3A). In contrast, the antimicrobial-active lysozyme (LYZ)-expressing cells, which also expressed *PPP1R1B*, *GP2*, and *AQP5*, represented a large fraction (12%) of the antral glands (Figures 3D and 3E; Table S5). The abundant isthmus cells in the antral grand, defined by *MKI67*, *STMN1*, and *TOP2A* transcripts, also expressed *NMU* (Figure 3F), encoding the neuropeptide neuromedin U, which has pleiotropic roles in controlling gastric acid secretion, gastric emptying, intestinal motility, and feeding behavior (Martinez and O'Driscoll, 2015).

The corpus and pylorus regions are known to be anatomically comparable between human and mouse (Willet and Mills, 2016). However, it is not known how similar or different the gene expression patterns of the distinct gastric cell types are between the human and mouse, except for the known difference that the gastric intrinsic factor (GIF) is specifically expressed in chief cells

---

(C) Violin plots showing the expression of selected marker genes indicative of the distinct cell types, which are shown by different colors. Gene expression is shown as normalized read counts on the y axis, whereas the numbers on the x axis refer to the different cell types, as indicated by the key shown below.

(D) Immunohistochemical analysis of COL17A1, IL1R2, KRT15, KRT5, and MKI67 expression in human esophageal resection specimens. LP, lamina propria.

(E) Statistical evaluation of MKI67 expression in COL17A1^{high} KRT15^{high} basal cells. The percentage of MKI67$^+$ cells in COL17A1^{high} and KRT15^{high} cells is shown as mean value with SEM. Dots correspond to the average value of several sections from three or four patients analyzed (see Method details).

(F and G) Violin plots displaying the expression of DLL1, JAG2, and NOTCH3 (F) genes, with enriched expression in quiescent basal cells and IL1RN (G).

(H) Expression analysis of the indicated marker genes in esophageal (purple) and duodenal (orange) organoids (bottom), as determined by qRT-PCR analysis relative to the *ACTB* mRNA. The maximal expression of each gene in the two different organoids was set to 1, and the expression data are shown as mean values with SEM (n = 3). The genes are colored according to cell type, in which their expression is enriched based on the esophageal scRNA-seq data. For color code, see (A).

(I) Immunohistochemical staining of esophageal tissue and *in vitro* cultured esophageal organoid for expression of KRT5, KRT13, KRT14, COL17A1, and MKI67. H&E, staining with hematoxylin and eosin.

The scale bars (D and I) indicate 50 μm.
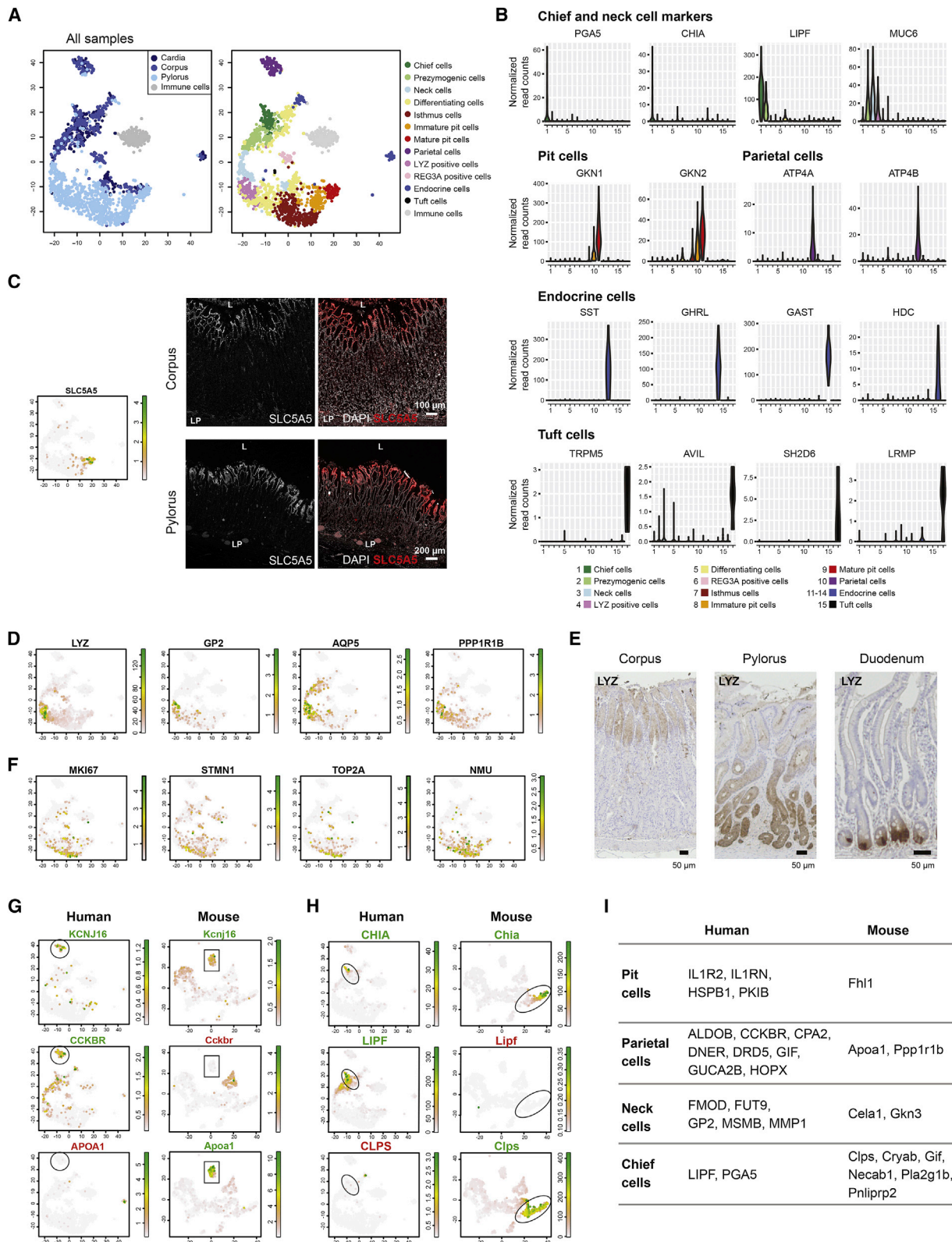
of the mouse and in parietal cells of humans (Levine et al., 1980). The stomach of the mouse has, however, not yet, to our knowledge, been analyzed by single-cell analysis. To compare the human and mouse gastric epithelia, we dissected the corpus and pylorus regions from adult mice to perform scRNA-seq by the SORT-seq method (Figures S1H and S4A–S4C; Tables S7 and S8). All known gastric cell types of the mouse, including the rare tuft cells, were identified by expression of their cell-type-specific marker genes (Figure S4; Table S7).

Next, we compared the expression profiles of chief, neck, parietal, and pit cells between the mouse and human datasets (Figures S5 and S6). To that end, we used a systematic filtering procedure to select for genes with enriched expression in a given cell type compared with the remaining cells of the stomach in each species (Table S9; Method details). We then analyzed the t-distributed stochastic neighbor embedding (t-SNE) maps of the enriched genes to determine their differential expression in the given cell type between the human and mouse stomach, as shown in Figures 3G, 3H, S5, and S6. As expected, this analysis revealed that the four gastric cell types expressed many genes in a similar manner in humans and mice (Figures 3G, 3H, S5, and S6; Table S9). However, the expression of several genes could be detected in pit, parietal, neck, or chief cells only in the human or mouse stomach (Table S9), and the expression differences of representative genes are shown in Figures 3G–3I, S5, and S6. Hence, in addition to the *GIF* gene, there are many more gene expression differences among the respective gastric cell types of humans and mice.

Molecular characterization of a rare CFTR[high] BEST4[high] cell type in the human duodenum

To investigate the small intestine, we used the standard endoscopic procedure to obtain biopsies from the most proximal small intestinal region, the duodenum. We separated the crypt and villus regions of each biopsy to enrich for the respective cell populations, which were then independently analyzed by scRNA-seq (Figure S1G; Method details). All main cell types were identified, which included Paneth cells, goblet cells, tuft cells, enterocytes, endocrine lineage cells and stem cells expressing *LGR5*, *ASCL2*, *OLFM4* and *RNF43* (Figures 4A, 4B, S7A, and S7B; Tables S10 and S11). Notably, the stem cell-supporting Paneth cells did not express niche factors of the WNT family (Figure S7C), contrary to the expression of Wnt3 and Wnt11 in mouse Paneth cells (Sato et al., 2011b). However,

expression of the NOTCH ligand DLL1 was detected in Paneth and goblet cells, whereas expression of the corresponding NOTCH1 receptor was observed in stem cells (Figure S7D). By performing immunohistochemical or *in situ* RNA hybridization analyses on histological resection specimens, we validated the expression of LYZ in Paneth cells (Figure 3F); the expression of *ASCL2* in stem cells; *DMBT1* in transit-amplifying cells; *KRT20*, *ANPEP*, and *CEACAM6* in enterocytes; as well as *REG4* in endocrine cells (Figure 4C).

Interestingly, we identified a small cell cluster (19 cells, 1.3% of all epithelial cells) that exhibited high expression of *BEST4*, *CFTR*, *GUCA2A*, and *GUCA2B* (Figures 4D–4F). We refer to this cluster as BEST4, CFTR high-expressor (BCHE) cells. Among all duodenal cells, the BCHE cells expressed the highest level of the CFTR (Figures 4D–4F), which functions as a chloride/bicarbonate transport channel (Moran, 2017) that, upon mutation, is known to cause cystic fibrosis, a disease associated with frequent mucosal infections because of the production of thick and sticky mucus (O'Neal and Knowles, 2018). We confirmed the observed co-expression of *CFTR*, *BEST4*, and *GUCA2B* in BCHE cells by *in situ* RNA hybridization (Figure 4E). *BEST4* (bestrophin 4) codes for a calcium-sensitive chloride channel (Hartzell et al., 2008), and *GUCA2A* and *GUCA2B* code for the small peptides guanylin and uroguanylin (Sindić and Schlatter, 2006; Rappaport and Waldman, 2018), both of which bind to the ubiquitously expressed receptor GUCY2C (Figure S7E). Additional genes with preferential expression in BCHE cells (Figure S7F) code for the cell surface proteins GPR112, CD24, and TMEM219 as well as the cystic fibrosis modifiers FAM13A and GSTP1 (Corvol et al., 2018; de Lima Marson et al., 2013) and gelsolin (GSN), a regulator of actin filament assembly, which is implicated in the trafficking and surface expression of CFTR (Edelman, 2014).

BEST4[+] cells were identified by scRNA-seq in the human colon (Parikh et al., 2019; Smillie et al., 2019). These colonic BEST4[+] cells and the duodenal BCHE cells share the expression of some genes including *BEST4*, *GUCA2A*, *GUCA2B*, and *OTOP2*, although GUCA2A and GUCA2B are broadly expressed in many colonic enterocytes in contrast to their BCHE-cell-specific expression in the duodenum (Figures 4F and S7G). Importantly, however, expression of the *CFTR*, *KRT72*, and *NAPSB* genes could be detected only in the duodenal BCHE cells, whereas OTOP3 expression was observed specifically in the colonic BEST4[+] cells (Figures 4F and S7G). The difference in

Figure 3. Single-cell analysis of the human stomach

(A) t-SNE map displaying the scRNA-seq data of all gastric cell types (Tables S5 and S6). Epithelial cells are colored according to their anatomic locations in the stomach (left) or according to the assigned cell types (right).

(B) Expression of representative genes in the indicated gastric cell types, as shown by violin plots (see Figure 2C).

(C) Validation of the correct assignment of *SLC5A5* expression in pit cells, as shown by t-SNE map (left) and *in situ* RNA hybridization (RNAScope) on corpus or pylorus resection specimens (right). L, lumen; lamina propria, LP.

(D) Co-expression of *GP2*, *AQP5*, and *PPP1R1B* in lysozyme (*LYZ*)-secreting cells, as displayed by t-SNE maps.

(E) Immunohistochemical staining of LYZ expression on resection specimens from the gastric corpus, pylorus, or duodenum.

(F) Co-expression of *MKI67*, *STMN1*, *TOP2A*, and *NMU* in isthmus cells, as shown by t-SNE maps.

(G and H) Common and distinct expression patterns of three selected genes in parietal (G) and chief (H) cells of the human and mouse stomachs. The positions of the parietal cell cluster on the t-SNE maps of the human and mouse stomach are shown by circles and squares, respectively (G), whereas the chief cell clusters are indicated by ovals (H). Genes with or without expression in parietal (G) or chief (H) cells are shown in green or red, respectively.

(I) List of genes that exhibit strongly differential gene expression in pit, parietal, neck, or chief cells between humans and mice. The expression pattern of each gene in the human and mouse gastric epithelium is shown in Figures 3G, 3H, S5, and S6.
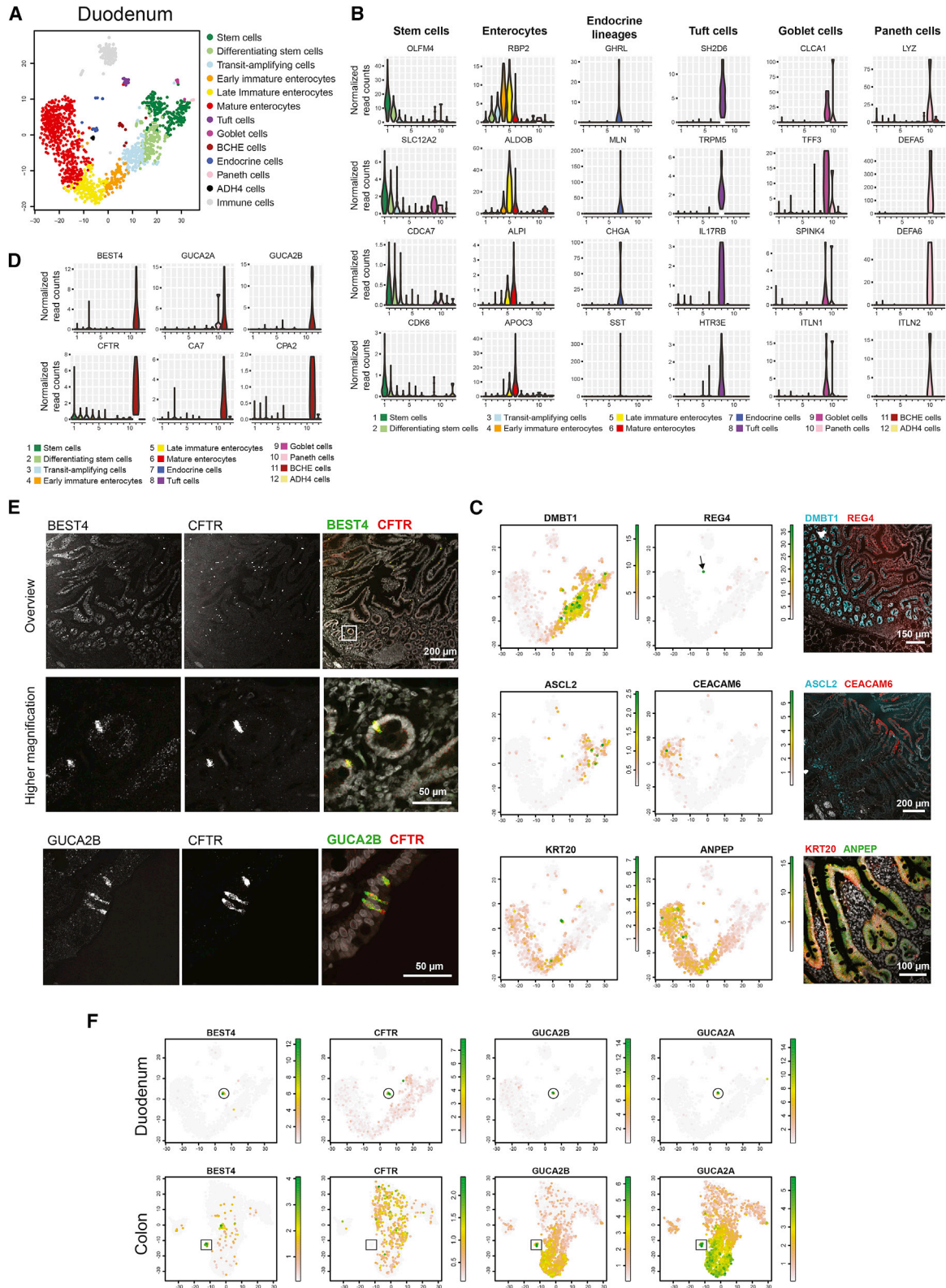
(legend on next page)

CFTR expression strongly suggests that the colonic BEST4+ cells and duodenal BCHE cells may have distinct functions. In addition to the BCHE cells, we detect another small cell cluster characterized by the expression of ADH4 (alcohol dehydrogenase 4) by scRNA-seq (Figures 4A and S7H; Table S10) and immunohistochemical analysis of the human duodenal epithelium (www.proteinatlas.org; Figure S7I).

Together, these single-cell analyses highlighted the absence of WNT expression in human Paneth cells and defined the gene expression signature of the unique BCHE cells in the human duodenum.

Molecular definition of the enteroendocrine cell types in the human stomach and duodenum

The rare EECs consist of multiple subtypes, which are defined by their primary hormone products or hormone-producing enzymes (Gribble and Reimann, 2017). The stomach and duodenum contain D cells (SST), G cells (GAST), X cells (GHRL), MX cells (MLN and GHRL), enterochromaffin (EC) cells (CHGA and REG4), and enterochromaffin-like (ECL) cells (CHGA and HDC) (Engelstoft et al., 2015; Gehart et al., 2019; Beumer et al., 2020). Our single-cell atlas of the human UGI tract offered the perfect opportunity to compare the gastric and duodenal EECs. To identify common and cell type-specific endocrine gene signatures, we combined, and then separately analyzed all EEC cell types, from which we obtained sequencing data from three (MX) or more (all others) cells per cell type (Figure 5A; Table S12). EECs from the stomach and duodenum shared a common endocrine expression signature consisting of 30 genes (Figure 5B). Those genes code for 13 proteins associated with hormone production and maturation such as secretogranins (CHGB, SCG3, SCG5, and SCGN) and proteolytic enzymes (BACE1, CPE, PCSK1, and PCSK1N); 10 transcription factors, many of which have known functions in endocrine cell differentiation such as ARX, INSM1, NEUROD1, NKX2-2, PROX1, and RFX6 (Gehart et al., 2019); and eight transmembrane proteins, mainly consisting of Ca- and K-voltage-gated channels (KCNMA1, KCTD12, and CACNA1 family members) (Figure 5B). We also identified a gastric-enriched expression signature comprising 19 genes, including the genes encoding the enteroendocrine transcription factor RUNX1T1 (Gehart et al., 2019), glutamine transporter SLC38A3, and the glutamate receptor GRIA2, suggesting that the gastric EECs may be responsive to glutamate stimuli.

Next, we searched for EEC subtype-enriched gene expression signatures in the ghrelin (GHRL)-producing gastric X cells and duodenal MX cells (Figure 5C). Both cell types share the expression of 18 genes, including *ACSL1* encoding the long-chain fatty-acid-coenzyme A (CoA) ligase required for acetyl modification of GHRL (Bando et al., 2016; Beumer et al., 2020) and *NPY1R* encoding the receptor for the neuropeptides NPY and PPY that may modulate the activity of both X and MX cells (Figure 5C). The gastric X cell-enriched expression signature consisted of 10 genes, which included the genes encoding CFC1, a member of the epidermal growth factor family, and the free fatty acid receptor FFAR3 sensing short fatty acids in the gut lumen (Figure 5C). As shown by *in situ* RNA hybridization, the *CFC1* gene was specifically co-expressed with *GHRL* in the corpus, but not in the duodenum, thus validating the correct assignment of the *CFC1* gene to the expression signature of the gastric X cells (Figure 5E).
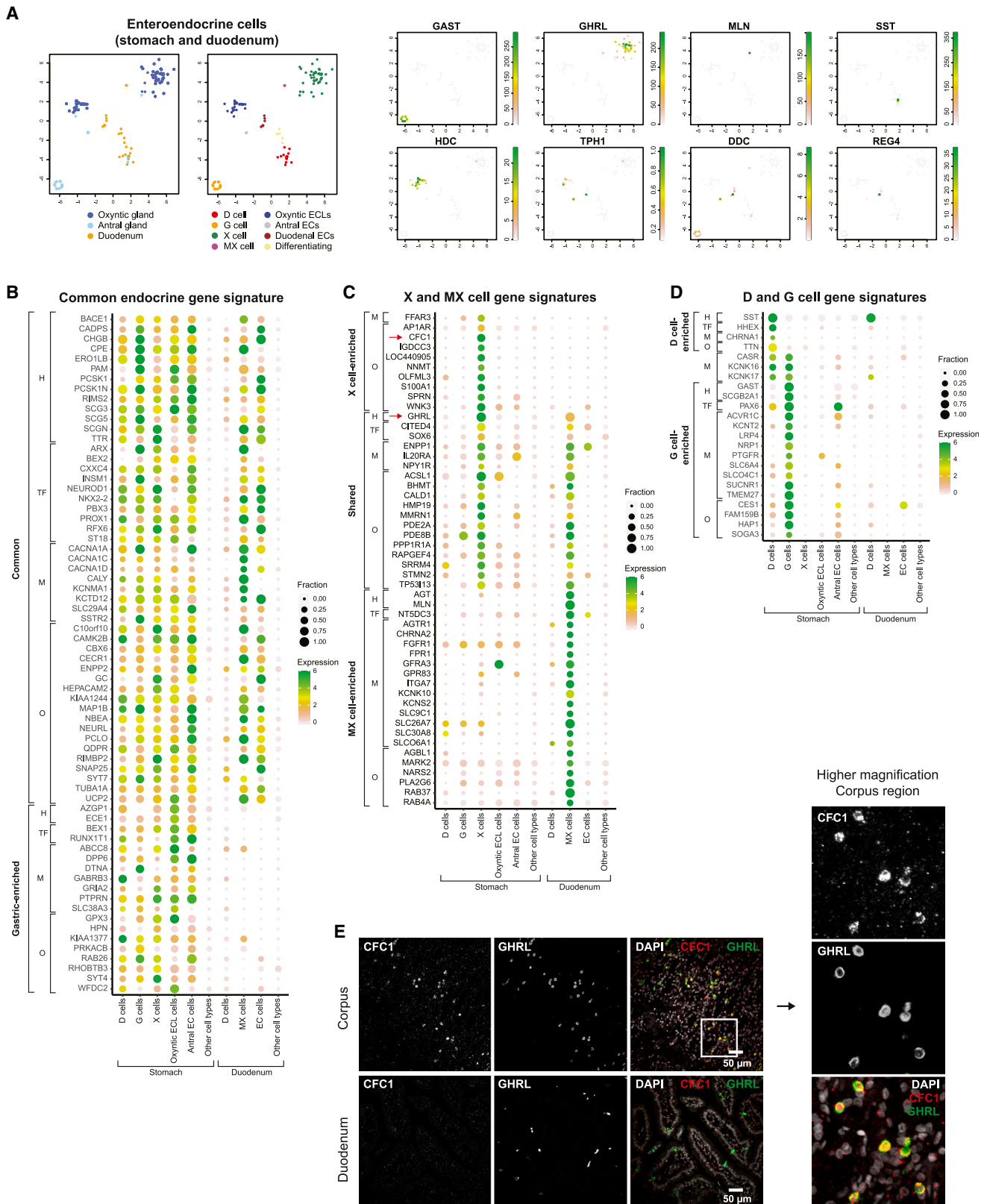
Twenty-two genes defined the duodenal MX cell-specific expression signature (Figure 5C). In addition to the hormone GHRL, the MX cells also expressed high levels of motilin (MLN) and angiotensinogen (AGT; the angiotensin II precursor protein), both of which are known to induce contraction of the human intestinal wall musculature (Ewert et al., 2006; Worthington et al., 2018). Interestingly, co-expression of the angiotensin II type 1 receptor (AGTR1) suggests that the MX cells may be under autocrine feedback control by angiotensin II (Figure 5C). GPR83, the receptor for the neuropeptide PEN, has been implicated in the control of food uptake but was so far shown to be expressed in the brain and thymus (Mack et al., 2019). The specific expression of GPR83 in MX cells now suggests a direct role of this receptor in the duodenum upon interaction with PEN, which may be secreted by adjacent neurons.

Somatostatin (SST), which was most highly and specifically expressed by D cells in the stomach and duodenum (Figures 5A and 5D), is a master regulator preventing hormone release in other EECs (Corleto, 2010). Consequently, expression of the respective receptor SSTR2 was detected in all EECs, with the highest levels being observed in G cells (Figure 5B). In addition to gastrin, the G cells in the antral region of the stomach (Table S5) highly expressed the transformin growth factor β receptor ACVR1C, the prostaglandin F receptor (PTGFR), and the serotonin transporter SLC6A4 (Figure 5D). Once gastrin is secreted by G cells, it binds to its cognate receptor CCKBR on oxyntic ECL cells (Figure 6A), where it promotes the synthesis and secretion of histamine (Prinz et al., 2003).

The duodenal EC cells are known to produce and secrete serotonin (Gershon, 2013), whereas the gastric ECL cells are the site of histamine production and secretion within the

Figure 4. Molecular characterization of a rare CFTR^high BEST4^high cell type in the human duodenum

(A) t-SNE map displaying the scRNA-seq data of the duodenum (Tables S10 and S11). Distinct cell types are shown in different colors, as indicated.

(B) Expression of genes that are representative of the indicated duodenal cell types, as shown by violin plots (see Figure 2C).

(C) Validation of the cell-type-specific expression of the indicated genes, as demonstrated by t-SNE maps (left) and *in situ* RNA staining (RNAScope) on duodenal resection specimens (right). An arrow marks the rare REG4+ cell cluster.

(D) Selective expression of the indicated genes in BCHE cells, as shown by violin plots.

(E) Co-expression of *BEST4*, *CFTR*, and *GUCA2B* in duodenal BCHE cells, as revealed by *in situ* RNA staining on duodenum resection specimens. Co-localization of *BEST4* and *CFTR* expression is shown at low (top) and high (middle) magnification and by an overlay of both stainings (right). Co-expression of *GUCA2B* and *CFTR* in BCHE cells is shown at high magnification (bottom).

(F) Expression of *BEST4*, *CFTR*, *GUCA2A*, and *GUCA2B* in the duodenum (top, our data) and healthy human colon (bottom, data published by Parikh et al. [2019]). Circles or squares denote the positions of the duodenal BCHE cells or colonic BEST4+ cells on the respective t-SNE maps. The scRNA-seq data of the human colon were analyzed with the RaceID program.

A Enteroendocrine cells (stomach and duodenum)

Oxyntic gland / Antral gland / Duodenum

D cell / G cell / X cell / MX cell / Oxyntic ECLs / Antral ECs / Duodenal ECs / Differentiating

GAST / GHRL / MLN / SST / HDC / TPH1 / DDC / REG4

B Common endocrine gene signature

C X and MX cell gene signatures

D D and G cell gene signatures

E

Higher magnification Corpus region

(legend on next page)

gastrointestinal epithelium (Prinz et al., 2003). Rare, serotonin-expressing cells have also been detected in the human stomach with a particular enrichment in the antral glands (Choi et al., 2014; Fakhry et al., 2019; Ito et al., 2013). The EC and ECL cells are both defined by high expression of chromogranin A (CHGA) (Figure 6A). Interestingly, the reanalysis of all EECs revealed three distinct clusters for ECL cells in oxyntic glands, EC cells in antral glands, and the duodenal EC cells (Figure 5A). Consistent with this finding, all three cell types were characterized by different gene expression signatures (Figure 6A). The oxyntic ECL cells expressed high levels of histidine decarboxylase (HDC), the enzyme converting histidine to histamine, as well as the histamine transporter SLC18A2 (Figures 5A and 6A), consistent with the fact that these cells are the primary source of histamine synthesis in the stomach. In contrast, the antral EC cells, like the duodenal EC cells, expressed high levels of tryptophan hydroxylase 1 (TPH1) and dopa decarboxylase (DDC), the two enzymes that generate serotonin from tryptophan (Figures 5A and 6A).

The expression signature of the oxyntic ECL cells consisted of 20 genes, including the genes coding for the gastrin receptor CCKBR, regulating histamine synthesis (Prinz et al., 2003), the purinergic receptor P2RY14, involved in histamine release (Gao and Jacobson, 2017), the alpha (CGA) and beta (LHB) subunits of the luteinizing hormone and the transcription factor PTF1A (Figure 6A). As shown by *in situ* RNA hybridization, the *LHB* and *PTF1A* genes were co-expressed with *CHGA* in the corpus but not in the pylorus or duodenum, thus confirming the correct identification of both genes as part of the expression signature of oxyntic ECL cells (Figures 6B and 6D). The serotonin-producing antral and duodenal EC cells differed substantially in their expression pattern because 27 genes were specifically expressed in the antral EC cells, whereas the duodenal EC cells revealed strongly enriched expression of 15 genes (Figure 6A), for instance, coding for REG4 (Grün et al., 2015) and the adhesion receptor GPR112 (Figure 6A). *In situ* RNA hybridization revealed co-expression of *GPR112* and *CHGA* as well as specific expression of *REG4* in the duodenum but not in the pylorus or corpus, thus confirming the correct assignment of these genes to the expression signature of the duodenal EC cells (Figures 6D and 6E).

In summary, the single-cell analysis of gastric and duodenal EECs revealed common and cell-type-specific gene expression signatures and identified MX cells as a potential target of PEN action and oxyntic ECL cells as producers of the luteinizing hormone. Importantly, the molecular definition of the serotonin-pro-

ducing EC cells in the antral gland revealed that these cells significantly differ in their gene expression pattern from the duodenal EC cells.

Organ- and cell-type-specific expression of transporter genes along the upper gastrointestinal tract

The composition of the luminal content in the UGI tract is modified by the secretory and absorptive capacities of epithelial cells. This is mainly achieved by the expression of transmembrane channels and transporters, which are encoded by about 10% of the entire coding human genome (Hediger et al., 2013). Solute carrier (SLC) proteins and ABC transporters constitute the largest group of transporters, although only a small number of them has been molecularly and functionally characterized (César-Razquin et al., 2015). As the expression of transporter genes has not yet been systematically analyzed for any organ, our scRNA-seq data offered the opportunity to investigate the organ-specific expression patterns of transporters along the human UGI tract at single-cell resolution. Because SLC5A1, SLC11A2, SLC15A1, SLC16A1, SLC46A1, and SLCO2B1 are well-characterized transporters expressed in the small intestine (Thwaites and Anderson, 2007), we first analyzed their expression in our dataset. We could demonstrate duodenal-specific expression of *SLC5A1*, *SLC15A1*, and *SLC46A1*, whereas *SLC11A2*, *SLC16A1*, and *SLCO2B1* displayed broader expression patterns in the UGI tract (Figures S8A and S8B). This prompted us to investigate which transporters were preferentially expressed in one of the three UGI organs (Figures 7A and S8C–S8E). This analysis identified 16 esophageal-, 8 gastric-, and 29 duodenal-enriched transporters (Figures 7A and S8C–S8E) with different substrate specificities (Table S13; César-Razquin et al., 2015), which provides a systematic description of transporter genes with selective expression in the three UGI organs.

We next identified transporter genes that exhibited cell-type-enriched expression within each organ of the UGI tract (Figures 7B–7D and S8F–S8N; Table S13). First, we focused on transporter genes with expression in surface epithelial cells. Esophageal suprabasal cells mainly expressed transporters with exporting properties such as SLC16A6, SLC16A9, SLC24A3, SLC42A3, and SLCO2B1 (Figure 7B), indicating that these cells have little absorptive capacity, as expected for the esophagus. Gastric pit cells expressed the iodide transporter SLC5A5, cystine, and glutamate importer SLC7A11 and the prostaglandin exporter SLCO2A1 (Figures 3C and 7C). Consistent with the known, highly absorptive potential of the duodenal epithelium, the most SLCs were detected in enterocytes, which include importers for

---

Figure 5. Comparison of gene expression patterns between gastric and duodenal endocrine lineages

(A) Overview t-SNE maps of gastric and duodenal enteroendocrine cells (Table S12), which are colored based on their anatomical origin or endocrine cell type (left). t-SNE maps of endocrine signature genes are shown to the right.
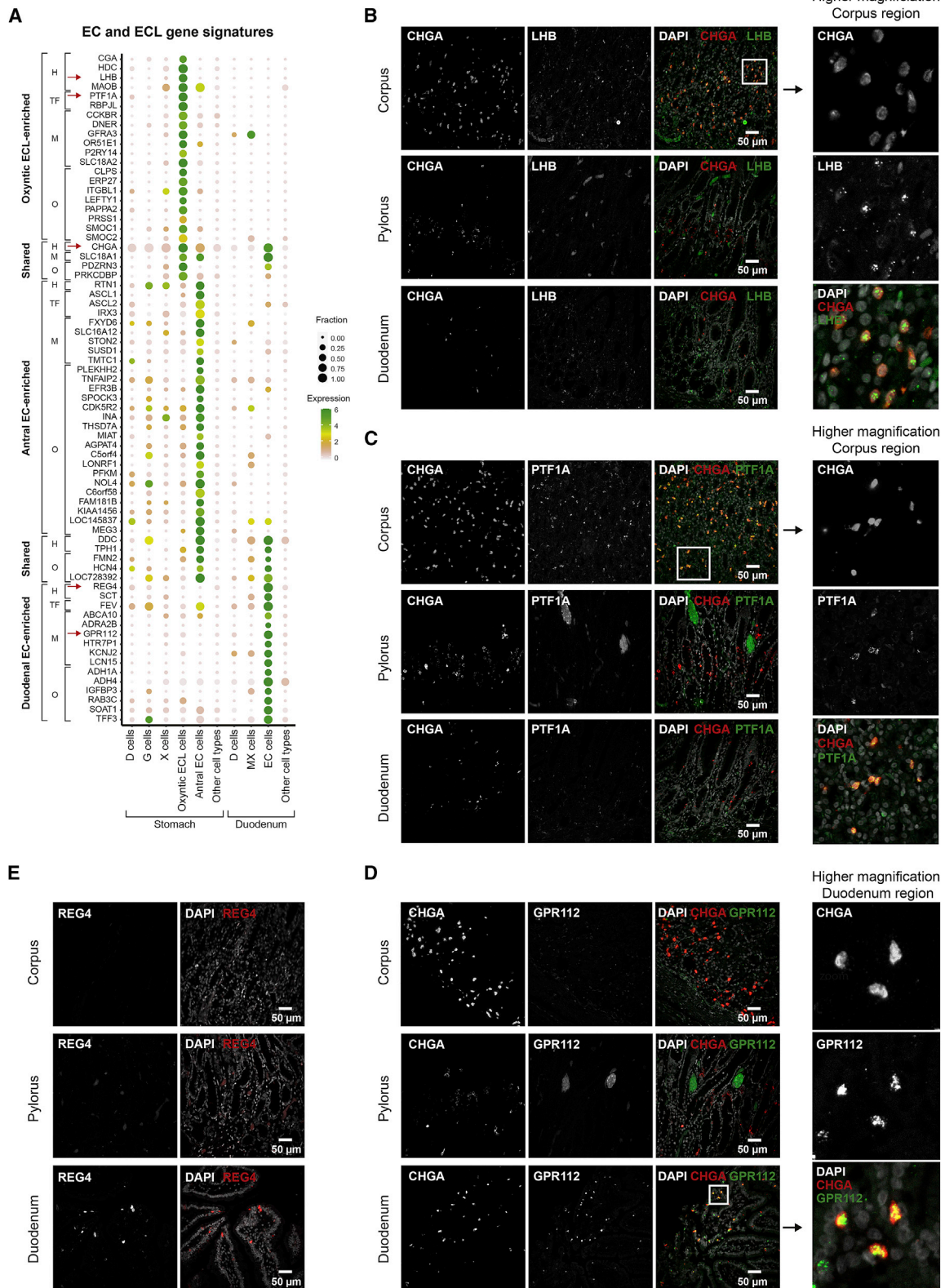
(B–D) Gene expression in endocrine cells, as shown by fraction dot plots. The different gastric and duodenal EECs are shown on the x axis, with EEC-enriched genes on the y axis. Each dot corresponds to the expression of one gene in one cell type. The dot size indicates the fraction of cells expressing a given gene within the indicated cell type, and the color denotes the mean expression level (scaled by *Z* score transformation). Genes on the y axis are grouped according to their function: H, hormones and proteins involved in hormone production and maturation; M, membrane proteins including receptors and channels; TF, transcription factors; O, other functions.

(B) Common and gastric-enriched gene expression signatures in EECs of the stomach and duodenum.

(C) Gene expression signatures of X and MX cells. Red arrows indicate CFC1 and GHRL, which were also analyzed by *in situ* RNA hybridization (see E).

(D) Gene expression signatures of D and G cells.

(E) Co-expression of *CFC1* and *GHRL* in gastric X cells, as determined by *in situ* RNA hybridization (RNAscope). A higher magnification is shown to the right.

A EC and ECL gene signatures

B CHGA / LHB / DAPI CHGA LHB — Corpus, Pylorus, Duodenum. Higher magnification Corpus region

C CHGA / PTF1A / DAPI CHGA PTF1A — Corpus, Pylorus, Duodenum. Higher magnification Corpus region

E REG4 / DAPI REG4 — Corpus, Pylorus, Duodenum

D CHGA / GPR112 / DAPI CHGA GPR112 — Corpus, Pylorus, Duodenum. Higher magnification Duodenum region
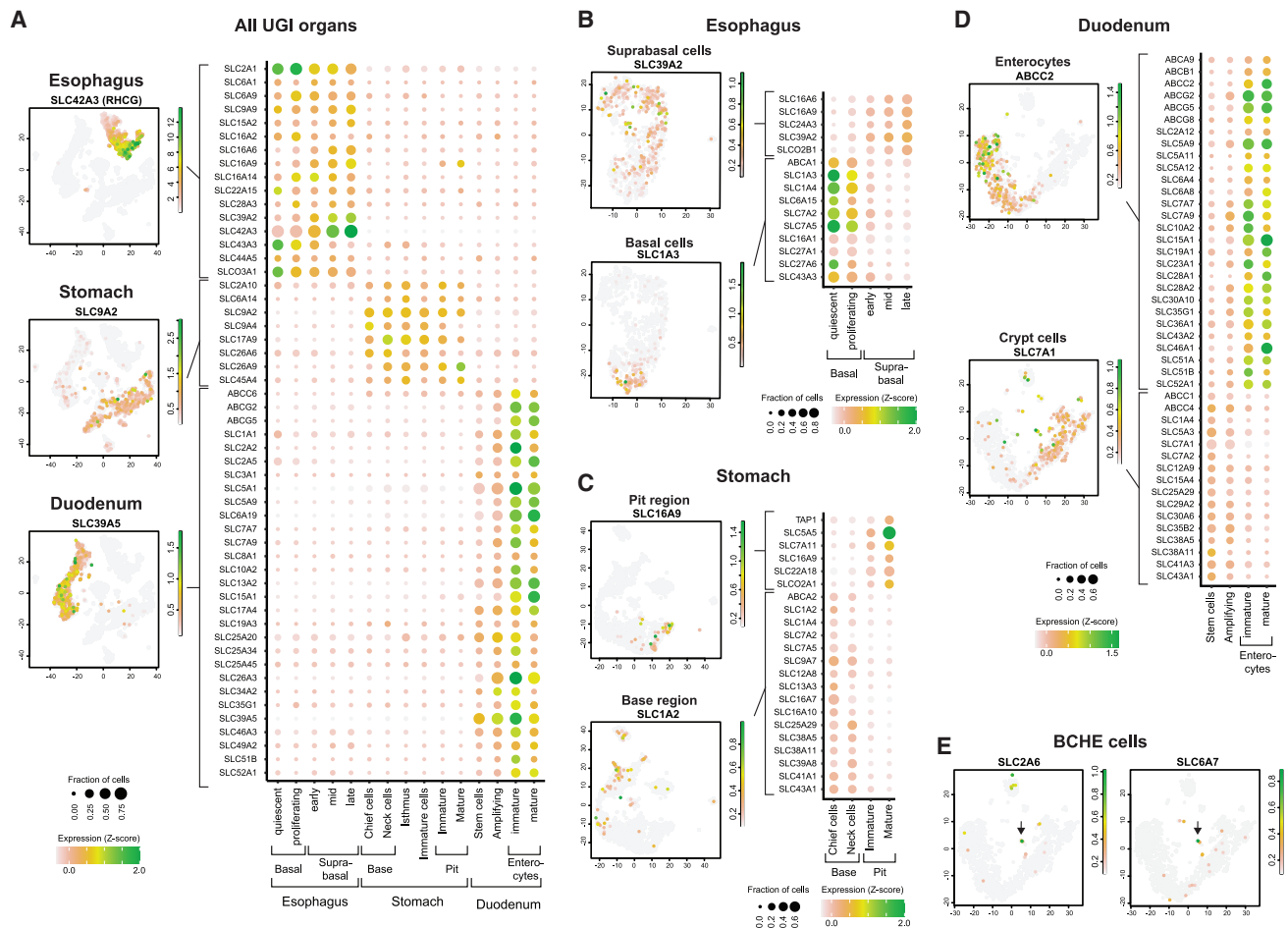
(legend on next page)

Figure 7. Organ-specific and cell-type-enriched expression of transporter genes along the upper gastrointestinal tract
(A–D) Localized expression of solute carriers (SLCs) and ABC transporters along the upper gastrointestinal tract, as shown by fraction dot plots (see Figure 5B). A t-SNE plot (left) indicates the expression of a representative transporter gene for each organ- or cell-type-specific expression pattern. Organ-enriched (A) or cell-type-enriched (B–D) expression of transporter genes is shown for the esophagus, stomach, or duodenum.
(E) t-SNE plots indicating *SLC2A6* and *SLC6A7* expression in duodenal BCHE cells.

different sugar molecules (SLC2A12 and SLC5A9) and amino acids (SLC5A11, SLC7A7, SLC7A9, SLC28A1, SLC28A2, and SLC43A2) (Figure 7D). Next, we looked for common features within the deeper layer of the different epithelia, which include the basal cell layer of the esophagus, the base region of the stomach, and the stem cell compartment of the duodenum. All three regions shared expression of SLC1A4 and SLC7A2, two importers for neutral and cationic amino acids, respectively (Figures 7B–7D), whereas the expression of SLC38A11, an exporter of neutral amino acids, and SLC43A1, an importer of branched-chain amino

acids, were selectively enriched in cells of the gastric base region and duodenal stem cells (Figures 7C and 7D). Notably, the three transporters of the SLC1 family, which are the only members importing the essential amino acids Glu and Asp, were differentially expressed along the UGI epithelia. SLC1A1 was preferentially expressed in duodenal cells (Figure 7A), SLC1A2 in gastric chief cells (Figure 7C), and SLC1A3 in esophageal basal cells (Figure 7B), indicating that Glu and Asp absorption is differentially regulated along the UGI tract. Lastly, by searching for selective expression of transporters in BCHE cells, we identified the glucose-specific

Figure 6. Differential gene expression between ECL and EC cells in the stomach and duodenum
(A) Gene expression signatures of antral and duodenal EC cells as well as oxyntic ECL cells, as shown by fraction dot plots (see Figure 5B). Genes are grouped according to their function (see Figure 5B). Red arrows indicate genes, which were also analyzed by *in situ* RNA hybridization.
(B–E) Validation of the co-expression of genes in the distinct signatures of ECL and EC cells by *in situ* RNA hybridization (RNAscope). All scale bars indicate 50 μm. The *LHB* (B) and *PTF1A* (C) genes are co-expressed with *CHGA* in oxyntic ECL cells of the corpus region (top row), which is shown at higher magnification (right). (D) Co-expression of *GPR112* and *CHGA* cells in EC cells of the duodenum (bottom row), which is shown at higher magnification (right). (E) Expression of *REG4* in the duodenal epithelium (bottom), but not in the stomach epithelium (top and middle).

importers SLC2A6 and the proline transporter SLC6A7 to be expressed in this cell type (Figure 7E). In summary, our scRNA-seq data have provided a comprehensive expression analysis of transporters that determine the secretory and absorptive capacities of the epithelial cells in the human UGI tract.

DISCUSSION

The epithelia of the gastrointestinal tract have been extensively studied by genetic means in the mouse model system (Beumer and Clevers, 2016; Clevers, 2013; Kim and Shivdasani, 2016; Willet and Mills, 2016; Zhang et al., 2017). However, less is known about the cellular composition and expression patterns of the human UGI epithelia. Although a recent single-cell study of the UGI tract focused primarily on the transcriptional similarity between Barrett's esophagus and esophageal submucosal glands (Owen et al., 2018), we have now performed an in-depth analysis of the healthy human epithelia of the esophagus, three distinct stomach regions, and the duodenum, which provided interesting insight into cellular and molecular functions of the distinct cell types of these organs.

The human esophagus was previously shown, by histological analysis, to contain asymmetrically dividing stem/progenitor cells in the basal cell layer (Seery and Watt, 2000). A subpopulation of basal cells with properties of stem cells was subsequently identified in the mouse esophagus (Kalabis et al., 2008). By identifying keratin 15 (Krt15) as an esophageal stem cell marker, Giroux et al. (2017) demonstrated, by lineage-tracing experiments, that the Krt15[+] basal cells are long-lived stem/progenitor cells able to self-renew, proliferate, and differentiate into keratinocytes. Similar to those mouse data, our single-cell analysis separated the basal cells of the human esophagus into a quiescent and proliferating cell population and revealed high expression of KRT15 in the quiescent basal cells. Like KRT15, the hemidesmosome component COL17A1 is also most highly expressed in the quiescent basal cells, thus identifying COL17A1 and KRT15 as markers of the human esophageal stem/progenitor cell compartment. In addition to *COL17A1*, *COL7A1* and *LAMB3* are also specifically co-expressed in the esophageal stem/progenitor cells. Notably, mutations in these three genes cause epidermolysis bullosa (EB), a well-characterized blistering disease of the skin (Hirsch et al., 2017; Natsuga et al., 2019; Shinkuma, 2015). Because EB is a hereditary disease, it is conceivable that the esophageal epithelium may also be affected in these patients, as is suggested by a report describing esophageal stenosis as a consequence of esophageal blistering in a patient with EB (Michalak et al., 2018). The COL17A1[high] KRT15[high] stem/progenitor cells are furthermore characterized by high expression of *NRTK2* and *IL1R2*, indicating that these stem cells may depend on signaling via the neurotrophic receptor NRTK2 and on the secreted IL-1 decoy receptor IL1R2 to prevent IL-1-induced inflammation (Peters et al., 2013).

The stomach of the mouse is constantly regenerated by fast-cycling stem cells in the isthmus region (Han et al., 2019) and by chief cells at the base of the gastric glands (Barker et al., 2010; Stange et al., 2013). Here, we describe expression of *NMU* in the human isthmus cells, which was, however, not detected in the mouse dataset. Systematic comparison of the

gene expression patterns between identical cell types of the human and mouse stomach revealed many more species-specific expression differences (Figures 3G–3I, S5, and S6) in addition to the known case of the gastric intrinsic factor (GIF) (Levine et al., 1980; Figure S5B). Mouse chief cells express the triglyceride lipase, Pnliprp2, and the colipase, Clps, whereas the gastric lipase (LIPF) is expressed in human chief cells, suggesting that different enzymes may contribute to the lipid metabolism of this cell type in the two species. Expression of the secreted uroguanylin (GUCA2B), the atypical NOTCH ligand DNER, the dopamine receptor D5 (DRD5), and the gastrin/cholecystokinin B receptor (CCKBR) could be detected only in human, not mouse, parietal cells, indicating that parietal cells in the two species may differ in certain signaling aspects. The pit cells in the human, but not mouse, stomach expressed the secreted IL-1 decoy receptors IL1R2 and IL1RN, suggesting that pits cells of the two species use different mechanisms to control IL-1-mediated inflammation. In addition, LYZ-secreting cells are abundant in human antral glands, whereas expression of the two homologous genes *Lyz1* and *Lyz2* was not detected in the mouse stomach. Finally, the stem-cell-supporting Paneth cells in the mouse intestinal crypt express Wnt3 and Wnt11 (Sato et al., 2011b), whereas no expression of any WNT family member could be observed in Paneth cells of the human duodenum, suggesting that the human intestinal stem cells entirely rely on WNT signals from adjacent mesenchymal niche cells. The absence of WNT expression in human Paneth cells provides a molecular explanation for why the *in vitro* growth of human small intestinal organoids depends on the addition of exogenous WNT, in contrast to the respective mouse organoids (Sato et al., 2009; 2011a; 2011b). Hence, the observed differences between human and mouse are more extensive than expected and may complicate the extrapolation of gene expression and phenotypic data obtained in the mouse model organism to the respective human gastrointestinal organs.

The human intestinal epithelium was previously shown to contain BEST4[+] cells (Ito et al., 2013), and the duodenum of both human and rat was reported to comprise CFTR[high] epithelial cells at a low frequency (Ameen et al., 1995; Jakab et al., 2013). Here, we have discovered that BEST4 and CFTR are highly expressed in the same duodenal cell type, referred to as BCHE cell. These cells are present already in the human fetal intestine, as revealed by reanalysis of published scRNA-seq data (Gao et al., 2018). BEST4[+] cells have also been identified in the human colon (Ito et al., 2013; Parikh et al., 2019; Smillie et al., 2019). Although the colonic BEST4[+] and duodenal BCHE cells share the expression of some genes, they also significantly differ in their expression pattern and may thus have distinct functions, which is best exemplified by the absence of *CFTR* transcripts in the colonic BEST4[+] cells. Interestingly, both BEST4 and CFTR are channels with chloride/bicarbonate transport function (Hartzell et al., 2008; Moran, 2017). In particular, the CFTR channel, which is localized at the apical surface of epithelial cells, is responsible for high-volume fluid secretion of water, chloride, and bicarbonate anions into the lumen of the gut (Moran, 2017). The high-volume fluid secretion by CFTR-expressing epithelial cells such as the BCHE cells may contribute to the removal of adherent mucus to clear the surfaces of adjacent

absorptive enterocyte to allow nutrient absorption (Jakab et al., 2013). Activation of the CFTR channel requires phosphorylation by the cAMP-dependent protein kinase A and the cGMP-dependent kinase PRKG2, which, in turn, are stimulated by signaling through the transmembrane receptor guanylate cyclase C (GUCY2C; Moran, 2017; Rappaport and Waldman, 2018). A direct connection between GUCY2C signaling and high-volume fluid secretion by CFTR is supported by genetic inactivation of *Prkg2* and *Gucy2c* in the mouse (Pfeifer et al., 1996; Schulz et al., 1997) as well as by gain- and loss-of-function mutations of *GUCY2C* in humans (Fiskerstrand et al., 2012; Romi et al., 2012). Remarkably, guanylin (*GUCA2A*) and uroguanylin (*GUCA2B*), the two physiological ligands of the GUCY2C receptor (Rappaport and Waldman, 2018), are highly and selectively expressed in BCHE cells, whereas their receptor is ubiquitously expressed in epithelial cells of the duodenum including the BCHE cells. Hence, an autocrine signaling loop initiated by guanylin and uroguanylin expression may continually activate the CFTR channels in BCHE cells, thus leading to constitutive high-volume fluid secretion.

Our single-cell analysis of the human UGI epithelia has also provided a molecular definition of the different enteroendocrine cell types, which secrete hormones to regulate the gastrointestinal activity (Gribble and Reimann, 2017). These enteroendocrine cell types share a large common expression signature of genes that code for endocrine-cell-specific transcription factors and proteins involved in hormone production, proteolytic processing, and secretion. This analysis also identified specific gene expression patterns for the different cell types. In that regard, the molecular definition of the antral EC cells is of particular interest, because so far, little is known about this rare, serotonin-producing cell type in the stomach of humans (Choi et al., 2014; Fakhry et al., 2019; Ito et al., 2013) and rodents (Engelstoft et al., 2015; Reynaud et al., 2016). The antral EC cells significantly differ in their expression pattern from the duodenal EC cells, as both cell types share the expression of only a few proteins, including chromogranin A and the serotonin-producing enzymes TPH1 and DDC.

Another unexpected aspect of the enteroendocrine cell analysis was the discovery that the oxyntic ECL cells express both subunits of LH. LH is best known for its role in controlling follicular maturation and ovulation during the female reproductive cycle and for stimulating testosterone production in males (Choi and Smitz, 2014). Notably, studies with male rhesus monkeys revealed that the fasting-induced suppression of LH levels in the blood is rapidly and progressively restored in response to increasing food intake because of a "metabolic signal" (Parfitt et al., 1991; Schreihofer et al., 1993). Moreover, LH has been implicated in controlling gastrointestinal motility by regulating the myoelectric complex in the small intestines of rats (Ducker et al., 1996). It is thus conceivable that the effects observed in rhesus monkeys and rats may be caused by local secretion of LH by the oxyntic ECL cells in the stomach. Future investigations will be required to elucidate the role of locally produced LH in the stomach. Importantly, our single-cell analyses have identified the LH as yet another member of the large family of hormones that are produced by the enteroendocrine cells in the gastrointestinal tract.

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 patients
- ● METHOD DETAILS
  - ○ Sample preparation for scRNA-seq
  - ○ RNA *in situ* hybridization (RNAScope)
  - ○ Immunohistochemical analysis
  - ○ Organoid cultures
  - ○ RT-qPCR analysis
  - ○ Library preparation
- ● QUANTIFICATION AND STATISTICAL ANALYSIS
  - ○ Analysis of single-cell RNA sequencing data
  - ○ Comparison of differential gene expression between the human and mouse stomach
  - ○ Re-analysis of human colon scRNA-seq data
  - ○ Analysis of endocrine cells
  - ○ Analysis of tissue- and cell type-specific transporter gene expression
  - ○ Quantification of MKI67 positive signal in esophageal COL17A1[high] and KRT15[high] basal cells

REFERENCES

Ameen, N.A., Ardito, T., Kashgarian, M., and Marino, C.R. (1995). A unique subset of rat and human intestinal villus cells express the cystic fibrosis transmembrane conductance regulator. Gastroenterology 108, 1016–1023.

Ayyaz, A., Kumar, S., Sangiorgi, B., Ghoshal, B., Gosio, J., Ouladan, S., Fink, M., Barutcu, S., Trcka, D., Shen, J., et al. (2019). Single-cell transcriptomes of the regenerating intestine reveal a revival stem cell. Nature 569, 121–125.

Bando, M., Iwakura, H., Koyama, H., Hosoda, H., Shigematsu, Y., Ariyasu, H., Akamizu, T., Kangawa, K., and Nakao, K. (2016). High incorporation of long-chain fatty acids contributes to the efficient production of acylated ghrelin in ghrelin-producing cells. FEBS Lett. 590, 992–1001.

Barker, N., Huch, M., Kujala, P., van de Wetering, M., Snippert, H.J., van Es, J.H., Sato, T., Stange, D.E., Begthel, H., van den Born, M., et al. (2010). Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. Cell Stem Cell 6, 25–36.

Bartfeld, S., and Clevers, H. (2015). Organoids as model for infectious diseases: culture of human and murine stomach organoids and microinjection of Helicobacter pylori. J. Vis. Exp. 105, 53359.

Beumer, J., and Clevers, H. (2016). Regulation and plasticity of intestinal stem cells during homeostasis and regeneration. Development 143, 3639–3649.

Beumer, J., Puschhof, J., Bauzá-Martinez, J., Martínez-Silgado, A., Elmentaite, R., James, K.R., Ross, A., Hendriks, D., Artegiani, B., Busslinger, G.A., et al. (2020). High resolution mRNA and secretome atlas of human enteroendocrine cells. Cell 181, 1291–1306.e19.

César-Razquin, A., Snijder, B., Frappier-Brinton, T., Isserlin, R., Gyimesi, G., Bai, X., Reithmeier, R.A., Hepworth, D., Hediger, M.A., Edwards, A.M., and Superti-Furga, G. (2015). A call for systematic research on solute carriers. Cell 162, 478–487.

Choi, J., and Smitz, J. (2014). Luteinizing hormone and human chorionic gonadotropin: origins of difference. Mol. Cell. Endocrinol. 383, 203–213.

Choi, E., Roland, J.T., Barlow, B.J., O'Neal, R., Rich, A.E., Nam, K.T., Shi, C., and Goldenring, J.R. (2014). Cell lineage distribution atlas of the human stomach reveals heterogeneous gland populations in the gastric antrum. Gut 63, 1711–1720.

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

Cocco, E., Scaltriti, M., and Drilon, A. (2018). NTRK fusion-positive cancers and TRK inhibitor therapy. Nat. Rev. Clin. Oncol. 15, 731–747.

Coebergh, J.W.W., van Veen, E.-B., Vandenbroucke, J.P., van Diest, P., and Oosterhuis, W. (2006). One-time general consent for research on biological samples: opt out system for patients is optimal and endorsed in many countries. BMJ 332, 665.

Corleto, V.D. (2010). Somatostatin and the gastrointestinal tract. Curr. Opin. Endocrinol. Diabetes Obes. 17, 63–68.

Corvol, H., Rousselet, N., Thompson, K.E., Berdah, L., Cottin, G., Foussigniere, T., Longchampt, E., Fiette, L., Sage, E., Prunier, C., et al. (2018). FAM13A is a modifier gene of cystic fibrosis lung phenotype regulating rhoa activity, actin cytoskeleton dynamics and epithelial-mesenchymal transition. J. Cyst. Fibros. 17, 190–203.

de Lima Marson, F.A., Bertuzzo, C.S., Secolin, R., Ribeiro, A.F., and Ribeiro, J.D. (2013). Genetic interaction of GSH metabolic pathway genes in cystic fibrosis. BMC Med. Genet. 14, 60.

Ducker, T.E., Boss, J.W., Altug, S.A., Mehrabian, H., Dekeratry, D.R., Clench, M.H., and Mathias, J.R. (1996). Luteinizing hormone and human chorionic gonadotropin fragment the migrating myoelectric complex in rat small intestine. Neurogastroenterol. Motil. 8, 95–100.

Edelman, A. (2014). Cytoskeleton and CFTR. Int. J. Biochem. Cell Biol. 52, 68–72.

Engelstoft, M.S., Lund, M.L., Grunddal, K.V., Egerod, K.L., Osborne-Lawrence, S., Poulsen, S.S., Zigman, J.M., and Schwartz, T.W. (2015). Research resource: a chromogranin A reporter for serotonin- and histamine-secreting enteroendocrine cells. Mol. Endocrinol. 29, 1658–1671.

Ewert, S., Spak, E., Olbers, T., Johnsson, E., Edebo, A., and Fändriks, L. (2006). Angiotensin II induced contraction of rat and human small intestinal wall musculature in vitro. Acta Physiol. (Oxf.) 188, 33–40.

Fakhry, J., Stebbing, M.J., Hunne, B., Bayguinov, Y., Ward, S.M., Sasse, K.C., Callaghan, B., McQuade, R.M., and Furness, J.B. (2019). Relationships of endocrine cells to each other and to other cell types in the human gastric fundus and corpus. Cell Tissue Res. 376, 37–49.

Fiskerstrand, T., Arshad, N., Haukanes, B.I., Tronstad, R.R., Pham, K.D.-C., Johansson, S., Håvik, B., Tønder, S.L., Levy, S.E., Brackman, D., et al. (2012). Familial diarrhea syndrome caused by an activating GUCY2C mutation. N. Engl. J. Med. 366, 1586–1595.

Galloway, A., Saveliev, A., Łukasiak, S., Hodson, D.J., Bolland, D., Balmanno, K., Ahlfors, H., Monzón-Casanova, E., Mannurita, S.C., Bell, L.S., et al. (2016). RNA-binding proteins ZFP36L1 and ZFP36L2 promote cell quiescence. Science 352, 453–459.

Gao, Z.-G., and Jacobson, K.A. (2017). Purinergic signaling in mast cell degranulation and asthma. Front. Pharmacol. 8, 947.

Gao, S., Yan, L., Wang, R., Li, J., Yong, J., Zhou, X., Wei, Y., Wu, X., Wang, X., Fan, X., et al. (2018). Tracing the temporal-spatial transcriptome landscapes of the human fetal digestive tract using single-cell RNA-sequencing. Nat. Cell Biol. 20, 721–734.

Gehart, H., and Clevers, H. (2019). Tales from the crypt: new insights into intestinal stem cells. Nat. Rev. Gastroenterol. Hepatol. 16, 19–34.

Gehart, H., van Es, J.H., Hamer, K., Beumer, J., Kretzschmar, K., Dekkers, J.F., Rios, A., and Clevers, H. (2019). Identification of enteroendocrine regulators by real-time single-cell differentiation mapping. Cell 176, 1158–1173.e16.

Gershon, M.D. (2013). 5-Hydroxytryptamine (serotonin) in the gastrointestinal tract. Curr. Opin. Endocrinol. Diabetes Obes. 20, 14–21.

Giroux, V., Lento, A.A., Islam, M., Pitarresi, J.R., Kharbanda, A., Hamilton, K.E., Whelan, K.A., Long, A., Rhoades, B., Tang, Q., et al. (2017). Long-lived keratin 15[+] esophageal progenitor cells contribute to homeostasis and regeneration. J. Clin. Invest. 127, 2378–2391.

Gribble, F.M., and Reimann, F. (2017). Signalling in the gut endocrine axis. Physiol. Behav. 176, 183–188.

Grün, D., and van Oudenaarden, A. (2015). Design and analysis of single-cell sequencing experiments. Cell 163, 799–810.

Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature 525, 251–255.

Grün, D., Muraro, M.J., Boisset, J.-C., Wiebrands, K., Lyubimova, A., Dharmadhikari, G., van den Born, M., van Es, J., Jansen, E., Clevers, H., et al. (2016). De novo prediction of stem cell identity using single-cell transcriptome data. Cell Stem Cell 19, 266–277.

Haber, A.L., Biton, M., Rogel, N., Herbst, R.H., Shekhar, K., Smillie, C., Burgin, G., Delorey, T.M., Howitt, M.R., Katz, Y., et al. (2017). A single-cell survey of the small intestinal epithelium. Nature 551, 333–339.

Han, S., Fink, J., Jörg, D.J., Lee, E., Yum, M.K., Chatzeli, L., Merker, S.R., Josserand, M., Trendafilova, T., Andersson-Rolf, A., et al. (2019). Defining the identity and dynamics of adult gastric isthmus stem cells. Cell Stem Cell 25, 342–356.e7.

Hartzell, H.C., Qu, Z., Yu, K., Xiao, Q., and Chien, L.-T. (2008). Molecular physiology of bestrophins: multifunctional membrane proteins linked to best disease and other retinopathies. Physiol. Rev. 88, 639–672.

Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., de Leeuw, Y., Anavy, L., Gennert, D., Li, S., Livak, K.J., Rozenblatt-Rosen, O., et al. (2016). CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. Genome Biol. 17, 77.

Hatzis, P., van der Flier, L.G., van Driel, M.A., Guryev, V., Nielsen, F., Denissov, S., Nijman, I.J., Koster, J., Santo, E.E., Welboren, W., et al. (2008). Genome-wide pattern of TCF7L2/TCF4 chromatin occupancy in colorectal cancer cells. Mol. Cell. Biol. 28, 2732–2744.

Hediger, M.A., Clémençon, B., Burrier, R.E., and Bruford, E.A. (2013). The ABCs of membrane transporters in health and disease (SLC series): introduction. Mol. Aspects Med. 34, 95–107.

Herman, J.S., Sagar, and Grün, D. (2018). FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. Nat. Methods 15, 379–386.

Hirsch, T., Rothoeft, T., Teig, N., Bauer, J.W., Pellegrini, G., De Rosa, L., Scaglione, D., Reichelt, J., Klausegger, A., Kneisz, D., et al. (2017). Regeneration of the entire human epidermis using transgenic stem cells. Nature 551, 327–332.

Hunt, R.H., Camilleri, M., Crowe, S.E., El-Omar, E.M., Fox, J.G., Kuipers, E.J., Malfertheiner, P., McColl, K.E.L., Pritchard, D.M., Rugge, M., et al. (2015). The stomach in health and disease. Gut 64, 1650–1668.

Ito, G., Okamoto, R., Murano, T., Shimizu, H., Fujii, S., Nakata, T., Mizutani, T., Yui, S., Akiyama-Morio, J., Nemoto, Y., et al. (2013). Lineage-specific expression of bestrophin-2 and bestrophin-4 in human intestinal epithelial cells. PLoS ONE 8, e79693.

Jakab, R.L., Collaco, A.M., and Ameen, N.A. (2013). Characterization of CFTR high expresser cells in the intestine. Am. J. Physiol. Gastrointest. Liver Physiol. 305, G453–G465.

Kalabis, J., Oyama, K., Okawa, T., Nakagawa, H., Michaylira, C.Z., Stairs, D.B., Figueiredo, J.-L., Mahmood, U., Diehl, J.A., Herlyn, M., and Rustgi, A.K. (2008). A subpopulation of mouse esophageal basal cells has properties of stem cells with the capacity for self-renewal and lineage specification. J. Clin. Invest. 118, 3860–3869.

Kim, T.H., and Shivdasani, R.A. (2016). Stomach development, stem cells and disease. Development 143, 554–565.

Künzli, K., Favre, B., Chofflon, M., and Borradori, L. (2016). One gene but different proteins and diseases: the complexity of dystonin and bullous pemphigoid antigen 1. Exp. Dermatol. 25, 10–16.

Lersch, R., and Fuchs, E. (1988). Sequence and expression of a type II keratin, K5, in human epidermal cells. Mol. Cell. Biol. 8, 486–493.

Levine, J.S., Nakane, P.K., and Allen, R.H. (1980). Immunocytochemical localization of human intrinsic factor: the nonstimulated stomach. Gastroenterology 79, 493–502.

Liu, N., Matsumura, H., Kato, T., Ichinose, S., Takada, A., Namiki, T., Asakawa, K., Morinaga, H., Mohri, Y., De Arcangelis, A., et al. (2019). Stem cell competition orchestrates skin homeostasis and ageing. Nature 568, 344–350.

Lu, J., Chatterjee, M., Schmid, H., Beck, S., and Gawaz, M. (2016). CXCL14 as an emerging immune and inflammatory modulator. J. Inflamm. (Lond.) 13, 1.

Mack, S.M., Gomes, I., and Devi, L.A. (2019). Neuropeptide PEN and its receptor GPR83: distribution, signaling, and regulation. ACS Chem. Neurosci. 10, 1884–1891.

Martinez, V.G., and O'Driscoll, L. (2015). Neuromedin U: a multifunctional neuropeptide with pleiotropic roles. Clin. Chem. 61, 471–482.

Michalak, A., Cichoż-Lach, H., Prozorow-Król, B., Buk, L., and Dzida, M. (2018). A rare case of skin blistering and esophageal stenosis in the course of epidermolysis bullosa—case report and literature review. BMC Gastroenterol. 18, 47.

Moran, O. (2017). The gating of the CFTR channel. Cell. Mol. Life Sci. 74, 85–92.

Muraro, M.J., Dharmadhikari, G., Grün, D., Groen, N., Dielen, T., Jansen, E., van Gurp, L., Engelse, M.A., Carlotti, F., de Koning, E.J.P., and van Oudenaarden, A. (2016). A single-cell transcriptome atlas of the human pancreas. Cell Syst. 3, 385–394.e3.

Natsuga, K., Watanabe, M., Nishie, W., and Shimizu, H. (2019). Life before and beyond blistering: the role of collagen XVII in epidermal physiology. Exp. Dermatol. 28, 1135–1141.

O'Neal, W.K., and Knowles, M.R. (2018). Cystic fibrosis disease modifiers: complex genetics defines the phenotypic diversity in a monogenic disease. Annu. Rev. Genomics Hum. Genet. 19, 201–222.

Owen, R.P., White, M.J., Severson, D.T., Braden, B., Bailey, A., Goldin, R., Wang, L.M., Ruiz-Puig, C., Maynard, N.D., Green, A., et al. (2018). Single cell RNA-seq reveals profound transcriptional similarity between Barrett's oesophagus and oesophageal submucosal glands. Nat. Commun. 9, 4261.

Parfitt, D.B., Church, K.R., and Cameron, J.L. (1991). Restoration of pulsatile luteinizing hormone secretion after fasting in rhesus monkeys (Macaca mulatta): dependence on size of the refeed meal. Endocrinology 129, 749–756.

Parikh, K., Antanaviciute, A., Fawkner-Corbett, D., Jagielowicz, M., Aulicino, A., Lagerholm, C., Davis, S., Kinchen, J., Chen, H.H., Alham, N.K., et al. (2019). Colonic epithelial cell diversity in health and inflammatory bowel disease. Nature 567, 49–55.

Peters, V.A., Joesting, J.J., and Freund, G.G. (2013). IL-1 receptor 2 (IL-1R2) and its role in immune regulation. Brain Behav. Immun. 32, 1–8.

Pfeifer, A., Aszódi, A., Seidler, U., Ruth, P., Hofmann, F., and Fässler, R. (1996). Intestinal secretory defects and dwarfism in mice lacking cGMP-dependent protein kinase II. Science 274, 2082–2086.

Prinz, C., Zanner, R., and Gratzl, M. (2003). Physiology of gastric enterochromaffin-like cells. Annu. Rev. Physiol. 65, 371–382.

Rappaport, J.A., and Waldman, S.A. (2018). The guanylate cyclase C-cGMP signaling axis opposes intestinal epithelial injury and neoplasia. Front. Oncol. 8, 299.

Reynaud, Y., Fakhry, J., Fothergill, L., Callaghan, B., Ringuet, M., Hunne, B., Bravo, D.M., and Furness, J.B. (2016). The chemical coding of 5-hydroxytryptamine containing enteroendocrine cells in the mouse gastrointestinal tract. Cell Tissue Res. 364, 489–497.

Romi, H., Cohen, I., Landau, D., Alkrinawi, S., Yerushalmi, B., Hershkovitz, R., Newman-Heiman, N., Cutting, G.R., Ofir, R., Sivan, S., and Birk, O.S. (2012). Meconium ileus caused by mutations in GUCY2C, encoding the CFTR-activating guanylate cyclase 2C. Am. J. Hum. Genet. 90, 893–899.

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

Sato, T., Stange, D.E., Ferrante, M., Vries, R.G.J., Van Es, J.H., Van den Brink, S., Van Houdt, W.J., Pronk, A., Van Gorp, J., Siersema, P.D., and Clevers, H. (2011a). Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. Gastroenterology 141, 1762–1772.

Sato, T., van Es, J.H., Snippert, H.J., Stange, D.E., Vries, R.G., van den Born, M., Barker, N., Shroyer, N.F., van de Wetering, M., and Clevers, H. (2011b). Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. Nature 469, 415–418.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682.

Schreihofer, D.A., Golden, G.A., and Cameron, J.L. (1993). Cholecystokinin (CCK)-induced stimulation of luteinizing hormone (LH) secretion in adult male rhesus monkeys: examination of the role of CCK in nutritional regulation of LH secretion. Endocrinology 132, 1553–1560.

Schulz, S., Lopez, M.J., Kuhn, M., and Garbers, D.L. (1997). Disruption of the guanylyl cyclase-C gene leads to a paradoxical phenotype of viable but heat-stable enterotoxin-resistant mice. J. Clin. Invest. 100, 1590–1595.

Seery, J.P., and Watt, F.M. (2000). Asymmetric stem-cell divisions define the architecture of human oesophageal epithelium. Curr. Biol. 10, 1447–1450.

Shinkuma, S. (2015). Dystrophic epidermolysis bullosa: a review. Clin. Cosmet. Investig. Dermatol. 8, 275–284.

Sindić, A., and Schlatter, E. (2006). Cellular effects of guanylin and uroguanylin. J. Am. Soc. Nephrol. 17, 607–616.

Smillie, C.S., Biton, M., Ordovas-Montanes, J., Sullivan, K.M., Burgin, G., Graham, D.B., Herbst, R.H., Rogel, N., Slyper, M., Waldman, J., et al. (2019).

Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. Cell 178, 714–730.e22.

Stange, D.E., Koo, B.-K., Huch, M., Sibbel, G., Basak, O., Lyubimova, A., Kujala, P., Bartfeld, S., Koster, J., Geahlen, J.H., et al. (2013). Differentiated Troy+ chief cells act as reserve stem cells to generate all lineages of the stomach epithelium. Cell 155, 357–368.

Thiagarajah, J.R., Donowitz, M., and Verkman, A.S. (2015). Secretory diarrhoea: mechanisms and emerging therapies. Nat. Rev. Gastroenterol. Hepatol. 12, 446–457.

Thompson, C.A., DeLaForest, A., and Battle, M.A. (2018). Patterning the gastrointestinal epithelium to confer regional-specific functions. Dev. Biol. 435, 97–108.

Thwaites, D.T., and Anderson, C.M.H. (2007). H+-coupled nutrient, micronutrient and drug transporters in the mammalian small intestine. Exp. Physiol. 92, 603–619.

van der Doef, H.P.J., Kokke, F.T.M., van der Ent, C.K., and Houwen, R.H.J. (2011). Intestinal obstruction syndromes in cystic fibrosis: meconium ileus, distal intestinal obstruction syndrome, and constipation. Curr. Gastroenterol. Rep. 13, 265–270.

van der Lienden, M.J.C., Gaspar, P., Boot, R., Aerts, J.M.F.G., and van Eijk, M. (2018). Glycoprotein non-metastatic protein B: an emerging biomarker for lysosomal dysfunction in macrophages. Int. J. Mol. Sci. 20, 66.

Wang, F., Flanagan, J., Su, N., Wang, L.-C., Bui, S., Nielson, A., Wu, X., Vo, H.-T., Ma, X.-J., and Luo, Y. (2012). RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. J. Mol. Diagn. 14, 22–29.

Wang, Y., Song, W., Wang, J., Wang, T., Xiong, X., Qi, Z., Fu, W., Yang, X., and Chen, Y.-G. (2020). Single-cell transcriptome analysis reveals differential nutrient absorption functions in human intestine. J. Exp. Med. 217, 357.

Willet, S.G., and Mills, J.C. (2016). Stomach organ and cell lineage differentiation: from embryogenesis to adult homeostasis. Cell. Mol. Gastroenterol. Hepatol. 2, 546–559.

Worthington, J.J., Reimann, F., and Gribble, F.M. (2018). Enteroendocrine cells-sensory sentinels of the intestinal environment and orchestrators of mucosal immunity. Mucosal Immunol. 11, 3–20.

Zhang, Y., Jiang, M., Kim, E., Lin, S., Liu, K., Lan, X., and Que, J. (2017). Development and stem cells of the esophagus. Semin. Cell Dev. Biol. 66, 25–35.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-COL17A1	Sigma	HPA043673, RRID:AB_10960893
Rabbit anti-COL17A1	Abcam	ab28440, RRID:AB_731698
EPCAM-PE	BioLegend	324206, RRID:AB_756080
Goat anti-IL1R2	R&D Systems	AF-263, RRID:AB_354431
Rabbit anti-Keratin-5 (KRT5)	Biocompare	905501, RRID:AB_2565050
Mouse anti-Keratin-13 (KRT13)	Abnova Corporation	MAB1864, RRID:AB_1679330
Mouse anti-Keratin-14 (KRT14)	Thermo Scientific	MA5-11599, RRID:AB_10982092
Mouse anti-Keratin-15 (KRT15)	Santa Cruz	sc-47697, RRID:AB_627847
Sheep anti-Lysozyme	US Biological	L9200-16E, RRID:AB_2138782
Mouse anti-MKI67	Monosan	MONX10283, RRID:AB_1833494
Rabbit anti-MKI67	Abcam	ab16667, RRID:AB_302459
Rabbit anti-MKI67	Milipore Merck	AB9260, RRID:AB_2142366
Alexa Fluor 568 donkey anti-mouse	Life technologies	A10037, RRID:AB_2534013
Alexa Fluor 488 donkey anti-rabbit	Life technologies	A21206, RRID:AB_2535792
Alexa Fluor 568 donkey anti-goat	Life technologies	A11057, RRID:AB_142581
Biological samples		
Human biopsies	Utrecht Medical Center	N/A
Chemicals, peptides, and recombinant proteins		
Trypsin enzyme	Promega, Madison, USA	Catalogue # T1426
DNase I	Sigma-Aldrich, Missouri, USA	Catalogue #DN25
Advanced DMEM/F12	Thermo Fisher Scientific	12634-010
B-27 Supplement minus Vitamin A	Thermo Fisher Scientific	12587010
B-27 Supplement plus Vitamin A	Thermo Fisher Scientific	*17504044*
GlutaMAX	Thermo Fisher Scientific	35050061
HEPES	Thermo Fisher Scientific	15630080
Penicillin-Streptomycin	Thermo Fisher scientific	15140122
Wnt Surrogate	U-Protein Express	Custom order
Noggin conditioned medium	U-Protein Express	Custom order
N-Acetyl-L-cysteine	Sigma-Aldrich	A9165
Nicotinamide	Sigma-Aldrich	N0636
Human EGF	Peprotech	AF-100-15
Human KGF/FGF-7	Peprotech	100-19
Human FGF-10	Peprotech	100-26
A83-01	Tocris	2939
Prostaglandin E2	Tocris	2296
Forskolin	Tocris	1099
SB202190	Sigma-Aldrich	S7076
Y-27632 dihydrochloride	Abmole	M1817
Primocin	Invivogen	ant-pm-2
Fungin	Invivogen	ant-fn-1
Cultrex Basement Membrane Extract (BME), Growth Factor Reduced, Type 2	R&D Systems, Bio-Techne	3533-001-02
DAPI	Thermo Fisher Scientific	D1306

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Donkey serum	Golden Bridge International	E27-100
SYBR Green	Bio Rad	1725270
Formaldehyde solution 4%	Sigma-Aldrich	1.00496
SORT-seq reagents	Muraro et al. (2016)	N/A
TryplE	Thermo Fisher Scientific	12605010
Prolong Gold Antifade	Thermo Fisher Scientific	P10144
Critical commercial assays		
RNeasy Mini Kit	QIAGEN	74104
GoScript Reverse Transcriptase	Promega	A5003
Thermo Scientific reagents for CEL-Seq2	(Hashimshony et al., 2016)	N/A
Reagents for library preparation from CEL-Seq2	(Hashimshony et al., 2016)	N/A
RNAScope® Multiplex Fluorescent Reagent Kit v2	ACDBio	N/A
Wizard Genomic DNA Purification Kit	Promega	A1120
Deposited data		
Raw reads of all human scRNaseq	European Genome-Phenome archive (https://www.ebi.ac.uk/ega/home)	EGAS00001004695
Raw reads of all mouse stomach scRNaseq	Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/)	GSE157694
Experimental models: cell lines		
R-spondin expressing cell line	Bio-Techne	3710-001-01
Experimental models: organisms/strains		
Wild-type mice	C57BL/6	N/A
Oligonucleotides		
COL7A1 qPCR (fw, rev)	5' GCTCTTGGCATTCCTAGTGGTC 3'	5' GATGTACTGCTGAATCTCCTGGC 3'
COL7A1 qPCR (fw, rev)	5' GTTGGAGAGAAAGGTGACGAGG 3'	5' TGGTCTCCCTTTTCACCCACAG 3'
CXCL14 qPCR (fw, rev)	5' AGATCCGCTACAGCGACGTGAA 3'	5' GCAGTGCTCCTGACCTCGGTA 3'
IVL qPCR (fw, rev)	5' GGTCCAAGACATTCAACCAGCC 3'	5' TCTGGACACTGCGGGTGGTTAT 3'
KI67 qPCR (fw, rev)	5' GAAAGAGTGGCAACCTGCCTTC 3'	5' GCACCAAGTTTTACTACATCTGCC 3'
KRT13 qPCR (fw, rev)	5' GATGCTGAGGAATGGTTCCACG 3'	5' AGCTCCGTGATCTCTGTCTTGC 3'
KRT14 qPCR (fw, rev)	5' TGCCGAGGAATGGTTCTTCACC 3'	5' GCAGCTCAATCTCCAGGTTCTG 3'
KRT15 qPCR (fw, rev)	5' AGGACTGACCTGGAGATGCAGA 3'	5' TGCGTCCATCTCCACATTGACC 3'
KRT4 qPCR (fw, rev)	5' GCCGAGAATGACTTTGTGGTCC 3'	5' CTCCGCATCATAGAGGACCTTC 3'
KRT5 qPCR (fw, rev)	5' GCTGCCTACATGAACAAGGTGG 3'	5' ATGGAGAGGACCACTGAGGTGT 3'
KRT6A qPCR (fw, rev)	5' GAGGAGATTGCTCAGAGAAGCC 3'	5' CAATCTCCTGCTTGGTGTTGCG 3'
SPRR3 qPCR (fw, rev)	5' TGAACCAGGCAGCATCAAGGTC 3'	5' GAAGGACATGGCTCTGGTAGCT 3'
TROY qPCR (fw, rev)	5' CTGCTCATCCTCTGTGTCATCTATTG 3'	5' GCCGTTGTACTGAATGTCCTGTG 3'
OLFM4 qPCR (fw, rev)	5' GACCAAGCTGAAAGAGTGTGAGG 3'	5' CCTCTCCAGTTGAGCTGAACCA 3'
LYZ qPCR (fw, rev)	5' ACTACAATGCTGGAGACAGAAGC 3'	5' GCACAAGCTACAGCATCAGCGA 3'
Recombinant DNA		
hs-ANPEP	ACDBio	477531
hs-ASCL2-C2	ACDBio	311011-C2
hs-BEST4	ACDBio	481501
hs-CEACAM6	ACDBio	403031
hs-CFC1-C2	ACDBio	529761-C2
hs-CFTR-C2	ACDBio	603291-C2
hs-CHGA-C2	ACDBio	311111-C2

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
hs-DMBT1-C3	ACDBio	478711-C3
hs-GHRL	ACDBio	455131
hs-GPR112-E6	ACDBio	411781
hs-GUCA2B	ACDBio	433011
hs-KRT20-C3	ACDBio	549681-C3
hs-LHB	ACDBio	877421
hs-LIPF-C4	ACDBio	300031-C4
hs-MUC6-C2	ACDBio	312901-C2
hs-PTF1A	ACDBio	524421
hs-REG4	ACDBio	312071
hs-SLC5A5-C4	ACDBio	417531-C4
Software and algorithms		
RaceID3	Herman et al. (2018)	https://cran.r-project.org/web/packages/RaceID/
StemID	Grün et al. (2016)	https://cran.r-project.org/web/packages/RaceID/
CFX manager software	Bio-Rad	N/A
GraphPad PRISM 8	GraphPad	N/A
Las X	Leica	N/A
Fiji	NIH, Fiji developers	https://imagej.net/Fiji
Rstudio	Rstudio	https://rstudio.com/
Adobe Illustrator	Adobe inc.	N/A
Other		
EVOS FL Auto 2 Cell Imaging System	Thermo Fisher Scientific	N/A
SP8 confocal microscope	Leica	N/A
DM4000	Leica	N/A
FACS Aria	BD Biosciences	N/A
FACS BD Influx	BD Biosciences	N/A
FACS Jazz	BD Biosciences	N/A

RESOURCE AVAILABILITY

Lead contact
Further information and requests for reagents may be directed to and will be fulfilled by the lead contact, H. Clevers (h.clevers@hubrecht.eu).

Materials availability
This study did not generate new unique reagents.

Data and code availability
All mouse scRNA-seq data have been deposited at the Gene Expression Omnibus (GEO) under the accession number GSE157694. All human scRNA-seq data are available at the European Genome-Phenome Archive (EGA) under the accession number EGAS00001004695.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human patients
The study was approved by the ethical committee of the University Medical Center Utrecht (UMCU; the Netherlands) and was in accordance with the Declaration of Helsinki. It is also according to Dutch law and compliant with all relevant ethical regulations

regarding research involving human participants. Patients planned for upper gastrointestinal endoscopy at the UMCU for Barrett's esophagus surveillance were consented for four extra biopsies. A standard biopsy forceps was used to take one or two biopsies of healthy esophageal squamous, gastric or duodenal epithelial tissue. Particular care was given to collect biopsies of the gastric anatomical regions to minimize false-sampling as much as possible. Gastric cardia biopsies were taken within the 1 cm rim bellow the gastresophageal junction, identified based on the first gastric folds and the palisade vessels in Barrett's esophagus patients. Gastric corpus biopsies were taken from gastric folds, located 5 cm proximal away of the antral-corpus junction, and pyloric biopsies were taken distal of incisural angularis. All biopsies were directly placed in DMEM containing glutamine, HEPES, Pen/Strep, Primocin and RhoK (Y-27632) inhibitor and then transferred to the laboratory for further processing. A detailed overview about sex, age and diagnosis of every included patient is provided in Table S1.

For immunostainings, sections of formalin-fixed, paraffin embedded human esophageal, gastric and duodenal tissue were obtained from resections performed at the University Medical Center Utrecht, the Netherlands. Anonymized archival pathology material was used according to the guidelines of the UMC Utrecht's Research Ethics Committee (Coebergh et al., 2006).

Mice
Male mice of the C57BL/6 strain at the age of 8-10 weeks were used for dissection of the Corpus and Pylorus regions of the stomach. All mice were bred under a project license granted by the Dier Experiment Commissie / Animal Experimentation Committee (DEC) or Central Committee Animal Experimentation (CCD) of the Dutch government and approved by the Hubrecht Institute Animal Welfare Body (IvD).

METHOD DETAILS

Sample preparation for scRNA-seq
Esophageal biopsies were cut into small pieces and incubated by shaking at 160 rpm for 30 min at 37°C in 0.125% diluted trypsin solution. Freed single cells in the supernatant were collected and subsequently prepared for fluorescent activated cell sorting (FACS) by staining with an EPCAM antibody (see below). Gastric cells from human biopsies as well as dissected mouse stomach regions were isolated as previously described (Bartfeld and Clevers, 2015). In short, gastric human biopsies or mouse tissue regions were cut into pieces and incubated by rotating in chelating solution at 4°C for 30 min. Thereafter, gastric glands were squeezed out of the tissue pieces by applying pressure on a glass slide, which was placed on top, and collected for single cell digestion for 5 min in TrypLE solution. The resulting single-cell suspension was then used for FACS sorting (see below). Duodenal biopsies were enriched for crypt and villus fractions by combining the previously described protocols for crypt isolation (Sato et al., 2011b) with gastric gland isolation (Bartfeld and Clevers, 2015). In short, biopsies were cut into pieces and incubated at 4°C for 30 min with occasional vigorous shaking in chelating solution. The supernatant was enriched of cells from the villus region, which were collected and processed separately. Cells of the crypt regions remained mainly in the tissue pieces. These cells were released by applying pressure on a glass slide, positioned on top of the tissue pieces, similar to the procedure used for gastric gland isolation. The duodenal villus and crypt fractions were both digested for 5 min at 37°C using TrypLE solution. Cells from the villus region were stained with the EPCAM antibody prior to FACS sorting.

Cells from esophageal, gastric and duodenal biopsies were then FACS-sorted into 384-well plates as previously described (Grün et al., 2015; Muraro et al., 2016). In short, DAPI was added to the single-cell suspension prior to FACS sorting of DAPI-negative live cells. Of note, gastric and duodenal samples also yielded a significant fraction of immune cells, which were absent for the esophageal samples. This is most likely explained by the different FACS-sorting strategies. Esophageal biopsies were digested as a whole, and EPCAM FACS sorting was required to enrich for esophageal epithelial cells, whereas gastric and duodenal epithelial cells were already enriched by isolating the gastric glands and duodenal crypts.

RNA *in situ* hybridization (RNAScope)
The staining was performed using RNAScope® Multiplex Fluorescent Reagent Kit v2 (Advanced Cell Diagnostics) according to the manufacturer's protocol (standard condition) (Wang et al., 2012). Images were acquired by a Leica SP8 confocal microscope. A list of the ordered probes is provided below in the Key resources table.

Immunohistochemical analysis
Organoids were fixed in 4% paraformaldehyde followed by dehydration, paraffin embedding and sectioning. Paraffin-embedding and sectioning of endoscopic resections of healthy human donors was performed under standard conditions at the Pathology Department of the UMC Utrecht. The transfer of human histological sections was approved by the responsible ethical committee. Standard immunohistochemical analysis was performed using antibodies as specified below in the Key resources table. In short, histological sections were deparaffinized in xylene and re-hydrated. Thereafter, the sections were boiled for 20 min in 0.1 M citrate buffer pH 6 for antigen retrieval. After blocking with 2% Donkey serum in PBS, the sections were incubated with the primary antibody overnight at 4°C and subsequently with the secondary antibody for 1 h at RT. Immunofluorescent images were acquired using the Leica SP8 confocal microscope. HRP-stained images were recorded using the Leica DM4000 microscope.

Organoid cultures

Esophageal cells were isolated from fresh biopsies as described above and plated in basement membrane extract (BME) droplets without prior FACS sorting. Healthy esophageal organoids were established in ESO culture medium (Advanced DMEM/F12 supplemented with glutamine, HEPES and Pen/Strep, 10% R-spondin conditioned medium, 1% Noggin, B27 w/o vitamin A, 5mM nicotinamide, 10 μM FSK, 100 ng/ml FGF10, 25 ng/ml FGF7, 500 nM A83-01, 1 μM P38 inhibitor and Primocin. For the first three days of establishing new cultures, Fungin was added to prevent fungal contamination during culture initiation. RhoK inhibitor (Y-27632) was added every time, when single cells were platted, but was not required for the maintenance phase, when medium was exchanged. Esophageal organoids were dissociated into single cells using TrypLE solution and were split in a 1:5 ratio once per week. Of note, after 6-7 passages the organoid growth slowed down and the cultures subsequently stopped growing.

Duodenal organoids were established and cultured as previously described (Sato et al., 2011a). In short, crypt cells were platted in BME and covered with INT media (Advanced DMEM/F12 supplemented with glutamine, HEPES and Pen/Strep, 20% R-spondin conditioned media, 1% Noggin, B27 with vitamin A, 10 mM Nicotinamide, 0.5 nM WNT Surrogate, 50 μg/ml EGF, 500 nM A83-01, 10 nM Prostaglandine E2, 1 μM P38 (SB 202190 inhibitor) and Primocin. Duodenal organoids were split every week as previously described (Sato et al., 2011a).

RT-qPCR analysis

Organoids were resuspended in the RLT buffer and RNA was subsequently isolated using the RNAeasy kit (QIAGEN). RNA was then converted to cDNA using NEB GoScript™ reverse transcription system. Quantitative PCR (qPCR) was performed with the isolated cDNA as previously described (Hatzis et al., 2008). The Biorad CFX Maestro software was used to analyze to qPCR data. Primer sequences are listed below in the Key Resources Table.

Library preparation

The RNA of single cells was processed according to the CEL-seq2 protocol (Muraro et al., 2016). In short, total RNA of single cells was reverse-transcribed into cDNA in a 384-well plate, thereafter pooled and *in vitro* transcribed by T7 polymerase overnight. After fragmentation, the amplified RNA was once more reverse-transcribed before the library for next-generation sequencing was prepared. The quality of the fragmented RNA and cDNA library was checked by analysis with the PICO RNA or DNA hypersensitivity chip (Agilent), respectively. The samples were then paired-end sequenced on NextSeq500 (2x 75 bp) by the Utrecht Sequencing Facilities (http://www.useq.nl).

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of single-cell RNA sequencing data

Sequencing reads were mapped to the human transcriptome as previously described (Muraro et al., 2016). Reads that mapped equally well to multiple locations were discarded. The UMI (unique molecule identifier)-normalized reads were used as inputs for the RaceID3 algorithm (supplied as CRAN package in R, https://cran.r-project.org/web/packages/RaceID/index.html) (Herman et al., 2018). Within RaceID, the data was filtered for cells with more than 1,500 transcripts per cell and for genes that were expressed by at least three transcripts in at least one cell. An overview of cells retained after filtering, including average transcript counts is provided in Table S2. Genes associated with clustering artifacts such as mitochondrial genes, MALAT1 and KCNQ1OT1 (Gehart et al., 2019; Grün and van Oudenaarden, 2015) were excluded from cluster calculation using the built-in FGenes or CGenes function of the RaceID3 program package. Cluster calculation was performed using the 'hclust' method. To increase the sensitivity and to identify small cell clusters, the outlier calculation parameters were set to: probthr = 0.02 and outlg = 1. Of note, gastric tuft cells were not detected with this standard analysis, as described above. They were identified by changing two parameters of the 'clustexp' function: sat = FALSE and cln = 20. The 4 identified gastric tuft cells were then superimposed on the standard analysis output.

Differential gene expression was performed on the identified cell clusters, while excluding immune cells form the overall analysis (Table S4, S6, S8, and S11; Figures 2A, 3A, 4A, and S4A). Heatmaps were generated using 'plotmarkergenes' function (Figures 1C and 2B). The t-SNE plots were calculated with the 'plotexpmap' function and fraction dot plots using the 'fractDotPlot' function (Figures 5, 6, and 7).

The lineage tree (Figure S2E) was calculated according to the documentation of the StemID algorithm, provided in the RaceID3 R package.

Comparison of differential gene expression between the human and mouse stomach

The expression pattern of the main gastric cell types (parietal, chief, neck and pit cells) were compared between the mouse and human dataset. To this end, we determined the relative expression value for an individual gene in a specific cell type cluster by dividing the averaged normalized read counts of the gene in this cell type by the averaged normalized read counts in all remaining cells of the respective organ. We then selected genes that were > 2-fold expressed in a given cell type compared to the remaining stomach cells in each species and combined the lists of the two species (Table S9). Each gene was then analyzed for their expression in the respective cell types of the mouse or human stomach by visual inspection of the respective t-SNE maps (Figures 3G, 3H, S5, and S6).

Re-analysis of human colon scRNA-seq data

The mapped read counts were downloaded from GEO omnibus (accession number GSE116222). Cells of the healthy human colon, which end with .A1, .B1 or .C1, were extracted and used for further analysis. For consistency reasons, the same analysis pipeline (RaceID algorithm), that was used for the duodenal data, was applied to obtain t-SNE maps showing the expression of *BEST4*, *CFTR*, *GUCA2A*, *GUCA2B*, *OTOP2*, *OTOP3*, *NAPSB* and *KRT72* in the human colonic epithelium (Figures 4F and S7G).

Analysis of endocrine cells

The UMI-normalized read counts of the endocrine cells, as defined in the individual organ calculation, were combined and subjected to a separate RaceID3 analysis. In the initial analysis, some immature endocrine cells with low expression of mature endocrine marker genes were detected, which were removed from further analysis to prevent dilution of endocrine expression pattern. These more stringent criteria lead to the final RaceID3 analysis as depicted in Figure 5A. The identified cells were then superimposed on the overview analysis including all cells from esophagus, stomach and duodenum. All non-endocrine gastric or duodenal epithelial cells were then combined in the category called "other cell types" for the stomach or duodenum, respectively, which served as controls for the fraction dot plots (Figures 5B–5D and 6A). Common EEC and gastric EEC signatures were derived by calculating the differential gene expression signatures for all individual endocrine cells in their organ-specific epithelial context, which were then combined in one list for initial fraction dot plot analysis and subsequent visual selection for the different categories. Their specificity was also confirmed in t-SNE plots.

Analysis of tissue- and cell type-specific transporter gene expression

A list of SLC transporters was downloaded from https://www.bioparadigms.org, which was combined with annotated ABC transporters listed on https://www.guidetopharmacology.org. Enriched expression of transporters was determined by calculating average gene expression for each organ-specific cell clusters and using an expression cutoff of more than 0.02 to identify expressed genes. Subsequently, the gene expression levels were compared between cell clusters and organs. If transporters were expressed in more than 75% of the most prominent tissue-specific cell types in one organ (esophagus: all five cell clusters; stomach: gastric chief, neck, parietal, pit and differentiating cells; duodenum: stem and transit-amplifying cells as well as immature and mature enterocytes) and less than 30% of the main tissue-specific cell types in the other two organs, it was categorized as organ-enriched. Cell type-enriched transporters were derived by analyzing the gene expression levels of the indicated cell types within the respective tissue. All cell type-enriched or tissue-enriched transporters were further selected upon visual inspection for their correct assignment in fraction dot plots and t-SNE maps as shown in Figures 7A–7D and S8.

Quantification of MKI67 positive signal in esophageal COL17A1^{high} and KRT15^{high} basal cells

Images of MKI67 and COL17A1/KRT15 double staining were recorded using the Leica SP8 confocal microscope and were loaded into the image processing program, FIJI, for further analysis (Schindelin et al., 2012). The "cell counter" option was applied to count all COL17A1 or KRT15 positive cells as well as the overlapping MKI67 signal. For COL17A1 positive cells, a total of 540, 221 and 738 cells were counted from three patient resection specimen, respectively, and 292, 1'486, 491 and 425 cells from four patients for KRT15 positive cells. Graphs were then generated by importing the percentage of MKI67 positive cells into GraphPad Prism.