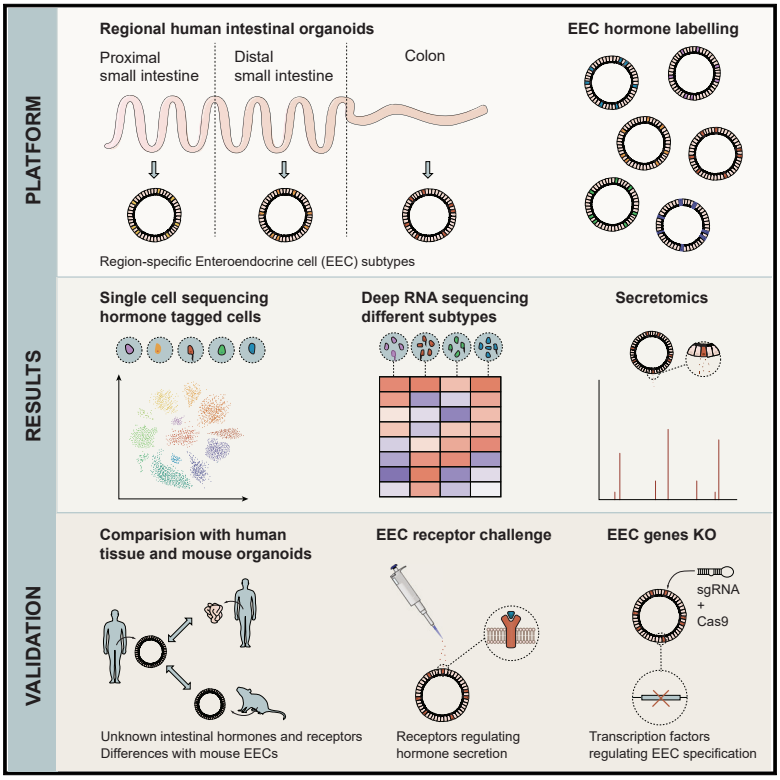
Cell

High-Resolution mRNA and Secretome Atlas of Human Enteroendocrine Cells

Graphical Abstract



Highlights

- A human organoid biobank combines hormone labeling and enteroendocrine cell generation
- Transcriptomic profiling of human enteroendocrine cells uncovers differences with mice
- Functional validation of EEC receptors and transcription factors
- Secretome analysis reveals the repertoire of enteroendocrine secreted products

Authors

Joep Beumer, Jens Puschhof, Julia Bauzá-Martinez, ..., Sarah A. Teichmann, Wei Wu, Hans Clevers

Correspondence

h.clevers@hubrecht.eu

In Brief

An organoid-based platform for studying human enteroendocrine cells, which sense intestinal content and release hormones to regulate many processes throughout the body, is developed by Beumer et al. and used to describe the landscape of mRNA expression and secreted products.

Beumer et al., 2020, Cell *181*, 1291–1306
June 11, 2020
https://doi.org/10.1016/j.cell.2020.04.036

CellPress

Cell

CellPress

Article

High-Resolution mRNA and Secretome Atlas of Human Enteroendocrine Cells

Joep Beumer,[1,2,14] Jens Puschhof,[1,2,14] Julia Bauzá-Martinez,[3,4,14] Adriana Martínez-Silgado,[1,2] Rasa Elmentaite,[5] Kylie R. James,[5] Alexander Ross,[6,7] Delilah Hendriks,[1,2] Benedetta Artegiani,[1,2] Georg A. Busslinger,[1,2] Bas Ponsioen,[8] Amanda Andersson-Rolf,[1,2] Aurelia Saftien,[1,2] Charelle Boot,[1,2] Kai Kretzschmar,[1,2] Maarten H. Geurts,[1,2] Yotam E. Bar-Ephraim,[1,2] Cayetano Pleguezuelos-Manzano,[1,2] Yorick Post,[1,2] Harry Begthel,[1,2] Franka van der Linden,[9] Carmen Lopez-Iglesias,[10] Willine J. van de Wetering,[1,10] Reinier van der Linden,[1,2] Peter J. Peters,[10] Albert J.R. Heck,[3,4] Joachim Goedhart,[9] Hugo Snippert,[8] Matthias Zilbauer,[7] Sarah A. Teichmann,[5,11,12] Wei Wu,[3,4,15] and Hans Clevers[1,2,13,15,16,*]

[1]Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences (KNAW) and UMC Utrecht, 3584 CT Utrecht, the Netherlands
[2]Oncode Institute, Hubrecht Institute, 3584 CT Utrecht, the Netherlands
[3]Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Padualaan 8, 3584 CH Utrecht, the Netherlands
[4]Netherlands Proteomics Centre, Padualaan 8, 3584 CH Utrecht, the Netherlands
[5]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton CB10 1SA, UK
[6]Department of Surgery, University of Cambridge, Cambridge CB2 0QQ, UK
[7]Department of Paediatrics, University of Cambridge, Cambridge CB2 0QQ, UK
[8]Oncode Institute, Center for Molecular Medicine, University Medical Centre Utrecht, Utrecht, the Netherlands
[9]Swammerdam Institute for Life Sciences, Section of Molecular Cytology, van Leeuwenhoek Centre for Advanced Microscopy, University of Amsterdam, Amsterdam, the Netherlands
[10]The Maastricht Multimodal Molecular Imaging Institute, Maastricht University, 6229 ER Maastricht, the Netherlands
[11]Theory of Condensed Matter, Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0HE, UK
[12]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton CB10 1SA, UK
[13]The Princess Maxima Center for Pediatric Oncology, 3584 CS Utrecht, the Netherlands
[14]These authors contributed equally
[15]Senior author
[16]Lead Contact
*Correspondence: h.clevers@hubrecht.eu
https://doi.org/10.1016/j.cell.2020.04.036

SUMMARY

Enteroendocrine cells (EECs) sense intestinal content and release hormones to regulate gastrointestinal activity, systemic metabolism, and food intake. Little is known about the molecular make-up of human EEC subtypes and the regulated secretion of individual hormones. Here, we describe an organoid-based platform for functional studies of human EECs. EEC formation is induced *in vitro* by transient expression of *NEUROG3*. A set of gut organoids was engineered in which the major hormones are fluorescently tagged. A single-cell mRNA atlas was generated for the different EEC subtypes, and their secreted products were recorded by mass-spectrometry. We note key differences to murine EECs, including hormones, sensory receptors, and transcription factors. Notably, several hormone-like molecules were identified. Inter-EEC communication is exemplified by secretin-induced GLP-1 secretion. Indeed, individual EEC subtypes carry receptors for various EEC hormones. This study provides a rich resource to study human EEC development and function.

INTRODUCTION

The principal function of the intestine is to digest food and absorb nutrients, but as the largest hormone producing organ, it also secretes hormones through its enteroendocrine cells (EECs) (Gribble and Reimann, 2017). EECs are rare secretory cells, comprising <1% of the epithelial cells. Apical EEC receptors sense chemicals in the intestinal lumen derived from food and microbiota (Furness et al., 2013). Hormones secreted by EECs signal to the local enteric nervous system and to distant organs including the pancreas and the brain, thus controlling food intake, insulin release, secretion of digestive enzymes, and bowel movement. EECs are therapeutic targets for metabolic diseases (i.e., obesity and diabetes), illustrated by recently introduced type 2 diabetes drugs that stabilize the hormone glucagon-like peptide 1 (GLP-1) or activate its receptor, leading to release of insulin from pancreatic β cells (Sharma et al., 2018).

EECs produce ~20 different hormones. GLP-1 and glucose-dependent insulinotropic peptide (GIP) are the incretin hormones that stimulate insulin secretion. The enterochromaffin (EC) cells produce 90% of body serotonin and regulate bowel movement (Worthington et al., 2018). Motilin (MLN) is a human EEC hormone, which controls gut contractions in the inter-digestive state (Worthington et al., 2018). Multiple hormones control appetite, including the appetite-inducing ghrelin (GHRL), the appetite-reducing peptide YY (PYY), and cholecystokinin (CCK). Gastrin (GAST) is secreted in the duodenum to control luminal acid by regulating proton secretion of stomach parietal cells. Somatostatin (SST) is an inhibitory peptide for most other intestinal hormones (Worthington et al., 2018).

Lgr5+ cells generate all differentiated intestinal cell types (Barker et al., 2007). The murine EEC subtypes are historically defined by their principle hormone product: L cells (Glp-1, Pyy), I cells (Cck), K cells (gastric inhibitory protein, Gip), N cells (neurotensin, Nts), S cells (secretin, Sct), EC cells (serotonin/5-HT), X cells (Ghrl), G cells (Gast), and D cells (Sst) (Engelstoft et al., 2013a; Gehart et al., 2019). Although this suggests that EEC phenotypes are hardwired, we have recently found that the crypt-villus BMP-signaling gradient induces hormone switching within individual murine EEC lineages (Beumer et al., 2018). The relative abundance of EEC subtypes greatly differs along the proximal-distal gastrointestinal axis. Studies on EECs have largely focused on murine models, exploiting a variety of reporter mice for subsets of EECs to monitor their responses to nutritional or genetic challenges (Goldspink et al., 2018). We have recently described the developmental hierarchy of murine subtypes EECs using a mouse model in which endogenous Neurogenin-3 expression, the main determinant of EEC fate, was coupled to the production of two separate fluorescent proteins with different half-lives (Gehart et al., 2019). Single-cell RNA sequencing of sorted EEC progenitors allowed for construction of a time-resolved development roadmap of the mouse EEC lineage.

Because the human diet and microbiome and that of rodents differ greatly (Nguyen et al., 2015), secretory hormone responses may also differ between these species. The study of human EECs is challenging because of their rarity and the lack of physiologically relevant *in vitro* models. Few human EEC-immortalized cell lines exist, and these differ substantially from their wild-type counterparts (Goldspink et al., 2018). There is currently no atlas of human EEC subtypes. Although some inducers of hormone secretion have been described in mice, there has been no experimental model to systematically assess such secretagogues for human EECs. Here, we describe an organoid-based platform to provide a detailed molecular and functional description of human EECs.

RESULTS

Production of Region-Specific Human EECs

Previous attempts to create human EECs *in vitro* have relied on growth-factor-based differentiation (Beumer et al., 2018) or overexpression of *NEUROG3*, the key transcription factor to instruct EEC fate (McCracken et al., 2014; Sinagoga et al., 2018). Both induced-pluripotent-stem-cell- (Zhang et al., 2019) and adult stem cell (ASC)-based (Chang-Graham et al., 2019) approaches allow studying of human EEC biology, such as modeling of hereditary *NEUROG3* mutations and virally induced serotonin release (Chang-Graham et al., 2019). However, imperfect differentiation and regional restriction of the donor material have limited these studies to a subset of human EECs.

To generate the full spectrum of human EECs, we established organoids from healthy adult proximal small intestine (duodenum), distal small intestine (ileum), and the ascending colon (Sato et al., 2011). These organoids were transduced with a doxycycline-inducible *NEUROG3* construct (Figure 1A). dTomato was inserted 3′ to the *NEUROG3* reading frame, separated by a self-cleavable P2A sequence to avoid creating a fusion protein. A 48 h-pulsed expression of *NEUROG3* in the basic medium "ENR" promoted the expression of the broad EEC marker chromogranin A (*CHGA)* (Figure 1B). Proximal small intestinal (SI) hormones such as *GAST*, *CCK*, and *MLN* were enriched in duodenal organoids, whereas *NTS*, *PYY*, and *GCG* were predominantly observed in distal SI organoids. Of note, *GCG* encodes the preproglucagon prehormone, a protein precursor to a set of hormones including GLP-1 (see below). *SST* was comparably expressed in proximal and distal organoids, consistent with its profile in the mouse gut. A recent single-cell RNA sequencing study generated the profile of 83 colonic EECs, suggesting that the human colon only contains serotonin-producing ECs and L cells positive for *GCG* and *PYY* (Parikh et al., 2019). Consistently, induced colon organoids only yielded serotonin-producing ECs and GCG-expressing EECs. We found that hormone expression peaked 5 days after initiation of *NEUROG3* expression (Figure S1A). Shorter doxycycline inductions stimulated the production of fewer EECs, whereas continuous doxycycline treatment throughout the differentiation window enhanced EC generation at the expense of L cells (Figure S1B). This suggests that Neurogenin-3 expression duration determines EEC subtype. We chose a 48 h doxycycline treatment in ENR, the maximum duration of *NEUROG3* expression *in vivo* (Gehart et al., 2019). Under these conditions, EECs in organoids displayed a normal morphology as visualized by transmission electron microscopy. Note the typical basal concentration of hormone vesicles (Figure 1C).

By immunofluorescent staining (Figure S1C), we observed mutually exclusive expression for MLN and GAST for GHRL and CHGA and for Serotonin and GLP-1, whereas a subset of GIP-positive cells co-expressed CCK. This closely resembled the co-expression patterns in mice (with the exception of MLN, a pseudogene in mice) (Haber et al., 2017) (Figure S1C). Virtually all EECs, as identified by the broad marker CHGA, were derived from NEUROG3-overexpressing cells (Figure S1C). A definitive hallmark of a mature EEC is its ability to secrete hormones. Indeed, exposure to forskolin, a stimulator of adenylate cyclase, greatly enhanced secreted GLP-1 levels (Figure 1D). We have previously reported that BMP signaling induces hormone expression changes in mature murine EECs in villi (Beumer et al., 2018). Consistent with our observations in murine EECs, we found that activation of BMP signaling enhances the expression of NTS while reducing GLP-1 (Figures S2A and S2B).

Because the initial expression of NEUROG3 occurs at random positions along the crypt axis in mice (Gehart et al., 2019), we

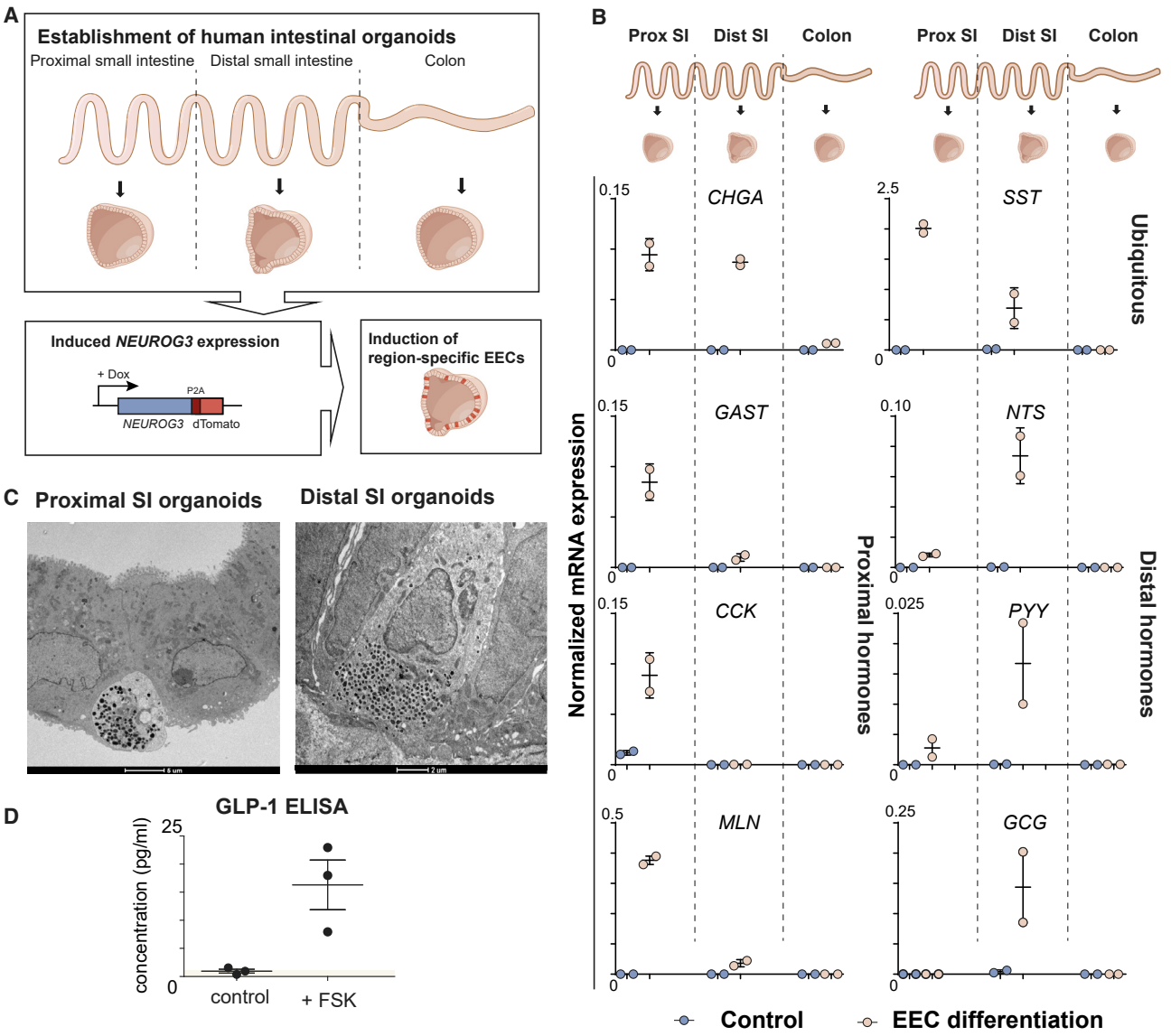


Figure 1. Production of Region-Specific Human Enteroendocrine Cells in Intestinal Organoids

(A) Schematic representation of the generation of region-specific enteroendocrine cells (EECs). Organoids are established from different regions of the intestinal tract of different patients, after which doxycycline (dox)–induced overexpression of neurogenin-3 (*NEUROG3*) can drive the production of EECs.

(B) qPCR analysis showing expression of hormones, control, or differentiation condition after a pulse of dox. Expression levels are normalized to *GADPH*. The experiment was performed in n = 2 independent experiments, and the mean expression and SEM are depicted.

(C) Transmission electron microscopy (TEM) of EECs in organoids showing polarized localization of hormone vesicles. Scale bar is 5 μm (left image) and 2 μm (right image).

(D) Concentration of supernatant GLP-1 determined by ELISA, in the absence (control) and presence of forskolin (FSK). The brown shaded area presents the detection threshold of GLP-1 (1 pg/mL). Forskolin induces secretion of GLP-1, confirming functionality of EECs. The experiment was performed in n = 3 independent experiments, and the mean concentration and SEM are depicted.

hypothesized that exposure to other crypt differentiation signals (i.e., Notch, Wnt) prior to this expression pulse could potentially determine EEC subtype. We modulated these signals prior to inducing NEUROG3 expression, mimicking the different initiation sites along the intestinal crypt axis (Figure S2C). As a control, we modulated the same signals after NEUROG3 induction (Figure S2D). Inhibition of Notch before or after expression of NEUROG3 did not affect EEC differentiation (Figure S2E). Inhibition of Wnt signaling before (but not after) the NEUROG3 pulse stimulated expression of *MLN* at the expense of *GCG*, while *SST* was unaffected (Figure S2E). Immunofluorescence revealed an increase in the number of MLN-producing cells rather than in the "per cell" expression levels (Figure S2F), resulting in a strong shift in L cell/M cell ratio (Figure S2G).

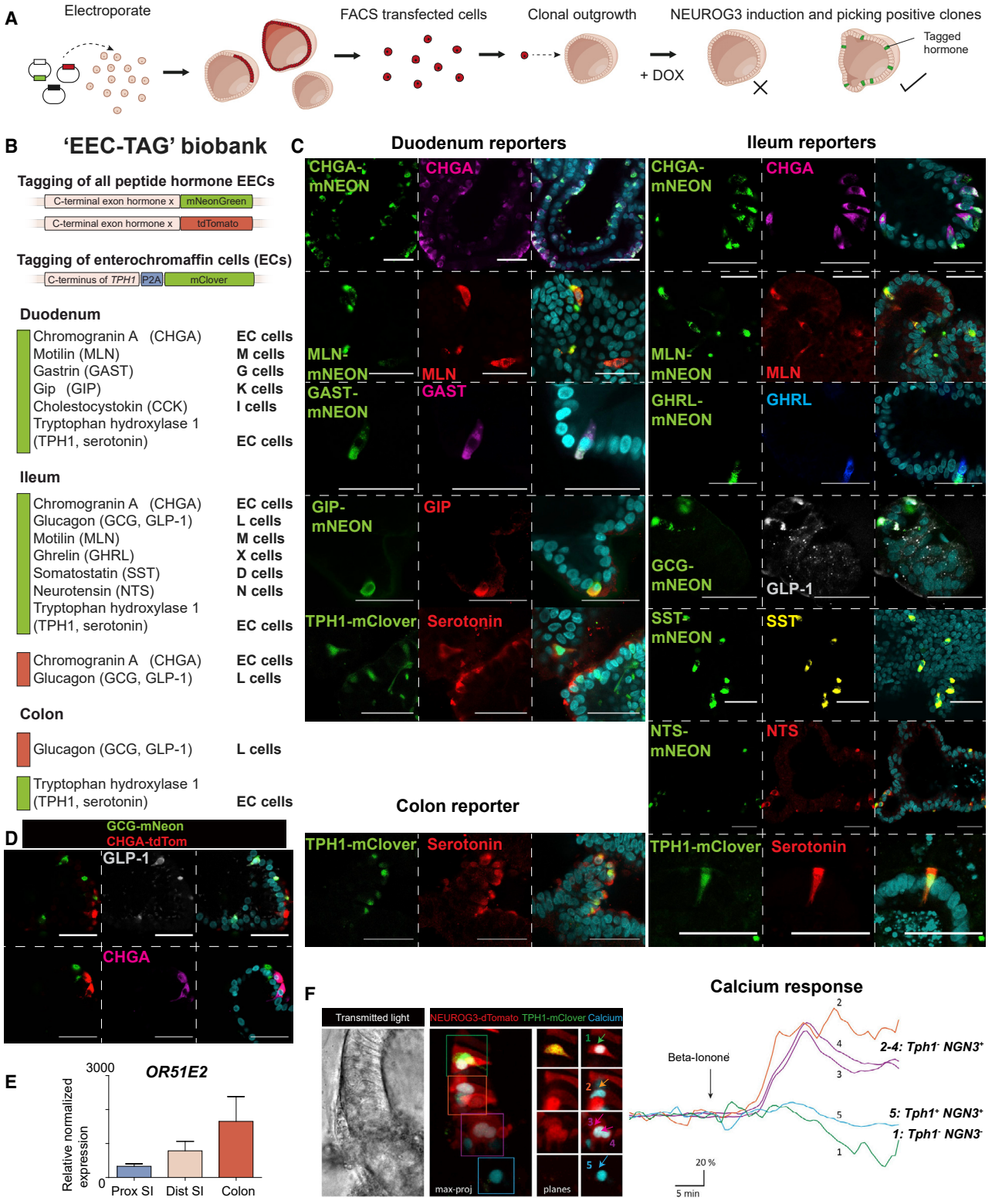


Figure 2. Generation of a Human Enteroendocrine Cell Organoid Toolbox

(A) Schematic workflow of generation of reporter organoids.

(B) Overview of EEC-TAG biobank-comprising reporters for hormones across human duodenum, ileum, and colon organoids. EECs are tagged with NHEJ (mNeon or tdTomato). TPH1-positive ECs are tagged using HDR with mClover.

(legend continued on next page)

Generation of Hormone Reporter Biobank EEC-TAG

Mouse models in which hormones are fluorescently tagged exist for several murine EEC hormones that were instrumental to study EEC subsets: Chga, Gcg, Gip, Cck, Ghrl, and Pyy (Engelstoft et al., 2013b; 2015; Gong et al., 2003; Parker et al., 2009; Reimann et al., 2008; Sommer and Mostoslavsky, 2014). CRISPR-Cas9 targeting followed by homology-directed repair (HDR) or non-homologous end joining (NHEJ) allows the introduction of exogenous genetic material (Bukhari and Müller, 2019; He et al., 2016; Schmid-Burgk et al., 2016). To mark ECs, we labeled tryptophan hydroxylase 1 (TPH1), the rate-limiting enzyme involved in serotonin synthesis. Using HDR, we tagged TPH1 with fluorescent mClover separated by a self-cleaving P2A site. We recently optimized a strategy for site-specific introduction of DNA into organoids using NHEJ (CRISPR-HOT) (Artegiani et al., 2020), which allowed fluorescently labeling of multiple secreted hormones (Figure 2A).

We generated a biobank of hormone reporter organoids termed EEC-TAG, consisting of duodenal, ileal, and colon organoids for a total of 10 human hormones (Figure 2B). All organoid lines showed complete overlap between fluorescent reporters and the corresponding hormone product (Figure 2C). The fluorescently tagged hormones localized to cytoplasmic vesicles. Serial tagging into the same organoid lines for multi-hormone labeling was feasible (Figure 2D).

Calcium signaling induces hormone secretion (Goldspink et al., 2018). We stably introduced a turquoise Ca^{2+} sensor (Tq-Ca-FLITS) into NEUROG3^{dTomato} TPH1^{mClover} reporter organoids using lentiviral transduction. The resulting genotype of the organoids is NEUROG3^{dTomato} TPH1^{mClover} CaFLITS^{Turqouise}. We stimulated the olfactory receptor OR51E2, of which the mouse homolog (*Olfr78*) is reported to be expressed in mouse EECs (Fleischer et al., 2015; Jovancevic et al., 2017). HEK cells overexpressing *OR51E2* elicit a calcium response when stimulated with the selective agonist beta-ionone (Pietraszewska-Bogiel et al., 2019). *OR51E2* is most strongly upregulated in distal organoids (Figure 2E). When reporter organoids were stimulated with beta-ionone, we observed calcium sparking in EECs that were TPH1-negative (Figure 2F), illustrating that sensors combined with hormone reporters can visualize activation of human EEC subtypes.

Single-Cell Transcriptomics

Studies have utilized reporter mice to enrich for hormone-producing cells when performing single-cell RNA sequencing. This approach cannot be used for primary human EECs, making the generation of a detailed atlas from SI tissue challenging. Murine EECs taken from primary tissue and from organoids are essentially identical (Gehart et al., 2019; Grün et al., 2015). We therefore exploited the human NEUROG3-induced organoids to perform single-cell RNA sequencing. NEUROG3 was induced in duodenal, ileal, and colon organoids in the absence or presence of BMP (to generate the crypt- and villus-"versions" of EECs; Figure S2). Data from 8,448 cells were generated and processed by sorting and robot-assisted transcriptome sequencing (SORT-seq; Muraro et al., 2016) (Figures S3A and S3B) and analyzed using RaceID3, a clustering method based on k-medoids (Herman et al., 2018). After filtering, a broad intestinal cell type atlas was built from 4,281 cells (Figures S3B and S3C). This atlas contained five large clusters: *CHGA*-positive EECs (2,255) and the following well-defined "contaminant" lineages: *FABP1*-positive enterocytes (585), *OLFM4*-positive stem cells (113), rare *MUC2*-positive goblet cells (33), *LYZ*/*MMP7*-positive Paneth cells (11), and several progenitor populations (Figure S3D).

Neuropeptide W (*NPW*) and *VGF*, recently observed in bulk EEC RNA sequencing (Roberts et al., 2019), were broadly expressed (Figure S3C). Although the function of VGF remains elusive, NPW is known to increase food intake when injected in the hypothalamus (Levine et al., 2005). Immunofluorescence confirmed protein expression of NPW by EECs in sections of human intestine (Figure 3A).

All EECs and their progenitors were identified by thresholding for expression of the generic EEC marker *CHGA* and thresholding against *MUC2*, *FABP1*, *LYZ*, and *OLFM4*. An EEC atlas was constructed from 2,255 cells (of which 805 cells were BMP-treated) (Figures 3B, 3C, and S4A). The major clusters overlapped with their mouse equivalents, and the different EEC subtypes followed regional identity (Figure S4B) (Haber et al., 2017). The human EEC atlas significantly differed from the mouse tissue EEC atlas (Figure S4C) (Gehart et al., 2019). To validate the *in vitro* EEC identities, we searched for EECs in a large single-cell dataset obtained from healthy and diseased human small intestines of various ages. Of 11,302 cells represented in this dataset, we derived mRNA signatures for 39 human EECs, underscoring the challenge of deriving single EEC mRNA signatures from human ileal biopsies (Figure S4D).

The largest cluster was formed by *TPH1*-expressing EC cells, highly expressing *CHGA* (as in mouse) and representing the most frequent EEC type *in vivo* (Figures 3B and 3C). ECs occurred in three "flavors": *REG4* high and *REG4* low cells (in cluster 4), which also exist in murine intestine (Haber et al., 2017). A third population of ECs, not observed in mice, expressed high levels of the secretogranin *SCG2* and occurred

(C) Immunofluorescent staining confirms faithful reporter activity (knockin left, stain middle, and merge on right). Reporter expression always overlaps with the corresponding hormone. Scale bar is 50 μm.

(D) Subsequent rounds of hormone tagging allow the generation of multiple-hormone reporter organoids. Immunofluorescent staining of GCG/CHGA double reporter organoid. Scale bar is 50 μm.

(E) qPCR analysis showing expression of the olfactory receptor *OR51E2* in different organoids enriched for EECs. Expressions levels are normalized to GADPH and relative to control organoids without EECs. The experiment was performed in n = 2 independent experiments, and the mean expression and SEM are depicted.

(F) Fluorescent image of a TPH1^{mClover} NEUROG3^{dTomato} organoid that is transduced with the turquoise calcium sensor Tq-Ca-FLITS, containing a nuclear localization signal. Five examples of nuclei are highlighted of which the calcium response is followed after treatment with beta-ionone, the agonist of OR51E2. EECs (marked in red) show increases in calcium flux ("2–4") with the exception of the TPH1^{+} cell ("1"). The non-EEC ("5∀) does not show calcium increases upon beta-ionone treatment.

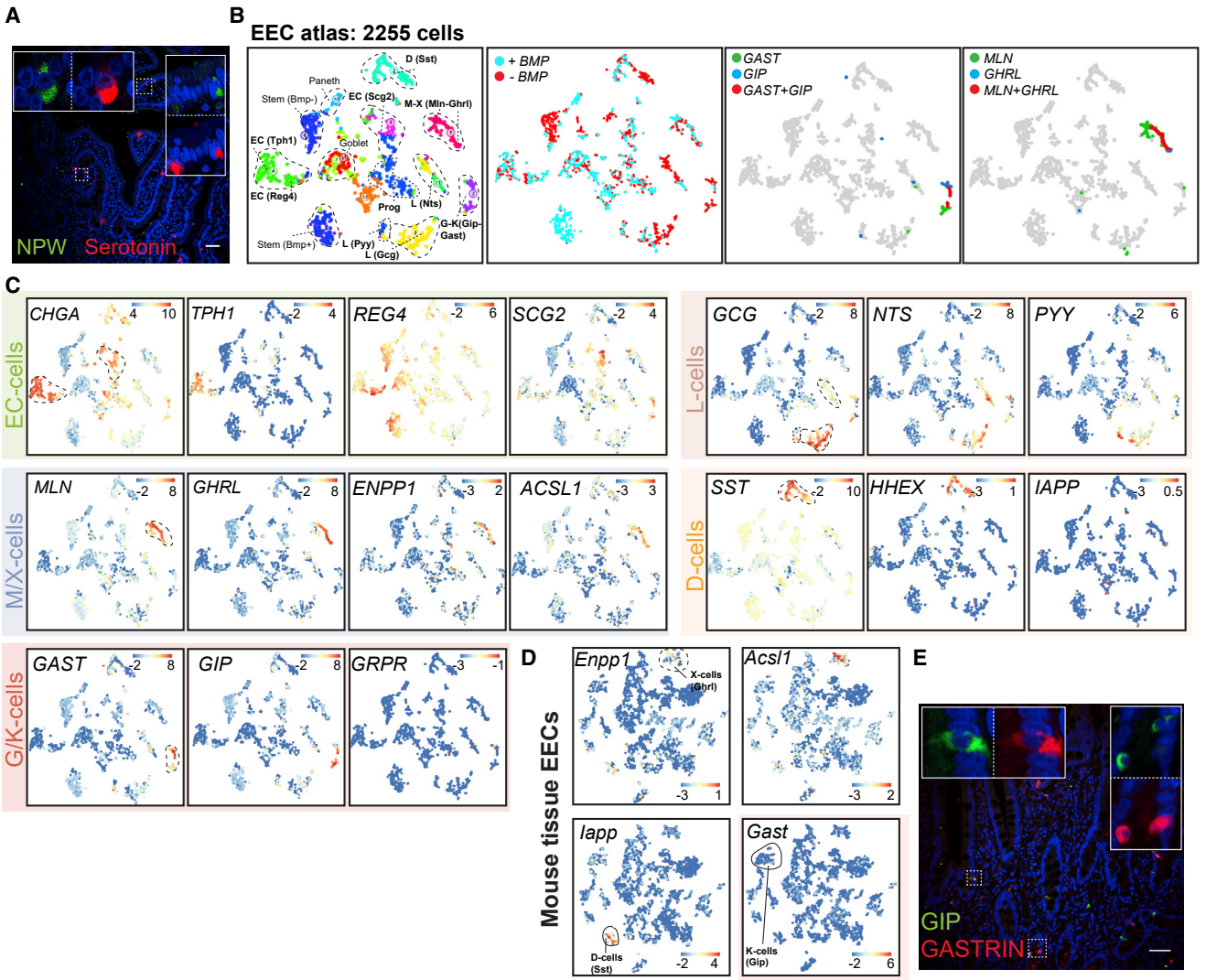


Figure 3. Single-Cell Transcriptome Atlas of Human Enteroendocrine Cells

(A) Immunofluorescent staining on human intestinal section (ileum) confirms EEC-specific expression of NPW. Scale bar is 50 μm.

(B) t-SNE map displaying the human EEC atlas (n = 2,255 cells). Different colors represent the 13 separate clusters, and BMP treated cells are highlighted. *GAST*- and *GIP*-positive cells (defined by a minimum of 25 unique transcripts per cell) show partly overlapping expression patterns (middle t-SNE map). *GHRL*- and *MLN*-positive cells (defined by a minimum of 25 unique transcripts per cell) also overlap partly (right t-SNE map).

(C) t-SNE maps displaying the expression levels of hormone and marker gene expression in the different human EEC subtypes from intestinal organoids. Bars display color-coded unique transcript expression (logarithmic scale).

(D) t-SNE maps displaying the expression levels of hormone and marker gene expression in murine tissue EECs. Bars display color-coded unique transcript expression (logarithmic scale).

(E) Immunofluorescent staining on duodenal sections confirms co-expression of GIP and GASTRIN. Scale bar is 50 μm.

mostly in proximal SI organoids (cluster 9) (Figures 3B and 3C). All human ECs highly expressed dopa decarboxylase (*DDC*) involved in serotonin biosynthesis, as well as *SLC18A1*, involved in serotonin transport (Figure S4E) (Lohoff et al., 2006). The prototypical EC markers *CHGB* and *GPR112* were broadly expressed by human ECs, as was the olfactory receptor OR51E1 (mouse homolog Olfr558), a marker of serotonin-producing neuroendocrine tumors in man (Figure S4E) (Cui et al., 2013).

G cells produce Gast and are largely restricted to the mouse stomach, whereas in man, expression continues more distally along the GI tract in EECs (Engelstoft et al., 2013a). Cells expressing *GAST* (cluster 3) co-expressed the receptor for Gast-releasing peptide, *GRPR*, a marker of G cells in mouse stomach (Figures 3B and 3C). *GAST*-expression often overlapped with high expression of the incretin *GIP* (same cluster 3), the main hormone product of murine K cells. We named these cells G/K cells (Figures 3B and 3C). In histological sections, we confirmed overlapping expression for these two hormones (Figure 3E). Some cells in cluster 3 almost exclusively produced either *GIP* or *GAST*, as seen on intestinal sections and in the atlas. The L

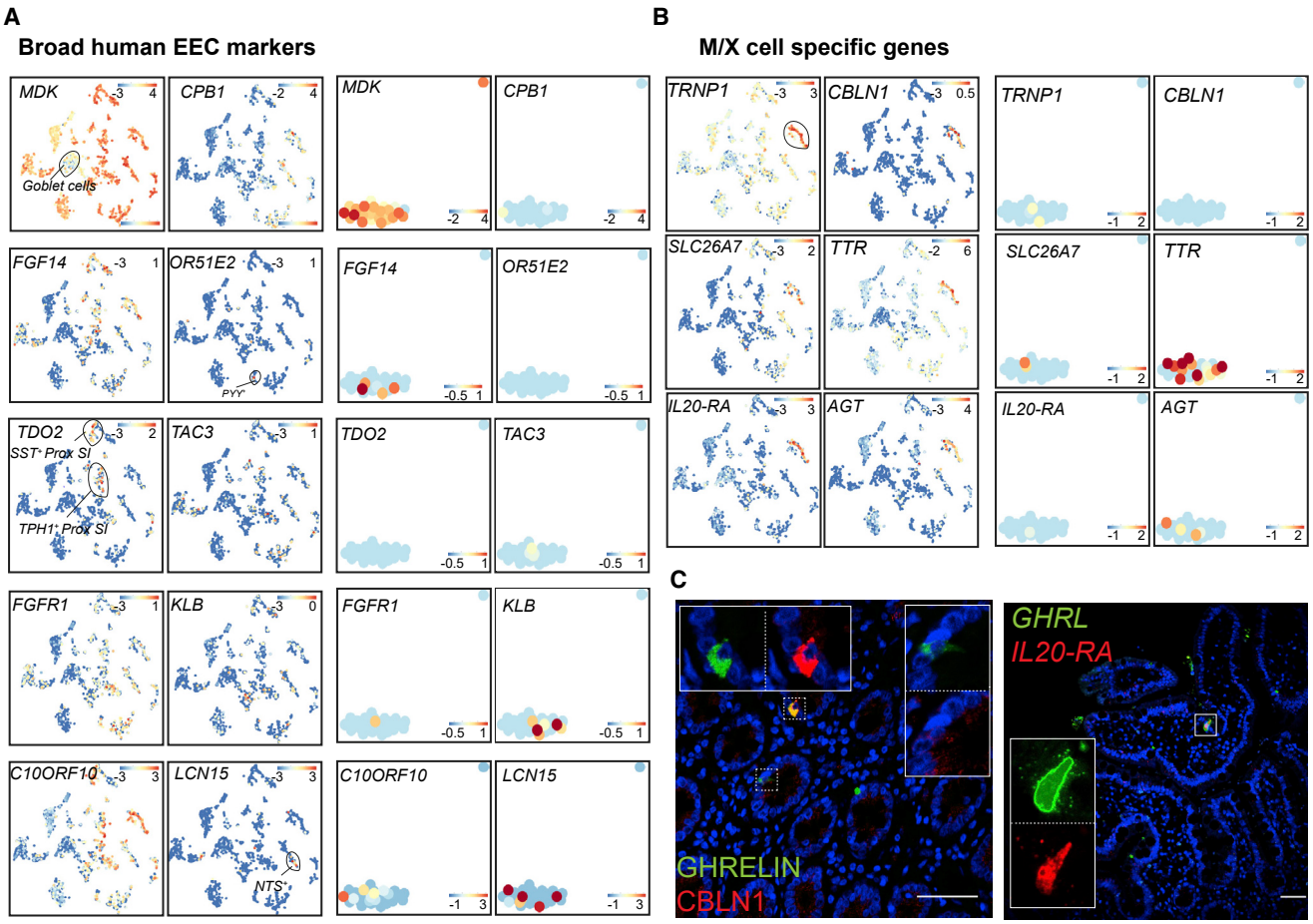


Figure 4. Human Enteroendocrine Cell Markers

(A) t-SNE maps displaying the levels of hormone and marker gene expression in the different human EEC subtypes from intestinal organoids (left) and intestinal tissue (right). Bars display color-coded unique transcript expression (logarithmic scale).

(B) t-SNE maps displaying the levels of hormone and marker gene in M/X cells. Bars display color-coded unique transcript expression (logarithmic scale).

(C) Immunofluorescent staining on intestinal sections confirms co-expression of GHRL and the EEC peptide CBLN1 (left panel). Fluorescent *in situ* hybridization shows that *GHRL*$^+$ cells express the cytokine receptor *IL-20RA* (right panel). All sections are from the human duodenum. Scale bar is 50 μm.

cell lineage clusters 8 and 13 displayed largely overlapping expression of *GCG*, *NTS*, and *PYY* (Figures 3B and 3C). Cluster 2 contained SST-positive D cells, also expressing the transcription factor *HHEX*. *HHEX* has been described in murine pancreatic and intestinal Sst-producing cells (Haber et al., 2017; Zhang et al., 2014). Notably, human D cells in tissue and organoids lacked expression of amylin (*IAPP*), a peptide hormone expressed in mouse D cells (Figures 3B, 3C, and S4D).

MIn$^+$ cells do not exist in mice. We identified a cluster of cells producing *MLN* and *GHRL* (cluster 5). A gradient from predominantly *MLN*$^-$ to predominantly *GHRL*-expressing cells can be observed in t-distributed stochastic neighbor embedding (t-SNE) space (Figures 3C and 3D). We termed these M/X or X/M cells (based on the highest expression of either *MLN* or *GHRL*, respectively) and speculate that these might represent different states of the same cell type. Indeed, BMP treatment reduced levels of *GHRL*, whereas *MLN* levels were slightly increased (Figure S6B). M/X cells were further characterized by *ENPP1* expression, a known regulator of insulin responses and extracellular ATP levels (Di Paola et al., 2011), similarly expressed by murine X cells (Figures 3C and 3D). GHRL requires a specific acyl modification by the acyl-CoA synthetase *Acsl1* in mouse stomach X cells (Bando et al., 2016). Human M/X cells and mouse intestinal X cells both expressed high levels of *Acsl1* (Figures 3C and 3D).

Genes Uniquely Expressed by Human EEC Subtypes

We next searched for EEC genes differentially expressed between human and mouse (Figure S4F). The heparin-binding growth factor midkine (*MDK*) was highly expressed by all human EEC types but not by, e.g., goblet cells (Figure 4A). MDK is a reported biomarker of human intestinal neuroendocrine tumors (Edfeldt et al., 2017). Midkine has been associated with obesity and inhibits insulin signaling in adipocytes (Fan et al., 2014). The carboxypeptidase *CPB1* was produced by most EECs (highest in M/X cells), with the exception of ECs (Figure 4A). Carboxypeptidases are typically involved in hormone processing (Sapio

and Fricker, 2014). Expression of Cpb1 has been observed in the rat pancreas (Yu et al., 2017). *FGF14* is a human pan-EEC marker—with very limited expression in murine EECs—and belongs to a set of intracellular FGFs, that play a role in the clustering of ion channels in neurons (Figures 4A and S4F) (Pablo and Pitta, 2017). The olfactory receptor *OR51E2* was sporadically expressed by different EEC subtypes, with highest levels occurring in *PYY*$^{+}$ cells (Figure 4A). The mouse homolog Olfr78 was lowly expressed in ECs only (Figure S4F). The enzyme tryptophan 2,3-dioxygenase (*TDO2*) was found in duodenal EECs from the proximal intestine (Figures 4A and S4A). TDO2 can metabolize tryptophan through the kynurenine pathway and is one of the primary regulators of availability of this amino acid. Tryptophan is the precursor of serotonin and *Tdo2* knockout mice experience increased serotonin levels (Too et al., 2016), suggesting that Tdo2 could locally regulate serotonin production in the gut. We noted the tachykinin peptide-coding *TAC3 w*as a broadly expressed gene in human EECs, whereas the mouse homolog *Tac2* is not expressed in the murine intestine (Figure 4A). *TAC3* codes for neurokinin B and has been described as a regulator of secretion of gonadotropin-releasing hormone in the human hypothalamus (Sanger, 2004). However, the main receptor for NKB, NK_3 (coded by *TACR3)*, has been implicated in the regulation of gastrointestinal motility (Sanger, 2004). The hepatokine FGF21 is a regulator of blood glucose. Several FGF21 mimetics are currently being tested for the treatment of diabetes (Kuro-O, 2019). Although the receptors for FGF21 are described as a complex of FGFR1 and B-Klotho (KLB), the site of action of FGF21 is debated. We observed broad expression of *FGFR1* and *KLB* by human EECs, suggesting that FGF21 effects could be partially mediated through the gut (Figure 4A). *Fgfr1* is absent in murine EECs, whereas *Klb* is expressed at very low levels (Figure S4F). We noted an upregulation of multiple hormones after FGF21 treatment, pointing to a potential role of signaling through FGFR1/KLB in EEC function (Figure S4G). *C10ORF10* (also known as *DEPP1*) was widely expressed by human EECs. This gene is negatively regulated by insulin in liver and adipocyte tissue; its product controls the ratio between ketogenesis and gluconeogenesis (Li et al., 2018) (Figure 4A). Finally, *LCN15* was produced by *NTS*$^{+}$ cells (Figure 4A). LCN15 is a lipocalin and one of the strongest glucose-regulated genes in Caco-2 cells (Boztepe and Gulec, 2018). Although some lipocalins have been implicated in insulin resistance, LCN15's function remains unknown.

We then focused on unique genes expressed by M/X cells but absent in murine X cells. *TRNP1*, involved in cortical folding in the brain (Stahl et al., 2013), was the only transcription factor specific to M/X cells (Figure 4B). A putative hormone, precerebellin 1 (*CBLN1*), was expressed in all M/X cells (Figure 4B). *CBLN1* stimulates food intake upon intracerebroventricular injection (like GHRE) (Gardiner et al., 2010). We confirmed CBLN1 expression in human GHRL^{+} cells *in vivo* using immunofluorescence (Figure 4C). We noted that M/X cells expressed the receptor for cytokines of the IL10-family (IL20-RA), an observation confirmed *in vivo* (Figures 4B and 4C). We detected high expression of the peptide hormone angiotensin (*AGT*), a regulator of blood pressure, but also of contraction of the human intestinal musculature (similar to motilin) (Ewert et al., 2006) (Figure 4B). Finally, M/X cells displayed the highest expression of all EECs of the sulfate transporter *SLC26A7* and of T_4 - and retinol-binding transthyretin (*TTR*) (Figure 4B). We confirmed the expression of these EEC genes in the (limited) number of single-cell sequenced human ileal EECs (Figures 4A and 4B).

To identify heterogeneity among the different EEC subtypes, we subclustered cells sorted from organoids carrying the individual hormone reporters. Expression of the fluorescent reporters directly correlated with the levels of the pertinent hormone transcripts within the same cell (Figure S5A). A substantial number of the cells sorted for MLN-reporter expression (yet with low MLN expression) were L cells (Figures S5A and S5B). Surprisingly, we identified a rare subcluster of *GCG*+-reporter cells that highly expressed *PPY* (Cox, 2007), a well-described pancreas hormone involved in appetite regulation, never seen in human or mouse small intestine. We confirmed its expression and partial overlap with GLP-1 by staining on human intestinal sections (Figures S5B and S5C).

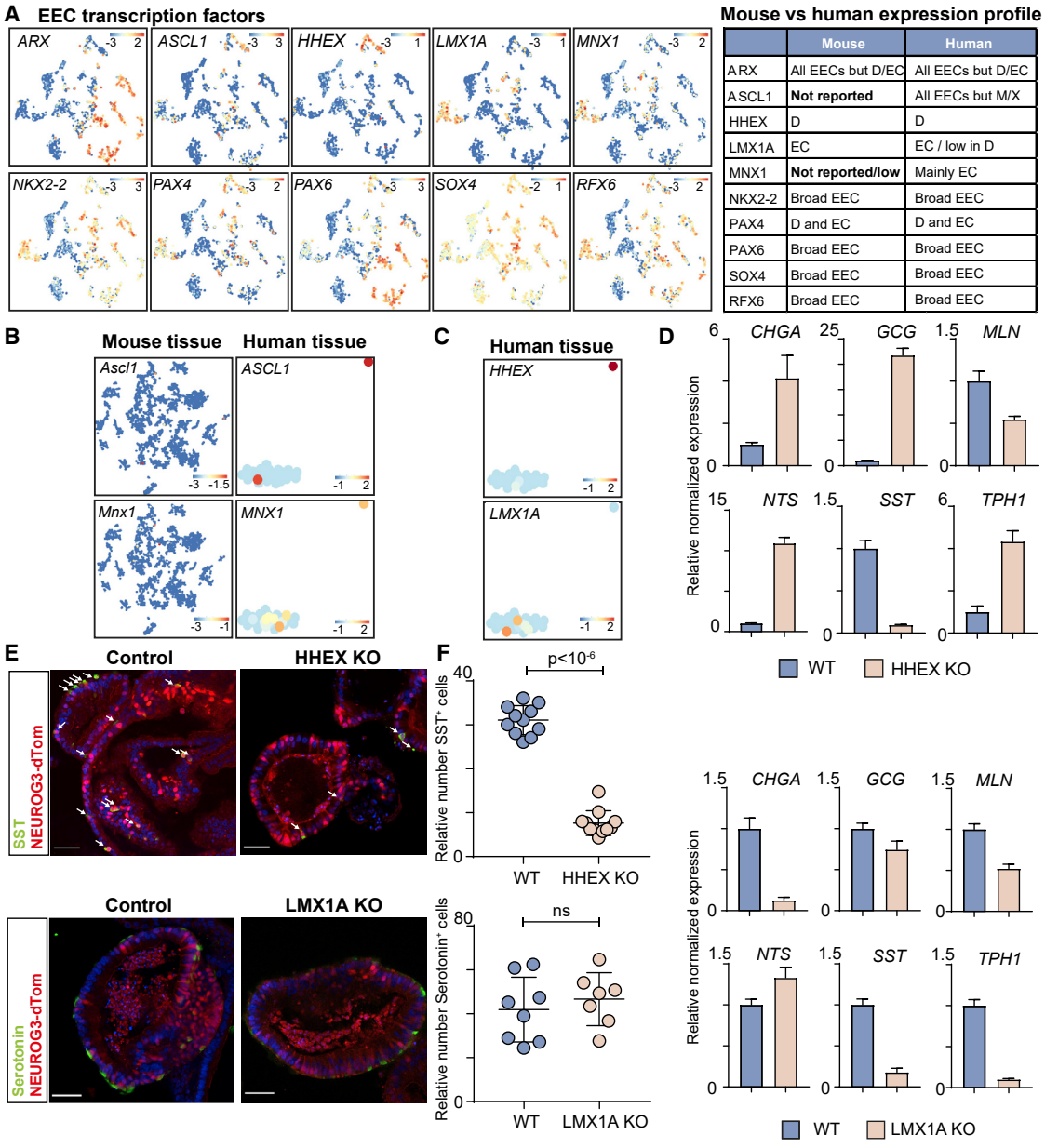
Transcriptional Networks

We analyzed expression of transcription factors known from mice to specify individual lineages (Figure 5A). PAX4 specifies D/EC cells, whereas expression of ARX promotes all other EEC fates (Beucher et al., 2012), in agreement with our expression profiles. HHEX and LMX1A defined human D and EC lineages respectively, consistent with mouse (Figure 5A) (Gross et al., 2016). The broad murine EEC transcription factors NKX2-2, PAX6, SOX4, and RFX6 (Gehart et al., 2019) were ubiquitously expressed in human EECs. We additionally identify ASCL1 as a broad human EEC transcription factor, absent from M/X cells and from all mouse EECs (Figures 5A and 5B). Ascl1 is expressed in endocrine cells in murine lung (Borges et al., 1997). *MNX1* was highly expressed by human ECs; it promotes neonatal diabetes when mutated (Figures 5A and 5B) (Pan et al., 2015). *MLN*$^{+}$ cells developmentally resembled murine Ghrl-producing X cells.

We chose to knock out the EC-specific *LMX1A* gene and the D-cell-specific *HHEX* gene (Figures 5A and 5C). Organoids were transiently transfected with a Cas9-EGFP coding plasmid that included the site-specific guide RNA (gRNA) (Ran et al., 2013). Genotyping of clonal organoids was performed to identify homozygous loss-of-function alleles. *Lmx1a*-null mice die shortly after birth, lacking intestinal *Tph1* and *Chga* expression, indicative of EC loss (Gross et al., 2016). *LMX1A* mutant human organoids displayed a strong reduction in *TPH1* (Figure 5D). We also observed a milder reduction in *SST* derived from D cells. In contrast to mouse EECs, human *LMX1A* is also expressed in D-cells (Figure 5A)

HHEX has been linked to type 2 diabetes (Scott et al., 2007). Although loss of *Hhex* in mice impairs the function of Sst-producing cells in pancreatic islets, effects in the murine intestinal tract were not described (Zhang et al., 2014). *HHEX* gene disruption blocked the production of *SST* (Figure 5D). The most striking increase was observed in *GCG* expression (over 20-fold). In *Hhex*-knockout mice, pancreatic glucagon similarly increases (Zhang et al., 2014).

Immunofluorescent staining in *HHEX* and *LMX1A* organoids revealed a 4-fold reduction in the number of SST^{+} D cells upon



	Mouse	Human
ARX	All EECs but D/EC	All EECs but D/EC
ASCL1	**Not reported**	All EECs but M/X
HHEX	D	D
LMX1A	EC	EC / low in D
MNX1	**Not reported/low**	Mainly EC
NKX2-2	Broad EEC	Broad EEC
PAX4	D and EC	D and EC
PAX6	Broad EEC	Broad EEC
SOX4	Broad EEC	Broad EEC
RFX6	Broad EEC	Broad EEC

Figure 5. Transcriptional Networks in Human EECs

(A) t-SNE map displaying the expression level of EEC lineage transcription factors. Bars display color-coded unique transcript expression (logarithmic scale). The murine and human patterns of expression of these genes among EEC subtypes are depicted (right table).

(B) t-SNE maps displaying the expression level of EEC transcription factor *ASCL1* and *MNX1* in mouse and human EECs. Bars display color-coded unique transcript expression (logarithmic scale).

(C) t-SNE maps displaying the expression level of EEC transcription factors *HHEX* and *LMX1A* in human tissue EECs. Bars display color-coded unique transcript expression (logarithmic scale).

(D) qPCR analysis showing expression of hormones in wild-type organoids and LMX1A and HHEX knockout (KO) organoids. Expression is normalized to GADPH, and relative to wild type. The experiment was performed as a technical duplicate, and the mean expression and SEM are depicted.

(E) Immufluorescent staining on wild-type and *HHEX* and *LMX1A* knockout organoids. D cells are reduced after *HHEX* loss, whereas EC cells do not decrease upon *LMX1A* knockout. Scale bar is 50 μm.

(F) Quantification of (E). Number of hormone positive cells are counted on organoid sections and shown normalized to number of tdTomato^{+} cells and relative to WT. At least n = 7 organoids were counted per condition; p value was calculated using an unpaired two-tailed t test.

HHEX loss, suggesting impaired allocation of progenitors into this lineage (Figures 5E and 5F). Serotonin-producing cells did however not decrease in *LMX1A* knockout organoids, despite a decline in *TPH1* expression (Figures 5D–5F). We conclude that *LMX1A* is not directly important for allocation to the EC fate.

BMP Signaling as Regulator of Hormone Switching in Human EECs

Crypt-villus gradients were observed for human hormones such as *GCG* (Figure S5D). We interrogated BMP dependency of hormone gene expression in the single-cell atlas. BMP activation induced *NTS* in L cells at the expense of *GCG* (Figure S5E). Live-cell imaging of *GCG*-reporter organoids confirmed that BMP activation decreased reporter expression in individual L cells (Figure S5F). Additionally, we observed BMP-mediated repression of *GHRL* in M/X cells, accompanied by a mild increase in *MLN* expression (Figure S5E). In murine intestine, we have found that expression of GHRL diminishes with migration of the X cell along the crypt-villus axis (Gehart et al., 2019). Thus, human *MLN*/*GHRL*-producing EECs appeared to undergo a BMP-controlled switch in hormone expression as previously described in mouse.

High-Definition Transcriptomic and Proteomic Profiling of EECs

Transcriptomics of pooled cells has a superior sensitivity compared to single-cell RNA sequencing. We thus generated a deep transcriptomic signature of sorted and pooled EC, L, and M cells (Figure S6A). In addition, CHGA-mNeon^{+} cells were sorted to generate a broad EEC signature. We identified the top 20 uniquely expressed markers from the RNA sequencing dataset for each population (Figure S6B). We thus uncovered multiple EEC subtype features, unnoticed in the single-cell atlas. The transcription factor *IRX3*, member of the Iroquois homeobox family, was one of the most defining markers of TPH1+ cells, yet has not been described in murine EECs (Haber et al., 2017) (Figure S6B). IRX3 has gained attention as a neuronal regulator of energy balance, and genetic variants in *IRX3* associate with obesity in humans (Schneeberger, 2019).

We analyzed our bulk transcriptomic datasets for subtype-specific receptors. We noted conserved expression of receptors known from mouse EECs, including *FFAR2* (broad EEC), *GPBAR1* (L cell), *SSTR5* (L cell), *OR51E1* (mouse homolog *Olfr558*; EC), *ADGRG4* (*Gpr112;* EC), and the extracellular calcium sensor *CASR* (broad EEC) (Furness et al., 2013) (Figure S6C). Human EECs expressed multiple orphan receptors, such as *GPR162* (L cells), not found in mice (Figure S6C), and reported to be expressed in brain to regulate food intake. Genetic variants in *GPR162* are linked to glucose deregulation (Caruso et al., 2016). *GPR68* is an orphan GPCR uniquely expressed by ECs (Figure S6C). The orphan peptide CART (cocaine- and amphetamine-regulated protein) activates GPR68 (Foster et al., 2019); it has a role in the regulation of anxiety, reward, and feeding behaviors (Shcherbina et al., 2018). We find broad expression of the subunit of the GABA-B receptor *GABBR2*, potentially allowing a GABA response (Figure S7C) (Hyland and Cryan, 2010). We identified production of multiple hormone receptors in EECs, including the melanocortin receptor MC1R (Figure S6C). MC4R in murine L cells is a regulator of hormone secretion and can be activated by MSH-like producing bacteria (Panaro et al., 2014). ECs selectively expressed the receptor for the thyroid-stimulating hormone, *TSHR* (Figure S6C). Serotonin regulates thyroid hormone levels (Sullo et al., 2011). *TSHR* expression by ECs suggests that this regulation could work bidirectionally. ECs also expressed the receptor for the L cell hormone PYY, *NPY1R* (Figure S6C), reported in murine enterocytes as a regulator of electrolyte transport (Goldspink et al., 2018). We did not confirm expression of *NPY1R* in the CHGA-mNeon^{-} population, which includes enterocytes. L cells highly expressed the Sct receptor *SCTR* as observed in our single-cell atlas but not in mice (Figures S6C and S6D). Fluorescent *in situ* hybridization (FISH) confirmed the expression of *SCTR* in EECs *in vivo* by overlap with *CHGA* (Figure S6E). Because we observed the highest expression of *SCTR* in L cells, we measured GLP-1 secretion upon a 24 h secretin treatment in organoids. Indeed, Sct induced GLP-1 secretion at levels comparable to forskolin as measured by ELISA, or as seen by the loss of intracellular fluorescence of GCG-neon (Figures S6F–S6H).

We next isolated intracellular proteins for mass spectrometry to establish subtype-specific proteomes (Figure S6A). PCA-analysis revealed a clear separation of reporter populations (Figure S7A). The analysis confirmed many of the novel markers for EEC populations, including the L cell hormone PPY, EEC marker MIDKINE, and the M cell peptidase CPB1 (Figure S7B). A gene identified at RNA level but not in the proteome was *CRYBA2*, a crystallin family member (Figure S7C), and reported as a marker of human endocrine cells in pancreas and colon (Muraro et al., 2016; Parikh et al., 2019). CRYBA2 protein was also absent on human intestinal sections using immunohistochemistry (IHC) (Figure S7D). In the human genome, *CRYBA2* maps adjacent to another EEC marker gene, *FEV* (Haber et al., 2017). *CRYBA2* and *FEV* RNAs were expressed in a virtually identical pattern in our single-cell atlas (Figures S7C and S7E).

The Human EEC Secretome

The human EEC organoid cultures uniquely allow proteomic analysis of hormones secreted basolaterally. We isolated supernatants of forskolin-stimulated proximal and distal SI EEC organoids and control organoids, separated into >10kDa and <10kDa fractions. The latter was directly analyzed using liquid chromatography-mass spectrometry (LC-MS), whereas the former was first trypsinised (Figure 6A). Proteins secreted by EEC organoids showed a large non-overlap with the bulk proteome of the different EEC populations (Figure 6B). These proteins were mainly annotated to extracellular processes such as ''secretion'' when compared to intracellular proteome (Figure 6C). In both the >10kDa and the <10kDa (representing processed hormones) fractions, the expected EEC marker hormones were found among the most abundant proteins and peptides identified (Figure 6D). Control organoids (not enriched for EECs) mostly secreted goblet cell products like mucins and trefoil factors. This provided strong evidence for specific hormone processing and secretion by proximal (e.g., GAST, GHRL, MLN, and CCK) and distal (e.g., PYY, NTS and GCG) intestinal organoids.

Hormones known to undergo proteolytic processing were detected in the processed peptides fraction. By contrast, REG4

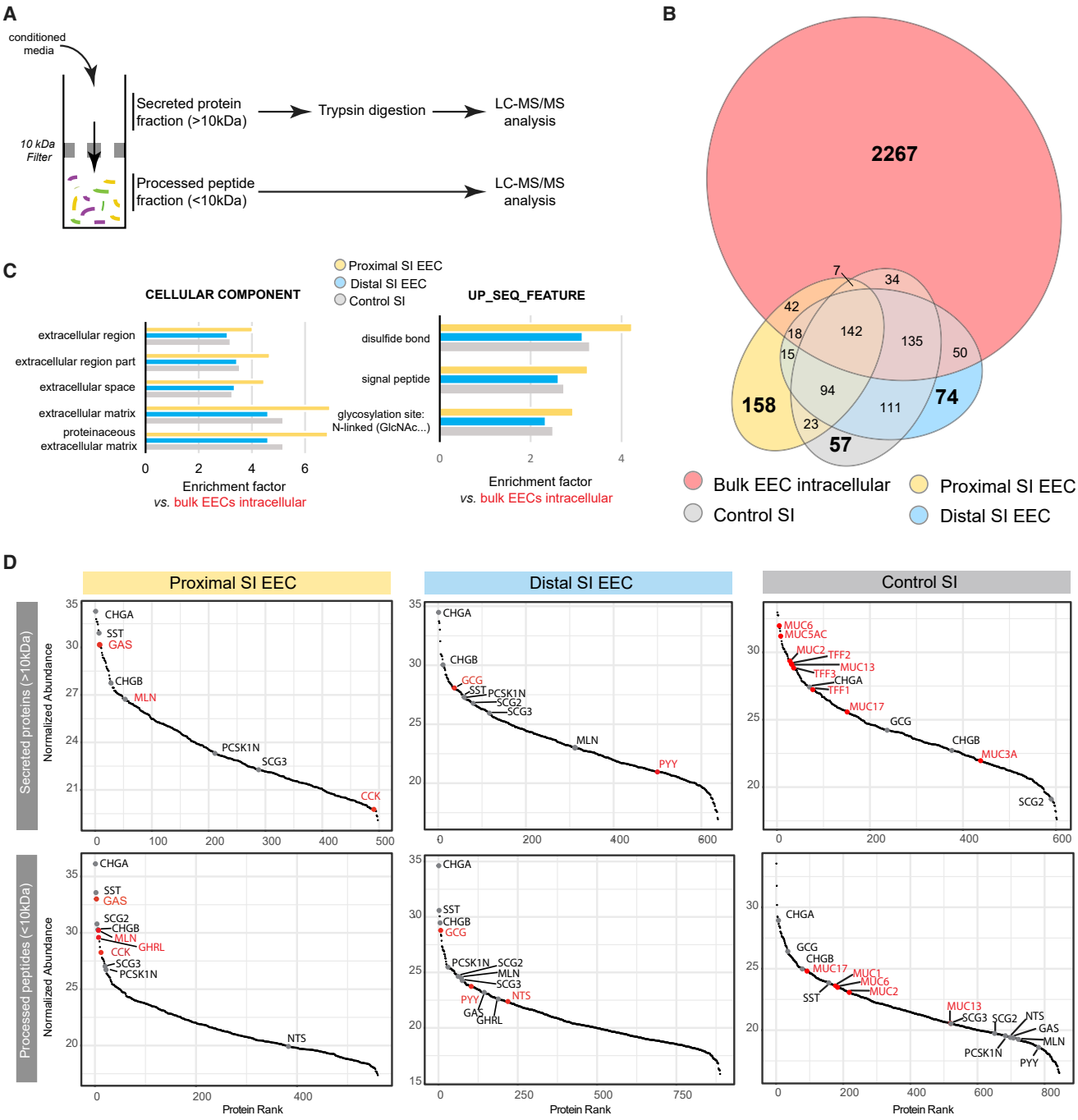


Figure 6. Human EEC Secretome Profiling by Mass Spectometry

(A) Workflow for the analysis of secreted proteins and processed peptides from conditioned media. NEUROG3 induced or control organoids were stimulated for 24 h with forskolin, after which conditioned media were collected. Before applying the 10 kDa cut-off filter, the conditioned media were denatured for efficient partitioning by molecular weight and to disrupt potential protein-protein and protein-peptide interactions.

(B) Venn diagram showing the overlap in protein identifications between the bulk EEC proteome (red) and the proximal EEC-enriched (orange), distal EEC-enriched (blue), and control (gray) small intestinal (SI) organoids secretomes.

(C) Top-enriched cellular component and sequence feature characteristics of proteins identified in the proximal EEC-enriched, distal EEC-enriched, and non-induced SI organoid secretomes. The proteins detected in the bulk EEC intracellular proteome were used as reference for enrichment.

(D) Ranked abundances of proteins detected in the secreted protein fraction and processed peptide fraction. EEC markers are annotated in gray. Products enriched in the different secretomes are annotated in red. Typical proximal or distal hormones are enriched in the respective organoid types, whereas control organoids uniquely secrete mucins and trefoil factors.

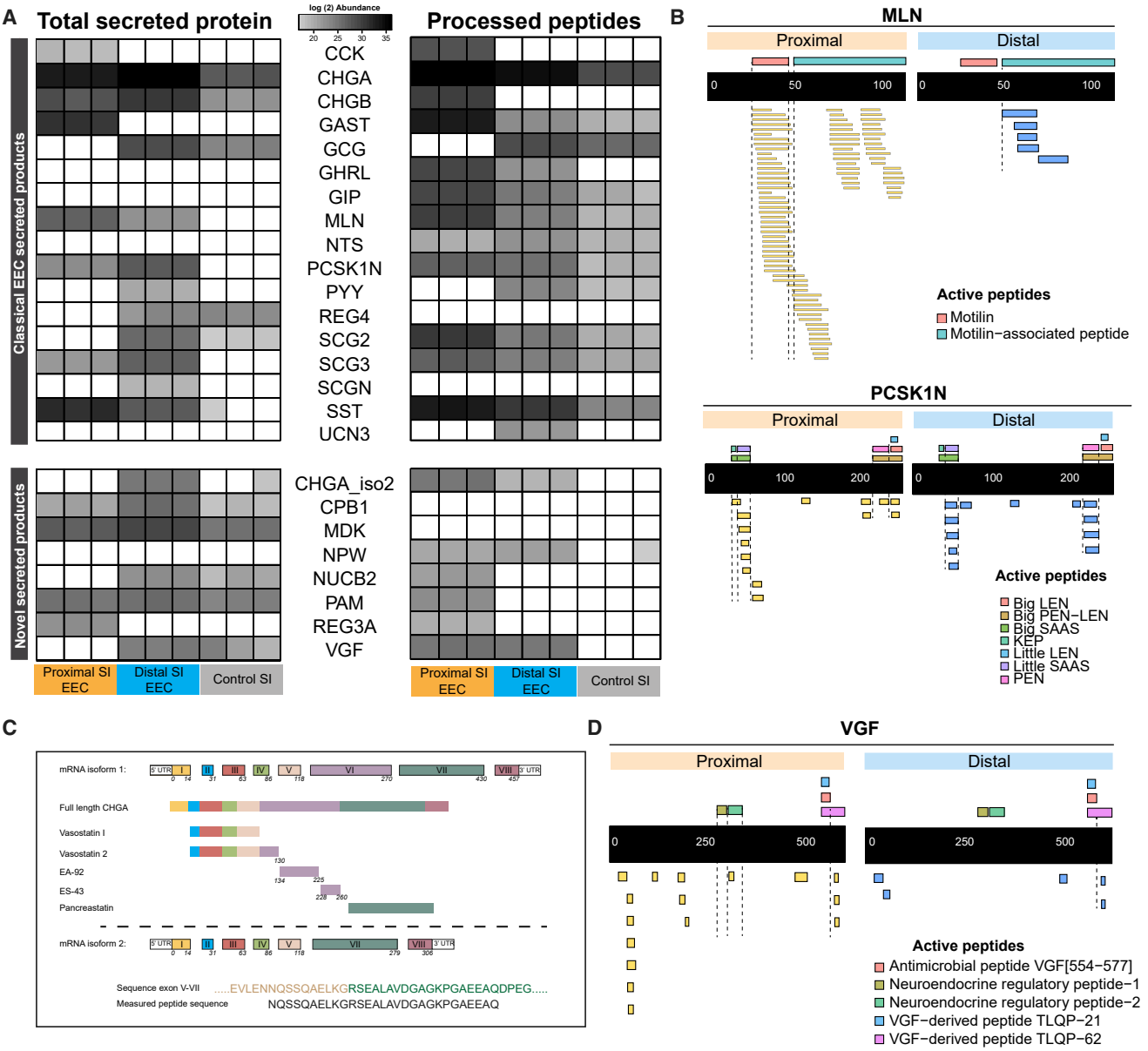


Figure 7. Processing of Region-Specific Human EEC Hormones

(A) Overview of relative abundance of known and novel secreted proteins and processed peptides. MS intensities are plotted as heatmaps for total secreted proteins (>10kDa fraction, left panel) and corresponding processed peptides (<10kDa fraction, right panel).

(B) Measured peptides (<10 kDa) in the supernatant of distal SI organoids mapping to the secreted prohormone MLN and the PCSKN1. Data from proximal (orange) and distal SI organoid (blue) supernatants are displayed. Known bioactive peptides are shown above the black bar.

(C) Alternative processing of chromogranin A (CHGA). Compared to Isoform 1, Isoform 2 of CHGA (lower panel) lacks exon VI. Peptides spanning sequences of both exon V and VII were detected in the supernatant, indicating isoform 2 is produced in human EECs.

(D) Measured peptides (<10 kDa) in the supernatant of proximal and distal SI organoids mapping to the secreted VGF protein. Other than sequences overlapping with known bioactive peptides, numerous other peptides were detected.

(biologically active as a full-length protein) was only found in the >10kDa fraction (Figure 7A). Processed peptides generally displayed C-terminal trimming, likely due to the endogenous activity of carboxypeptidases (Figures 7B and S7F). All known fragments of the proglucagon pro-hormone were observed (Figure S7F). Individual fragments rarely spanned more than a single biologically active peptide. Neuronostatin, a fragment of the pro-somatostatin hormone (Vainio et al., 2012), was found abundantly in the EEC secretome (Figure S7F). Processed peptides were biased toward bioactive fragments of multiple hormones and enzymes, including for MLN, PCSK1N, and GHRL (Figures 7B and S7F). The signal peptide (20–25 N-terminal amino acids of the

prohormones) was consistently cleaved from all hormones (Figures 7B and S7F). Apart from quantitative differences, we also detected region-specific biases in the ratio between bioactive peptides and those with no known activity. For example, duodenal bioactive peptides from proximal-enriched hormones (GHRL, MLN) were highly overrepresented.

Next, we looked for peptides not shown to be secreted from EECs before. CHGA codes for a 457-amino acid preproprotein, cleaved into many different bioactive products. A shorter second isoform lacks exon 6 and has not been shown to be translated (Loh et al., 2012). We now find abundant peptides spanning exon 5 and exon 7 (Figures 7A and 7C). Consistent with the observed RNA expression, other previously unknown EEC products were also observed (Figures 4A and S7). These include NPW, MIDKINE, VGF, and the peptidase CPB1 (Figures 7A and 7D). We found expression and secretion of the antimicrobial peptide REG3A from human EECs (Figures 7A and S7F). We further identified EEC-specific secretion of the enzyme PAM (peptidyl-glycine alpha-amidating monooxygenase), which activates endocrine peptides by C-terminal amidation (Figures 7A and S7F). Coding variants of PAM are associated with type 2 diabetes and can affect insulin secretion (Thomsen et al., 2018). We detected the nucleobindin-2 (NUC2B) precursor, processed to the neuropeptides nesfatin-1, -2, and -3 (Ramesh et al., 2015). Nesfatin-1 has recently gained attention as an anorexigenic and insulinotropic peptide, produced in the hypothalamus and pancreas. Nesfatin-1 has been shown to regulate GLP-1 and GIP secretion *in vitro* (Ramesh et al., 2015).

DISCUSSION

Human EECs are rare and have been largely inaccessible for *in vitro* studies. We have generated a high-resolution transcriptomic and proteomic profile of human EECs from three locations along the gastrointestinal tract, including a first assessment of their secreted products. This dataset yields new hormones, transcription factors, and receptors and can be mined for novel therapeutic targets. The expression atlas highlights key differences with mouse. The transcriptional networks generating the different EEC subtypes have been well worked out in mice (Beucher et al., 2012; Gehart et al., 2019; Gross et al., 2016; Piccand et al., 2019). These networks could result from a stochastically acting system that generates fixed ratios of different EECs. This would explain why organoids generate conserved ratios of EEC subtypes when compared to their tissue of origin (Beumer et al., 2018). A recent study has surveyed a broad human EEC population using antibody-based sorting approaches and bulk RNA sequencing (Roberts et al., 2019) and identified EEC features, such as the expression of neuropeptide W, confirmed by the current study.

We present the first transcriptomic and proteomic profiling of MLN-producing cells. MLN is a regulator of gut motility with intriguing evolutionary dynamics, inactivated independently in lineages leading to the mouse and rat, and guinea pigs (He et al., 2010). The MLN receptor underwent a similar fate (He et al., 2010). This raises questions as to how the cell type (the X cell) that produces MLN diverged from that point. For example, the production of a certain hormone is likely to be accompanied by the expression of dedicated sensory receptors. We found many similarities between mouse X cells and the human counterparts, M/X cells. Transcription factors are conserved (Figure 5A), as is expression of genes required for GHRL modifications such as Acsl1 (Figures 3C and 3D). We noted important differences also, like the expression of putative hormones including *CBLN1* and *AGT*. The latter has been proposed as motility regulator similar to MLN (Figure 4). We also identify a cytokine receptor in M/X cells, *IL-20RA*, which could link a sensory mechanism for pathogens to an expelling motility response.

The expression of receptors for some EEC hormones by EECs has been reported in mouse, particularly for *SST* (e.g., *Sstr5* in L cells) (Chisholm and Greenberg, 2002). We now find that human EECs can sense extracellular PYY (NPY1R) and Sct (SCTR). The PYY-receptor Npy1r has been suggested as an enterocyte marker in mouse, which we do not confirm in human (Goldspink et al., 2018). Rather, we observe exclusive expression in human serotonin-producing ECs. *SCTR* expression is low in ECs and enriched in EECs producing *GCG* and *GAST/GIP*. We show that Sct can stimulate L cells to secrete GLP-1. Importantly, a Sct stimulation test is commonly used in diagnostics of Zollinger-Ellison syndrome patients that suffer from gastrin-producing tumors (Berna et al., 2006). Sct normally represses blood gastrin by inhibiting the secretion of GAST from stomach G cells (the major site of GAST production), likely through modulating the luminal pH. Patients suffering from SI gastrinoma show sharp increases in serum GAST upon Sct administration. Our data suggest this to occur through SCTR expression by SI GAST-producing G cells. More broadly, our data indicate that human EECs have an extensive capacity to cross-communicate through their hormone products.

Taken together, the EEC atlas and EEC-TAG biobank represent rich resources to identify regulators of human EEC development and function.

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
- METHOD DETAILS
  - Cell culture of human intestinal organoids
  - Constructs for EEC-TAG reporter and knockout generation
  - Calcium sensor
  - Live cell imaging of calcium reporter organoids
  - Transmission electron microscopy
  - Immunostaining
  - Fluorescent *in situ* hybridization
  - ELISA
  - RNA isolation and quantitative PCR

- Processing human intestinal tissue for single cell RNA sequencing
- Single cell sorting for RNA sequencing from organoids
- Single cell RNA sequencing analysis from organoids
- Bulk RNA sequencing analysis
- Preparation of secreted peptides and proteins for LC-MS
- Preparation of FACS-sorted EECs for proteome analyses
- LC-MS
- Proteomics data analysis

● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2020.04.036.

ACKNOWLEDGMENTS

We thank Anko de Graaff and the Hubrecht Imaging Centre (HIC) for microscopy assistance, Single Cell Discoveries for single-cell sequencing, and Folkert Morsink and Johan Offerhaus of the University Medical Center Utrecht for providing sections of human intestinal biopsies. J.B.-M., A.J.R.H., and W.W. acknowledge financial support from the Horizon 2020 program INFRAIA project Epic-XS (Project 823839) and the NWO funded Netherlands Proteomics Centre through the National Road Map for Large-scale Infrastructures program X-Omics (Project 184.034.019). This work was supported by NETRF/Petersen Accelerator (J.B.), CRUK grant OPTIMISTICC (C10674/A27140) (J.P. and C.P.-M.), Netherlands Organ-on-Chip Initiative (024.003.001) from the Netherlands Organisation for Scientific Research (NWO) (J.P., A.M.-S., and C.P.-M.), and ERC Advanced Grant Agreement no. 67013e (H.C.). We thank the Microscopy CORE Lab at M4I Maastricht University for their support in electron microscopy.

AUTHOR CONTRIBUTIONS

J.B., J.P., and H.C. conceptualized the project, designed the experiments, interpreted the results, and wrote the manuscript. J.B.-M., W.W., and A.J.R.H. performed the proteomic and secretomic experiments and analysis. A.M.S. assisted in cell culture experiments supervised by J.B., J.P., and H.C.; J.B. and J.P generated and analyzed the organoid-derived EEC single-cell atlas. R.E., K.R.J., A.R., M.Z., and S.A.T. generated and analyzed the primary human EEC single-cell dataset. G.A.B. generated and provided the neurogenin-3 overexpression vector. D.H. and B.A. generated and provided the targeting vectors for CRISPR-HOT-mediated reporter organoids. A.A.-R. and M.H.G provided different constructs. A.S., C.B., Y.P., and C.P.-M. contributed to functional EEC assays. F.v.d.L. and J.G. provided the Tq-Ca-FLITS reporter. B.P. and H.S. performed imaging and quantification of calcium responses. Y.E.B.-E. and R.v.d.L assisted with FACS experiments. K.K. provided organoid lines. C.L.-I., W.J.v.d.W., and P.J.P. performed transmission electron microscopy.

DECLARATION OF INTERESTS

H.C. is inventor on several patents related to organoid technology; his full disclosure is given at https://www.uu.nl/staff/JCClevers/. H.C. is founder of OrganoidZ, which employs organoids for drug development. J.B., J.P., and H.C. are inventors on patents related to this work.

Received: November 29, 2019
Revised: March 10, 2020
Accepted: April 21, 2020
Published: May 13, 2020

REFERENCES

Artegiani, B., Hendriks, D., Beumer, J., Kok, R., Zheng, X., Joore, I., Chuva de Sousa Lopes, S., van Zon, J., Tans, S., and Clevers, H. (2020). Fast and efficient generation of knock-in human organoids using homology-independent CRISPR-Cas9 precision genome editing. Nat. Cell Biol. *22*, 321–331.

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

Berna, M.J., Hoffmann, K.M., Long, S.H., Serrano, J., Gibril, F., and Jensen, R.T. (2006). Serum gastrin in Zollinger-Ellison syndrome: II. Prospective study of gastrin provocative testing in 293 patients from the National Institutes of Health and comparison with 537 cases from the literature. evaluation of diagnostic criteria, proposal of new criteria, and correlations with clinical and tumoral features. Medicine (Baltimore) *85*, 331–364.

Beucher, A., Gjernes, E., Collin, C., Courtney, M., Meunier, A., Collombat, P., and Gradwohl, G. (2012). The homeodomain-containing transcription factors Arx and Pax4 control enteroendocrine subtype specification in mice. PLoS ONE *7*, e36449.

Beumer, J., Artegiani, B., Post, Y., Reimann, F., Gribble, F., Nguyen, T.N., Zeng, H., Van den Born, M., Van Es, J.H., and Clevers, H. (2018). Enteroendocrine cells switch hormone expression along the crypt-to-villus BMP signalling gradient. Nat. Cell Biol. *20*, 909–916.

Borges, M., Linnoila, R.I., van de Velde, H.J.K., Chen, H., Nelkin, B.D., Mabry, M., Baylin, S.B., and Ball, D.W. (1997). An achaete-scute homologue essential for neuroendocrine differentiation in the lung. Nature *386*, 852–855.

Boztepe, T., and Gulec, S. (2018). Investigation of the influence of high glucose on molecular and genetic responses: an *in vitro* study using a human intestine model. Genes Nutr. *13*, 11.

Bukhari, H., and Müller, T. (2019). Endogenous Fluorescence Tagging by CRISPR. Trends Cell Biol. *29*, 912–928.

Caruso, V., Sreedharan, S., Carlini, V.P., Jacobsson, J.A., Haitina, T., Hammer, J., Stephansson, O., Crona, F., Sommer, W.H., Risérus, U., et al. (2016). mRNA GPR162 changes are associated with decreased food intake in rat, and its human genetic variants with impairments in glucose homeostasis in two Swedish cohorts. Gene *581*, 139–145.

Chang-Graham, A.L., Danhof, H.A., Engevik, M.A., Tomaro-Duchesneau, C., Karandikar, U.C., Estes, M.K., Versalovic, J., Britton, R.A., and Hyser, J.M. (2019). Human Intestinal Enteroids With Inducible Neurogenin-3 Expression as a Novel Model of Gut Hormone Secretion. Cell. Mol. Gastroenterol. Hepatol. *8*, 209–229.

Chisholm, C., and Greenberg, G.R. (2002). Somatostatin-28 regulates GLP-1 secretion via somatostatin receptor subtype 5 in rat intestinal cultures. Am. J. Physiol. Endocrinol. Metab. *283*, E311–E317.

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.

Cox, H.M. (2007). Neuropeptide Y receptors; antisecretory control of intestinal epithelial function. Auton. Neurosci. *133*, 76–85.

Cui, T., Tsolakis, A.V., Li, S.C., Cunningham, J.L., Lind, T., Öberg, K., and Giandomenico, V. (2013). Olfactory receptor 51E1 protein as a potential novel tissue biomarker for small intestine neuroendocrine carcinomas. Eur. J. Endocrinol. *168*, 253–261.

Di Paola, R., Caporarello, N., Marucci, A., Dimatteo, C., Iadicicco, C., Del Guerra, S., Prudente, S., Sudano, D., Miele, C., Parrino, C., et al. (2011). ENPP1 affects insulin action and secretion: evidences from in vitro studies. PLoS ONE *6*, e19462.

Edfeldt, K., Daskalakis, K., Bäcklin, C., Norlén, O., Tiensuu Janson, E., Westin, G., Hellman, P., and Stålberg, P. (2017). DcR3, TFF3, and Midkine Are Novel Serum Biomarkers in Small Intestinal Neuroendocrine Tumors. Neuroendocrinology *105*, 170–181.

Engelstoft, M.S., Egerod, K.L., Lund, M.L., and Schwartz, T.W. (2013a). Enteroendocrine cell types revisited. Curr. Opin. Pharmacol. *13*, 912–921.

Engelstoft, M.S., Park, W.M., Sakata, I., Kristensen, L.V., Husted, A.S., Osborne-Lawrence, S., Piper, P.K., Walker, A.K., Pedersen, M.H., Nøhr, M.K., et al. (2013b). Seven transmembrane G protein-coupled receptor repertoire of gastric ghrelin cells. Mol. Metab. *2*, 376–392.

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.

Fan, N., Sun, H., Wang, Y., Zhang, L., Xia, Z., Peng, L., Hou, Y., Shen, W., Liu, R., and Peng, Y. (2014). Midkine, a potential link between obesity and insulin resistance. PLoS ONE *9*, e88299.

Fleischer, J., Bumbalo, R., Bautze, V., Strotmann, J., and Breer, H. (2015). Expression of odorant receptor Olfr78 in enteroendocrine cells of the colon. Cell Tissue Res. *361*, 697–710.

Foster, S.R., Hauser, A.S., Vedel, L., Strachan, R.T., Huang, X.P., Gavin, A.C., Shah, S.D., Nayak, A.P., Haugaard-Kedström, L.M., Penn, R.B., et al. (2019). Discovery of Human Signaling Systems: Pairing Peptides to G Protein-Coupled Receptors. Cell *179*, 895–908.e21.

Fujii, M., Matano, M., Nanki, K., and Sato, T. (2015). Efficient genetic engineering of human intestinal organoids using electroporation. Nat. Protoc. *10*, 1474–1485.

Furness, J.B., Rivera, L.R., Cho, H.-J., Bravo, D.M., and Callaghan, B. (2013). The gut as a sensory organ. Nat. Rev. Gastroenterol. Hepatol. *10*, 729–740.

Gardiner, J.V., Beale, K.E., Roy, D., Boughton, C.K., Bataveljic, A., Campbell, D.C., Bewick, G.A., Patel, N.A., Patterson, M., Leavy, E.M., et al. (2010). Cerebellin1 is a novel orexigenic peptide. Diabetes Obes. Metab. *12*, 883–890.

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.

Goldspink, D.A., Reimann, F., and Gribble, F.M. (2018). Models and Tools for Studying Enteroendocrine Cells. Endocrinology *159*, 3874–3884.

Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A., Leblanc, G., Hatten, M.E., and Heintz, N. (2003). A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. Nature *425*, 917–925.

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

Gross, S., Garofalo, D.C., Balderes, D.A., Mastracci, T.L., Dias, J.M., Perlmann, T., Ericson, J., and Sussel, L. (2016). The novel enterochromaffin marker Lmx1a regulates serotonin biosynthesis in enteroendocrine cell lineages downstream of Nkx2.2. Development *143*, 2616–2628.

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.

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.

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.

He, J., Irwin, D.M., Chen, R., and Zhang, Y.P. (2010). Stepwise loss of motilin and its specific receptor genes in rodents. J. Mol. Endocrinol. *44*, 37–44.

He, X., Tan, C., Wang, F., Wang, Y., Zhou, R., Cui, D., You, W., Zhao, H., Ren, J., and Feng, B. (2016). Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. Nucleic Acids Res. *44*, e85.

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.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Hyland, N.P., and Cryan, J.F. (2010). A gut feeling about GABA: Focus on GABAB receptors. Front. Pharmacol. https://doi.org/10.3389/fphar.2010.00124.

Jovancevic, N., Khalfaoui, S., Weinrich, M., Weidinger, D., Simon, A., Kalbe, B., Kernt, M., Kampik, A., Gisselmann, G., Gelis, L., and Hatt, H. (2017). Odorant receptor 51E2 agonist β-ionone regulates RPE cell migration and proliferation. Front. Physiol. *8*, 888.

Koo, B.K., Stange, D.E., Sato, T., Karthaus, W., Farin, H.F., Huch, M., van Es, J.H., and Clevers, H. (2011). Controlled gene expression in primary Lgr5 organoid cultures. Nat. Methods *9*, 81–83.

Kuro-O, M. (2019). The Klotho proteins in health and disease. Nat. Rev. Nephrol. *15*, 27–44.

Levine, A.S., Winsky-Sommerer, R., Huitron-Resendiz, S., Grace, M.K., and de Lecea, L. (2005). Injection of neuropeptide W into paraventricular nucleus of hypothalamus increases food intake. Am. J. Physiol. Regul. Integr. Comp. Physiol. *288*, R1727–R1732.

Li, W., Ji, M., Lin, Y., Miao, Y., Chen, S., and Li, H. (2018). DEPP/DEPP1/C10ORF10 regulates hepatic glucose and fat metabolism partly via ROS-induced FGF21. FASEB J. *32*, 5459–5469.

Lo, C.A., Kays, I., Emran, F., Lin, T.J., Cvetkovska, V., and Chen, B.E. (2015). Quantification of Protein Levels in Single Living Cells. Cell Rep. *13*, 2634–2644.

Loh, Y.P., Cheng, Y., Mahata, S.K., Corti, A., and Tota, B. (2012). Chromogranin A and derived peptides in health and disease. J. Mol. Neurosci. *48*, 347–356.

Lohoff, F.W., Dahl, J.P., Ferraro, T.N., Arnold, S.E., Gallinat, J., Sander, T., and Berrettini, W.H. (2006). Variations in the vesicular monoamine transporter 1 gene (VMAT1/SLC18A1) are associated with bipolar i disorder. Neuropsychopharmacology *31*, 2739–2747.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

McCracken, K.W., Catá, E.M., Crawford, C.M., Sinagoga, K.L., Schumacher, M., Rockich, B.E., Tsai, Y.H., Mayhew, C.N., Spence, J.R., Zavros, Y., and Wells, J.M. (2014). Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. Nature *516*, 400–404.

McInnes, L., Healy, J., Saul, N., and Großberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. J. Open Source Softw. *3*, 861.

Muñoz, J., Stange, D.E., Schepers, A.G., van de Wetering, M., Koo, B.-K., Itzkovitz, S., Volckmann, R., Kung, K.S., Koster, J., Radulescu, S., et al. (2012). The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers. EMBO J. *31*, 3079–3091.

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.

Nguyen, T.L.A., Vieira-Silva, S., Liston, A., and Raes, J. (2015). How informative is the mouse for human gut microbiota research? Dis. Model. Mech. *8*, 1–16.

Pablo, J.L., and Pitta, G.S. (2017). FGF14 is a regulator of KCNQ2/3 channels. Proc. Natl. Acad. Sci. U S A. https://doi.org/10.1073/pnas.1610158114.

Pan, F.C., Brissova, M., Powers, A.C., Pfaff, S., and Wright, C.V.E. (2015). Inactivating the permanent neonatal diabetes gene Mnx1 switches insulin-producing β-cells to a γ-like fate and reveals a facultative proliferative capacity in aged β-cells. Dev. *142*, 3637–3648.

Panaro, B.L., Tough, I.R., Engelstoft, M.S., Matthews, R.T., Digby, G.J., Møller, C.L., Svendsen, B., Gribble, F., Reimann, F., Holst, J.J., et al. (2014). The melanocortin-4 receptor is expressed in enteroendocrine L cells and regulates the release of peptide YY and glucagon-like peptide 1 in vivo. Cell Metab. *20*, 1018–1029.

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.

Parker, H.E., Habib, A.M., Rogers, G.J., Gribble, F.M., and Reimann, F. (2009). Nutrient-dependent secretion of glucose-dependent insulinotropic polypeptide from primary murine K cells. Diabetologia *52*, 289–298.

Piccand, J., Vagne, C., Blot, F., Meunier, A., Beucher, A., Strasser, P., Lund, M.L., Ghimire, S., Nivlet, L., Lapp, C., et al. (2019). Rfx6 promotes the differentiation of peptide-secreting enteroendocrine cells while repressing genetic programs controlling serotonin production. Mol. Metab. *29*, 24–39.

Pietraszewska-Bogiel, A., van Weeren, L., and Goedhart, J. (2019). Seeing cells smell: Dynamic optical measurements of Ca^{2+} and cAMP signaling from Olfactory Receptors transiently expressed in HEK293TN cells. bioRxiv. https://doi.org/10.1101/771261.

Ramesh, N., Mortazavi, S., and Unniappan, S. (2015). Nesfatin-1 stimulates glucagon-like peptide-1 and glucose-dependent insulinotropic polypeptide secretion from STC-1 cells in vitro. Biochem. Biophys. Res. Commun. *462*, 124–130.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.

Reimann, F., Habib, A.M., Tolhurst, G., Parker, H.E., Rogers, G.J., and Gribble, F.M. (2008). Glucose Sensing in L Cells: A Primary Cell Study. Cell Metab *8*, 532–539.

Roberts, G.P., Larraufie, P., Richards, P., Kay, R.G., Galvin, S.G., Miedzybrodzka, E.L., Leiter, A., Li, H.J., Glass, L.L., Ma, M.K.L., et al. (2019). Comparison of human and murine enteroendocrine cells by transcriptomic and peptidomic profiling. Diabetes *68*, 1062–1072.

Sachs, N., Papaspyropoulos, A., Zomer-van Ommen, D.D., Heo, I., Böttinger, L., Klay, D., Weeber, F., Huelsz-Prince, G., Iakobachvili, N., Amatngalim, G.D., et al. (2019). Long-term expanding human airway organoids for disease modeling. EMBO J. *38*, e100300.

Sanger, G.J. (2004). Neurokinin NK1 and NK3 receptors as targets for drugs to treat gastrointestinal motility disorders and pain. Br. J. Pharmacol. *141*, 1303–1312.

Sapio, M.R., and Fricker, L.D. (2014). Carboxypeptidases in disease: insights from peptidomic studies. Proteomics Clin. Appl. *8*, 327–337.

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. (2011). Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. Gastroenterology *141*, 1762–1772.

Schmid-Burgk, J.L., Höning, K., Ebert, T.S., and Hornung, V. (2016). CRISPaint allows modular base-specific gene tagging using a ligase-4-dependent mechanism. Nat. Commun. *7*, 12338.

Schneeberger, M. (2019). Irx3, a new leader on obesity genetics. EBioMedicine *39*, 19–20.

Scott, L.J., Mohlke, K.L., Bonnycastle, L.L., Willer, C.J., Li, Y., Duren, W.L., Erdos, M.R., Stringham, H.M., Chines, P.S., Jackson, A.U., et al. (2007). A genome-wide association study of type 2 diabetes in finns detects multiple susceptibility variants. Science *316*, 1341–1345.

Sharma, D., Verma, S., Vaidya, S., Kalia, K., and Tiwari, V. (2018). Recent updates on GLP-1 agonists: Current advancements & challenges. Biomed. Pharmacother. *108*, 952–962.

Shcherbina, L., Lindqvist, A., Thorén Fischer, A.H., Ahlqvist, E., Zhang, E., Falkmer, S.E., Renström, E., Koffert, J., Honka, H., and Wierup, N. (2018). Intestinal CART is a regulator of GIP and GLP-1 secretion and expression. Mol. Cell. Endocrinol. *476*, 8–16.

Sinagoga, K.L., McCauley, H.A., Múnera, J.O., Reynolds, N.A., Enriquez, J.R., Watson, C., Yang, H.C., Helmrath, M.A., and Wells, J.M. (2018). Deriving functional human enteroendocrine cells from pluripotent stem cells. Development *145*, dev165795.

Sommer, C.A., and Mostoslavsky, G. (2014). RNA-Seq analysis of enteroendocrine cells reveals a role for FABP5 in the control of GIP secretion. Mol. Endocrinol. *28*, 1855–1865.

Stahl, R., Walcher, T., De Juan Romero, C., Pilz, G.A., Cappello, S., Irmler, M., Sanz-Aquela, J.M., Beckers, J., Blum, R., Borrell, V., and Götz, M. (2013). Trnp1 regulates expansion and folding of the mammalian cerebral cortex by control of radial glial fate. Cell *153*, 535–549.

Sullo, A., Brizzi, G., and Maffulli, N. (2011). Chronic peripheral administration of serotonin inhibits thyroid function in the rat. Muscles Ligaments Tendons J. *1*, 48–50.

The, M., MacCoss, M.J., Noble, W.S., and Käll, L. (2016). Fast and Accurate Protein False Discovery Rates on Large-Scale Proteomics Data Sets with Percolator 3.0. J. Am. Soc. Mass Spectrom. *27*, 1719–1727.

Thomsen, S.K., Raimondo, A., Hastoy, B., Sengupta, S., Dai, X.Q., Bautista, A., Censin, J., Payne, A.J., Umapathysivam, M.M., Spigelman, A.F., et al. (2018). Type 2 diabetes risk alleles in PAM impact insulin release from human pancreatic β-cells. Nat. Genet. *50*, 1122–1131.

Too, L.K., Li, K.M., Suarna, C., Maghzal, G.J., Stocker, R., McGregor, I.S., and Hunt, N.H. (2016). Deletion of TDO2, IDO-1 and IDO-2 differentially affects mouse behavior and cognitive function. Behav. Brain Res. *312*, 102–117.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740.

Vainio, L., Perjes, A., Ryti, N., Magga, J., Alakoski, T., Serpi, R., Kaikkonen, L., Piuhola, J., Szokodi, I., Ruskoaho, H., and Kerkelä, R. (2012). Neuronostatin, a novel peptide encoded by somatostatin gene, regulates cardiac contractile function and cardiomyocyte survival. J. Biol. Chem. *287*, 4572–4580.

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.

Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15.

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.

Yu, S.L., Han, S., Kim, H.R., Park, J.W., Jin, D.I., and Kang, J. (2017). Phosphorylation of carboxypeptidase B1 protein regulates β-cell proliferation. Int. J. Mol. Med. *40*, 1397–1404.

Zhang, J., McKenna, L.B., Bogue, C.W., and Kaestner, K.H. (2014). The diabetes gene Hhex maintains δ-cell differentiation and islet function. Genes Dev. *28*, 829–834.

Zhang, X., McGrath, P.S., Salomone, J., Rahal, M., McCauley, H.A., Schweitzer, J., Kovall, R., Gebelein, B., and Wells, J.M. (2019). A Comprehensive Structure-Function Study of Neurogenin3 Disease-Causing Alleles during Human Pancreas and Intestinal Organoid Development. Dev. Cell *50*, 367–380.e7.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Chromogranin A	Santa Cruz	sc-1488, RRID: AB_2276319
Anti-Cholestocystokinin	Santa Cruz	sc-21617, RRID: AB_2072464
Anti-Neurotensin	Santa Cruz	sc-20806, RRID: AB_2155562
Anti-Somatostatin	Santa Cruz	sc-7819, RRID: AB_2302603
Anti-Serotonin	Abcam	ab66047, RRID: AB_1142794
Anti-Gastric inhibitory polypeptide	Abcam	ab22624-50, RRID: AB_2109683
Anti-GLP1	Santa Cruz	sc-7782, RRID: AB_2107325
Anti-Motilin	Atlas antibodies	HPA069392, RRID: AB_2686136
Anti-GLP1	Abcam	ab22625, RRID: AB_447206
Anti-Gastrin	Proteintech	60346-1-Ig
Anti-Ghrelin	Santa Cruz	sc-10368, RRID: AB_2232479
Anti-beta-catenin	BD transduction laboratories	#610154, RRID: AB_397555
Anti-Neuropeptide W	Novus biologicals	NBP2-57337
Anti-Precerebellin	Sigma-Aldrich	ABN304
Anti-PPY	Atlas antibodies	HPA032122, RRID: AB_2674164
Alexa Fluor 488 donkey anti-rabbit	Thermo Fisher scientific	A21206, RRID: AB_2535792
Alexa Fluor 488 donkey anti-goat	Thermo Fisher scientific	A11055, RRID: AB_2534102
Alexa Fluor 568 donkey anti-rabbit	Thermo Fisher scientific	A10042, RRID: AB_2534017
Alexa Fluor 568 donkey anti-goat	Thermo Fisher scientific	A11057, RRID: AB_2534104
Alexa Fluor 647 donkey anti-rabbit	Thermo Fisher scientific	A31573, RRID: AB_2536183
Alexa Fluor 647 donkey anti-goat	Thermo Fisher scientific	A32849, RRID: AB_2762840
Alexa Fluor 647 donkey anti-mouse	Thermo Fisher scientific	A31571, RRID: AB_162542
Envision+ System-HRP polymer anti-rabbit	DAKO	K4002
Biological Samples		
Human intestinal tissue for organoids	Utrecht Medical Center	N/A
Human intestinal biopsies for RNA sequencing	Addenbrooke's Hospital, Cambridge	Ethics: REC 17/EE/0265
Chemicals, Peptides, and Recombinant Proteins		
10 kDa Vivaspin centrifugal device	Sartorius, Gottingen, Germany	Catalogue # VS0101
Reversed-phase C18 1cc columns	Waters Corporation, Milford, USA	Catalogue # WAT054925
Trypsin enzyme	Promega, Madison, USA	Catalogue # T1426
Lysyl endopeptidase enzyme (Lys C)	Wako Chemicals GmbH	Catalogue # 129-02541
DNase I	Sigma-Aldrich, Missouri, USA	Catalogue #DN25
RNase A	Sigma-Aldrich, Missouri, USA	Catalogue # R-6513
Advanced DMEM/F12	Thermo Fisher scientific	12634-010
B-27 Supplement	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
R-spondin conditioned medium	U-Protein Express	Custom order

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
N-Acetyl-L-cysteine	Sigma-Aldrich	A9165
Nicotinamide	Sigma-Aldrich	N0636
Human EGF	Peprotech	AF-100-15
A83-01	Tocris	2939
Prostaglandin E2	Tocris	2296
Forskolin	Tocris	1099
A83-01	Tocris	2939
SB 202190	Sigma-Aldrich	S7076
Y-27632 dihydrochloride	Abmole	M1817
Primocin	Invivogen	ant-pm-2
BMP-2	Peprotech	120-02C
BMP-4	Peprotech	120-05ET
Secretin	Tocris	1918
Cultrex Basement Membrane Extract (BME), Growth Factor Reduced, Type 2	R&D Systems, Bio-Techne	3533-001-02
DAPI	Thermo Fisher scientific	D1306
Formaldehyde solution 4%	Sigma-Aldrich	1.00496
SYBR Green	Bio Rad	1725270
Donkey serum	Golden Bridge International	E27-100
Triton X-100	Sigma-Aldrich	X100-100ML
SORT-seq reagents	(Muraro et al., 2016)	N/A
Beta-ionone	Sigma-Aldrich	I12603
DAPT	Sigma-Aldrich	D5942
IWP-2	Stemcell Techonologies	72122
SapI	New England Biolabs	R0569S
NotI	New England Biolabs	R0189S
Phusion High fidelity DNA polymerase	New England Biolabs	M0530S
TryplE	Thermo Fisher scientific	12605010
Vectashield	Vector Labs	H-1000-10
Hyaluronidase	Merck	#385931-25KU
BTXpress solution	BTX	45-0805
Critical Commercial Assays		
RNeasy Mini Kit	QIAGEN	74104
GLP-1 ELISA kit	Sigma-Aldrich	RAB0201
In-fusion cloning kit	Takara	638910
QIAquick PCR Purification Kit	QIAGEN	28104
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
Miniprep DNA isolation kit	Thermo Fisher scientific	K210003
Midiprep DNA isolation kit	Thermo Fisher scientific	K210005
Chromium Single Cell 3′ Library & Gel Bead Kit v2, 16 rxns	10x Genomics	PN-120237
Deposited Data		
Raw mass spectrometry data	PRIDE repository https://www.ebi.ac.uk/pride/	Accession number PXD017468
Raw and analyzed sequencing	Gene expression omnibus https://www.ncbi.nlm.nih.gov/geo/	GSE146799

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Proteome Discoverer 2.3	Thermo Fisher scientific	OPTON-30956 (https://www.thermofisher.com/us/en/home/technical-resources/request-a-quote.OPTON-30956.html?supportType=SL)
Sequest HT	Thermo Fisher scientific	With Proteome Discoverer 2.3 (OPTON-30956)
Percolator	Thermo Fisher scientific	With Proteome Discoverer 2.3 (OPTON-30956)
Perseus 1.6.2.2	Max Planck Institute of Biochemistry	https://maxquant.net/perseus/
Database for Annotation, Visualization and Integrated Discovery (DAVID) version 6.8	(Huang et al., 2009)	https://david.ncifcrf.gov/
Uniprot human database (Organism Species 9606)	Uniprot	https://www.uniprot.org/uniprot/?query=*&fil=organism%3A%22Homo+sapiens+%28Human%29+%5B9606%5D%22+AND+reviewed%3Ayes
CFX manager software	Bio-Rad	N/A
RaceID3	(Herman et al., 2018)	https://github.com/dgrun/RaceID3_StemID2
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
Cellranger (Version 2.1.0, reference transcriptome GRCh38-1.2.0)	10x Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/installation
Scanpy (Version 1.4)	N/A	https://icb-scanpy.readthedocs-hosted.com/en/stable/
Pandas (Version 0.25.2)	N/A	https://pandas.pydata.org/
NumPy (Version 1.16.2)	N/A	https://numpy.org/
UMAP	(McInnes et al., 2018)	Python package umap
Scipy (Version 1.2.1)	N/A	https://www.scipy.org/
Anndata (Version 0.6.19)	N/A	https://pypi.org/project/anndata/
Other		
EVOS Cell Imaging System	Thermo Fisher scientific	N/A
EVOS FL Auto 2 Cell Imaging System	Thermo Fisher scientific	N/A
SP8 confocal microscope	Leica	N/A
DM4000	Leica	N/A
NEPA21 electroporator	Nepagene	N/A
FACSAria	BD Bioschiences	N/A
FACS BD Influx	BD Bioschiences	N/A

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to the Lead Contact, Hans Clevers (h.clevers@hubrecht.eu).

Materials Availability

Unique/stable reagents generated in this study are available and can be requested from the Lead Contact, a completed Materials Transfer Agreement may be required.

Data and Code Availability

All bulk and single cell RNA sequencing data of this study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE146799.

The raw MS data is deposited in PRIDE, with accession number PXD017468.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Tissues from the human duodenum, ileum and colon were obtained from the UMC Utrecht with informed consent of each patient. All patients were males that were diagnosed with small intestinal or colon adenocarcinoma that was resected. A sample from non-transformed, normal mucosa was taken for this study. The study was approved by the UMC Utrecht (Utrecht, the Netherlands) ethical committee and was in accordance with the Declaration of Helsinki and according to Dutch law. This study is compliant with all relevant ethical regulations regarding research involving human participants.

For immunostainings, sections of formalin-fixed, paraffin embedded human intestinal 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).

METHOD DETAILS

Cell culture of human intestinal organoids

Human small intestinal cells were isolated, processed and cultured as described previously (Beumer et al., 2018; Sato et al., 2011). Instead of Wnt conditioned media, the medium was supplemented with Wnt surrogate (0,15 nM, U-Protein Express). Ileal organoids were splitted on average every 10 days, duodenal and colon organoids ever 7 days. For passaging, organoids were removed from the BME using ice-cold AdDMEM/F12 (GIBCO) and mechanically dissociated into small fragments using a Pasteur pipette. Fragments were replated in fresh BME.

For differentiation toward EECs, organoids were treated with 1 μg/mL doxycycline (Sigma) in 'ENR' medium (Sato et al., 2009). Secretin (Tocris) was used at a concentration of 10 μg/mL. Beta-ionone (Sigma) was used at 100 mg/mL. BMP activation was achieved by withdrawing Noggin from 'ENR' and addition of BMP-2 (Peprotech, 50 ng/mL) and BMP-4 (Peprotech, 50 ng/mL). Notch signaling was inhibited by treatment with the Gamma-secretase inhibitor DAPT (Sigma, 10 μM). Wnt inhibition was performed by treatment with the Porcupine inhibitor IWP-2 (StemCell Technologies, Inc., 5 μM). FGF21 was used at a concentration of 1 μg/mL (Peprotech).

Constructs for EEC-TAG reporter and knockout generation

The NEUROG3 was cloned in a two insert Gibson reaction into BSKS II vector. Of note, two PCR reactions were done: first, NEUROG3 was amplified from human genomic DNA, since the entire coding region lies in one exon. Second, the BSKS vector was amplified. The forward and reverse primers for Gly linker, FLAG, HA and P2A sequence were annealing to each other (Table S1). All three DNA fragments were then combined in BSKS-NEUROG3-Flag-HA-P2A. In the next step, NEUROG3-P2A sequence was excised using EcoRI enzyme and cloned into previously published pLX-NS2 vector (Sachs et al., 2019). Organoids were lentivirally transduced as described before (Koo et al., 2011). Lentiviral transduction was performed on small clumps of cells (2-10 cells) rather than single cells, achieved after TryplE (TrypLE Express; Life Technologies) dissociation.

For generation of the reporter organoid lines using CRISPR-HOT, we utilized a method described in (Artegiani et al., 2020). Briefly, we used a targeting plasmid containing a fluorescent protein (mNEON or tdTomato) which can be linearized at a defined base position by a specific sgRNA and Cas9 provided from a second plasmid, which also encodes mCherry (Schmid-Burgk et al., 2016). These two plasmids are co-electroporated with a plasmid encoding the sgRNA for the respective locus (Table S1).

The HDR donor plasmid allows C-terminal knock-in of the fluorescent reporter mClover3 in the TPH1 locus and was generated using pUC118 as a backbone. First, the endogenous SapI site in PUC118 was inactivated. Then, a selection cassette (PGK promoter driven expression of blasticidin) flanked by LoxP and two SapI sites was cloned into the SapI-inactivated pUC118 using infusion cloning (638910, Takara). Subsequently, a P2A sequence and the fluorescent protein mClover3 was PCR amplified (Phusion High fidelity DNA polymerase, M0530S, NEB) from the Addgene plasmid #74252 and cloned upstream of the selection cassette using infusion cloning (638910, Takara) and NotI (R0189S, NEB) digestion of the pUC118 selection-cassette containing plasmid. Next, homology arms corresponding to the genomic regions, approximately 1000bp, upstream and downstream of the *TPH1* stop codon were PCR amplified (Phusion High fidelity DNA polymerase, M0530S, NEB) from genomic DNA (extracted and purified from human small intestinal organoid DNA). The PCR primers contained overhangs allowing subsequent Golden Gate cloning (Table S1).

The PCR amplified homology arms were purified (QIAquick PCR Purification Kit, 28104, QIAGEN) and finally, the targeting vector was generated by SapI (R0569S, NEB) mediated Golden Gate insertion of the homology arms into the pUC118 selection-cassette containing plasmid.

The sgRNA was selected based on the WTSI website (https://wge.stemcell.sanger.ac.uk//) and chosen as close to the TPH1 stop codon as possible. The gRNA sequence overlapped with the stop codon, so that the homology vector was not cut. The target sequence was ordered as two complementary oligos (IDT) and cloned in the Cas9-EGFP vector (addgene plasmid #48138) following the protocol described before (Ran et al., 2013).

For the generation of *HHEX* and *LMX1A* knockout organoids, gRNAs were selected using the WTSI website and cloned in the Cas9-EGFP vector (addgene plasmid #48138) following the protocol described before (Ran et al., 2013). gRNAs used in this story are presented in Table S1.

Human intestinal organoids were transiently transfected using a NEPA21 electroporator and a previously developed protocol (Fujii et al., 2015). For electroporation, organoids were dissociated into small clumps of cells (2-10 cells) and washed twice with Optimem. The resulting pellet was resuspended in BTXpress solution (BTX) with 15 μg of plasmids, after which electroporation was performed. 3-7 days after electroporation, either mCherry (for generation of NHEJ-mediated reporter organoids) or EGFP (for generation of *HHEX* and *LMX1A* knockout lines) positive cells were sorted using a FACS-ARIA (BD Biosciences). Rho kinase inhibitor (Y-27632 dihydrochloride; 10 μM, Abmole) was added to the culture medium up to 1 week after sorting to enhance single cell outgrowth. All reporter organoids were generated in organoid lines also transduced with *NEUROG3*-overexpression (with or without dTomato) vector.

For generation of CRISPR-HOT reporter organoids: Cells were transfected 1) with a gRNA targeting the hormone locus near its stop codon, 2) a vector encoding mNeon or tdTomato and 3) a vector encoding Cas9, a constitutively produced mCherry fluorescent molecule and a gRNA linearizing the vector encoding the fluorescent molecule. Five days later, transfected cells were sorted for mCherry and plated as single cells. After two weeks, NEUROG3 was induced in the resulting clonal organoids to visualize expression of the fluorescent fusion hormones. Typically, the first fluorescent organoids appeared 2-3 days later and were then clonally expanded. Organoids where fluorescent cells appeared during EEC differentiation were picked, digested using TrypLE (TryplE Express; Life Technologies) and clonally expanded to establish stable knock-in organoid lines.

Organoids grown from Cas9-EGFP transfected cells were genotyped for *HHEX* and *LMX1A* to confirm homozygous frameshift mutation (primers in Table S1).

Calcium sensor

A red calcium probe (pTorPE-R-GECO1, addgene plasmid #32465) was used as a template to engineer a cyan genetically encoded calcium probe. The cpApple was replaced with a circular permuted mTurquoise. The resulting probe was dubbed Tq-Ca-FLITS (Turquoise Calcium Fluorescence Lifetime Indicator for Truthful Sensing). A triple nuclear localization signal (3xnls) was added to the N terminus of the calcium probe to simplify analysis. Details of the engineering and characterization will be described elsewhere (van der Linden et al., unpublished).

PCRs were performed on Tq-Ca-FLITS (Fw AAACAAGCGGGAGACGTGGAGGAAAACCCTGGACCTCTCGAGatgggatcagatc caaaaaagaagag, Rev ATGGCACTAGGCTAGTTCTAGAcCTACTTCGCTGTCATCATTTGGAC) as well as H2B-mMaroon (Fw TC GGCGCGCCACGCGT, Rev CGTCTCCCGCTTGTTTCAGTAGACTAAAATTCGTCGCGCCAGATCCGCTAGCattaagtttgtgcccc) and the two PCRs were cloned into a lentiviral vector using InPhusion Cloning (Takara), to produce H2B-mMaroon-P2A- Tq-Ca-FLITS, two simultaneously expressed cistrons separated by a de-optimized P2A (Lo et al., 2015).

Live cell imaging of calcium reporter organoids

H2B-mMaroon-P2A-Tq-Ca-FLITS organoids were imaged on a Leica SP8 confocal laser scanning microscope, equipped with Argon laser and White Light Laser, the latter allowing spectral flexibility for optimal visualization of all fluorophores. For cell type identification, cells were first imaged in 5 channels (Tq-Ca-FLITS-mTurquoise2, Clover, TdTomato, H2B-Maroon and transmitted light) and subsequently Tq-Ca-FLITS and H2B-mMaroon were time lapse imaged during administration of beta-ionone in XYZT-mode. Post-acquisitional analysis was done with custom-made Fiji-script.

Transmission electron microscopy

Organoids were fixed with 1.5% glutaraldehyde in 0.1M cacodylate buffer. They were kept in the fixative for 24 h at 4°C. Then, they were washed with 0.1M cacodylate buffer and potsfixed with 1% osmium tetroxide in the same buffer containing 1.5% potassium ferricyanide for 1 h (dark) at 4°C. Then the samples were dehydrated in ethanol, infiltrated with Epon resin for 2 days, embedded in the same resin and polymerized at 60°C for 48 h. Ultrathin sections were obtained using a Leica Ultracut UCT ultramicrotome (Leica Microsystems, Vienna) and mounted on Formvar-coated copper grids. They were stained with 2% uranyl acetate in water and lead citrate. Then, sections were observed in a Tecnai T12 electron microscope equipped with an Eagle 4kx4k CCD camera (Thermo Fisher Scientific, the Netherlands).

Alternatively, organoids were chemically fixed at 4°C with a mixture of 2% paraformaldehyde and 0.2% glutaraldehyde in PB buffer. After washing with PB containing 50 mM glycine, cells were embedding in 12% gelatine and infused in 2.3 M sucrose. Mounted gelatine blocks were frozen in liquid nitrogen. Thin sections were prepared in an ultracryomicrotome (Leica EM Ultracut UC6/FC6, Leica Microsystems, Vienna, Austria). Ultrathin cryosections were collected with 2% methylcellulose in 2.3 M sucrose. The observations were performed in an Electron Microscope Tecnai T12 as mentioned.

Immunostaining

Organoids were stained as described before (Beumer et al., 2018). In brief, organoids were removed from the BME using ice-cold AdDMEM/F12 (GIBCO), after which these were fixed in formalin for at least 2 h at room temperature. Next, the organoids were washed and blocked for at least 15 min in 2% donkey serum in PBS. After blocking, the cells were permeabilized in 0.5% Triton X-100 (Sigma) in PBS for at least 15 min. Primary antibodies used were goat anti-chromogranin A (1:500; Santa Cruz), goat anti-cholestocystokin (sc-21617,1:100; Santa Cruz), rabbit anti-neurotensin (sc-20806,1:100; Santa Cruz), goat anti-somatostatin (sc-7819, 1:100; Santa Cruz), goat anti-serotonin (ab66047, 1:1,000, Abcam), rabbit anti-gastric inhibitory polypeptide (ab22624-50, 1:500; Abcam), goat anti-GLP1 (sc-7782, 1:100; Santa Cruz), rabbit anti-GLP1 (ab22625, 1:200; Abcam), rabbit anti-MLN (HPA069392, 1:200, Atlas antibodies), mouse anti-Gastrin (60346, 1:200, Proteintech), mouse anti beta-Catenin (610154, 1:100; BD transduction laboratories), goat anti-Ghrelin (sc-10368, 1:200; Santa Cruz), rabbit anti-Neuropeptide W (NBP2-57337, 1:100; Novus), rabbit anti-Precerebellin (ABN304, 1:100; Sigma) and rabbit anti-PPY (HPA032122, 1:200; Atlas antibodies). Organoids were incubated with the corresponding secondary antibodies Alexa488-, 568- and 647-conjugated anti-rabbit and anti-goat (1:1,000; Molecular Probes) in blocking buffer containing 4′,6-diamidino-2-phenylindole (DAPI; 1;1,000, Invitrogen). Sections were embedded in Vectashield (Vector Labs) and imaged using a Sp8 confocal microscope (Leica). Image analysis was performed using ImageJ software.

Fluorescent *in situ* hybridization

FISH was performed using the RNAScope® Multiplex Fluorescent Reagent Kit v2 (Advanced Cell Diagnostics) according to the manufacturer's protocol (Wang et al., 2012). In brief, paraffin embedded ileal surgical sections were deparaffinized, treated with hydrogen peroxide for 10 min and boiled in target retrieval buffer for 15 min before a 30-min protease treatment. Probes directed against *CHGA/ SCTR, CHGA/GCG and GHRL/ IL20RA* were multiplexed, respectively, amplified and detected using fluorescent probes based on opal dyes. Slides were counterstained with DAPI for 30 s, mounted using ProLong Gold Antifade Mountant (Thermo Fisher scientific) and images were obtained using a SP8 confocal fluorescent microscope (Leica).

ELISA

The supernatant from organoids either cultured in ENR for 5 days or differentiated toward EECs were collected after a 24 h stimulation with 10 μM forskolin (Tocris). GLP-1 concentration was measured using a GLP-1 EIA kit (Rab0201 from Sigma that detects both full-length and N-terminal cleaved GLP-1) following the manufacturer's protocol.

RNA isolation and quantitative PCR

Organoid RNA was isolated using a RNAeasy kit (QIAGEN), following the manufacturer's protocol. Quantitative PCR (qPCR) analysis was performed using biological and technical duplicates as described described before (Muñoz et al., 2012). Primers were designed using the NCBI primer design tool, tested using a standard curve, and are presented in Table S1.

Processing human intestinal tissue for single cell RNA sequencing

Human intestinal mucosal biopsies were obtained from patients undergoing colonoscopy at Addenbrooke's Hospital, Cambridge, UK. All patients gave informed consent for extra biopsy samples to be taken for research use when undergoing elective colonoscopy (REC 17/EE/0265). Only those patients with a macroscopically normal mucosa and subsequent histological confirmation of a normal mucosa with no intestinal pathology were included in this study.

Once acquired, biopsies were immediately placed into Hanks Buffered Saline Solution (HBSS) and washed three times in fresh HBSS. Biopsies were then placed into an HBSS solution containing 1.07 Wünsch units/mL Liberase DH (Roche) and 70 U/mL hyaluronidase (Merck) and incubated at 37°C for 15 min while on a plate shaker at 750 rpm. The samples were then mechanically disrupted by pipetting the solution up and down using a p1000 pipette, and then incubated at 37°C for a further 15 min on a plate-shaker at 750rpm. The samples were then washed three times by pelleting the cells using centrifugation at 400 g for 4 min, removal of the supernatant and resuspending in DMEM/F12 (Thermofisher). On the third wash the cells were left in suspension, and a 10 μL aliquot was placed into a Countess ® Automated Cell Counter to estimate the cellular concentration.

3,000 cells suspended in DMEM/F12 (Thermofisher) were loaded into an individual channel of a 10x single cell chip as per the manufacturer's protocol (version 2, 3′) and run in standard conditions by the chromium controller (10X genomics). cDNA libraries were prepared according to the manufacturer's protocol and sequenced on an Illumina Hi-seq 4000 (2x50bp paired-end reads).

Raw sequence reads in FASTQ format were aligned to the human transcriptome GRCh38-1.2.0 using Cellranger v2.1.1(10x Genomics) with default parameters. Scanpy v1.4 (Wolf et al., 2018) was used for data processing, quality control and dimensionality reduction. We excluded cells with less than 200 genes and genes expressed in fewer than 3 cells. After quality control, the dataset contained 11,302 cells. EECs were identified and subclustered based on the expression of at least 10 CHGA counts per cell. ‘‘

Single cell sorting for RNA sequencing from organoids

Organoids were dissociated to single cells using a 10-min incubation with TrypLE (TrypLE Express; Life Technologies) and repeated mechanical disruption by pipetting. Cells were sorted using a BD FACS Aria (BD Biosciences) based on fluorescence levels. For single cell RNA sequencing, individual cells were collected in 384-well plates with ERCC spike-ins (Agilent), reverse transcription primers and dNTPs (both Promega). Single cell sequencing was performed according to the Sort-seq method (Muraro et al.,

2016). Sequencing libraries were generated with TruSeq small RNA primers (Illumina) and sequenced paired-end at 60 and 26 bp read length, respectively, on the Illumina NextSeq.

For bulk RNA sequencing, cells were sorted into Eppendorf tubes containing RLT buffer (RNeasy kit, QIAGEN). 5,000 – 30,000 cells were sorted per reporter in duplicates (and triplicates for tdTomato negative cells). RNA was extracted using the RNeasy mini kit (QIAGEN) following the manufacturer's instructions. Sequencing libraries were generated using a modified CELseq2 protocol (Hashimshony et al., 2016). 75 bp paired-end sequencing of libraries was performed on an Illumina NextSeq platform.

Single cell RNA sequencing analysis from organoids

Reads were mapped to the human GRCh37 genome assembly. Sort-seq read counts were filtered to exclude reads with identical library-, cell- and molecule barcodes. UMI counts were adjusted using Poisson counting statistics (Muraro et al., 2016). Cells with fewer than 2,000 unique transcripts were excluded from further analysis. The remaining cells in the EEC atlas were derived from the following sources: 1446 cells from duodenum, 2145 cells from ileum, 690 cells from colon.

Subsequently, RaceID3 was used for k-medoids based clustering (knn = 10) of cells and differential gene expression analysis between clusters using the standard settings described at https://github.com/dgrun/RaceID3_StemID2_package.

The dataset was then subsetted to require expression of EEC markers and exclude cells based on expression of markers of other cell types with the following transcript count cutoffs: CHGA > 5; MUC2 < 5; FABP1 < 15; LYZ < 15; OLFM4 < 10. The resulting set of EECs was again subjected to clustering (knn = 5) and differential gene expression as described above.

For reporter analyses, cells sorted by fluorescent reporter positivity were analyzed as one dataset per reporter to gain more detailed insights into single EEC subpopulations. The following deviations from standard settings were made per reporter: GCG: knn = 5; outlg = 1; probthr = 0.00001; perplexity = 10; MLN: knn = 10; probthr = 0.0000001; SST: knn = 10; perplexity = 20

For mouse validation, the tissue-derived single cell count tables from Gehart el al. (2019) were reanalyzed using the procedure and settings described above. No subsetting for EECs was performed.

Bulk RNA sequencing analysis

Reads were mapped to the human GRCh37 genome assembly. The counted reads were filtered to exclude reads with identical library- and molecule barcodes. Differential gene expression analysis was performed using the DESeq2 package (Love et al., 2014). For display in heatmaps, genes were ranked by fold change compared against tdTomato negative cells. After filtering for an adjusted p value < 0.05, the row z-score for the top 20 genes was calculated.

Preparation of secreted peptides and proteins for LC-MS

Organoids differentiated for 5 days to EECs were washed extensively in PBS and stimulated with 10 μM forskolin (Tocris). Conditioned media was collected for 24 h and supplemented with 1x cOmplete Protease Inhibitor Cocktail on harvest (Roche). Potential cell debris was removed by centrifugation at 10,000 x *g*, for 5 min at 4°C. Conditioned media supernatant was denatured in final 4 M Urea, 50 mM ammonium bicarbonate and fractionated by molecular weight with a 10 kDa Vivaspin centrifugal device (Sartorius, Göttingen, Germany), at 12,000 x *g*, for 10 min at 4°C. (i) Endogenously processed peptides recovered from the filtrate were acidified to 5% formic acid, desalted by reversed phase C18 1cc columns (Waters Corporation, Milford, USA), further purified by home-made strong cation exchange STAGE tip, and dried by vacuum centrifugation. (ii) Longer secreted proteins in the 10 kDa retentate were recovered and diluted to final 2 M Urea, 50 mM ammonium bicarbonate, for reduction with dithiothreitol, alkylation with iodoacetamide, and overnight digestion with trypsin (Promega, Madison, USA) at 37°C. Digested peptides were similarly acidified to 5% formic acid, desalted by reversed phase C18 1 cc columns (Waters), and dried by vacuum centrifugation.

Preparation of FACS-sorted EECs for proteome analyses

FACS sorted enteroendocrine cells were lysed in 8M Urea, 50 mM Ammonium bicarbonate, 0.5% Sodium deoxycholate, 1x cOmplete protease inhibitor, 50 μg/mL DNase I, and sonicated with the Biorupter (3 cycles, 20 s on, 20 s off at 4°C) (Diagenode, Liege, Belgium). Cell debris was pelleted by centrifugation at 14,000 x *g* for 1 h at 15°C, and supernatant containing extracted proteins were reduced, alkylated, diluted 4 times with 50 mM ammonium bicarbonate, and digested sequentially with Lys-C (Wako) and trypsin (Promega). Peptide digests were quenched to 5% formic acid, and sodium deoxycholate was precipitated and removed by centrifugation at 14,000 x g, 4°C for 10 min. Peptides in the supernatant were diluted to final 20% acetonitrile and purified by SCX STAGE tips. Eluted peptides were dried by vacuum centrifugation.

LC-MS

Peptides were reconstituted in 2% FA for LC-MS injection. Data was acquired using an UHPLC 1290 system (Agilent, California, USA) coupled to an Orbitrap HF-X mass spectrometer (Thermo Scientific, Massachussets, USA). Peptides were first trapped in a 2 cm × 100 μM Reprosil C18 trap column (Dr Maisch, Ammerbuch, Germany) of 3 μM pore size for 5 min in solvent A (0.1% formic acid in water). After trapping, samples separated in an analytical column (Agilent Poroshell, EC-C18, 2.7 μm, 50 cm × 75 μm) using a gradient of 0.1% formic acid in 80% acetonitrile (solvent B). Depending on total peptide input, species complexity, and elution profiles, different LC gradient lengths were used for FACS-sorted whole proteomes (35min, 13%–40% solvent B), secreted proteins (65min, 13%–40% solvent B) and endogenously processed peptides (95min, 13%–44% solvent B). MS acquisition was performed

in data-dependent mode. Full scans (MS1) were acquired from 375 to 1600 m/z at resolution 60,000, with 20 ms injection time and $3x10^6$ AGC target value. The TOP 15 most intense precursor ions where selected for fragmentation using 1.4 m/z isolation window. Isolated precursors were fragmented using high energy C-trap dissociation (HCD) at normalized collision energy (NCE) of 27%. MS2 scans were acquired at resolution 30,000, with 50 ms injection time and an AGC target value of $1x10^5$. Exclusion times were set to 8, 12 or 16 s for proteomics, digested retentates and endogenously processed peptides respectively.LC-MS injection loads were adjusted to the sample of least quantity, such that all LC-MS measurements to be compared were matched in TIC intensity.

Proteomics data analysis

Collected spectral data was processed using Proteome Discoverer 2.3 (Thermo Scientific, Massachussets, USA), and searched using Sequest HT search engine, against UniProt *Homo sapiens* database (173235 entries, downloaded in August 2019, including common contaminants). Precursor and fragment mass tolerance were set to 10 ppm and 0.02 Da respectively. Protein N-terminal Acetylation and methionine oxidation were set as variable modifications. For the endogenously processed peptides, peptides of length 6 - 50 amino acids were generated from the database upon unspecific cleavage. For digested retentates and analyses of the FACS-sorted EECs proteomes, cysteine carbamidomethylation was set as fixed modification and up to 2 missed trypsin cleavages were allowed. Identified peptides were filtered to 1% FDR using the Percolator algorithm (The et al., 2016).

In proteome analyses, intensities of proteins detected in 2 out of 3 replicates in at least one group were *log(2)* transformed and missing values were imputed from the normal distribution independently for each sample using Perseus software (v_1.6.2.2)(Tyanova et al., 2016). Processed data was assessed for statistical significance among the groups using One-way ANOVA or Student's t test, and resulting *p-value*s were corrected for type I error using Benjamini-Hochberg approach (q-value). Tukey Honest Significant Difference test was performed to assess difference between the groups when required. Hormone processing plots were generated from peptides identified at high confidence (1% FDR) in at least 2 out of 3 replicates. For novel secreted products, peptides identified in 3 out of 3 replicates and not observed in the secretome of control organoids were used. Peptide sequences identified based on retention time alignment (not supported by spectral evidence) were not used in hormone processing plots. Statistical analysis and plots were generated using *in-house* built R scripts (R version 3.6.0). Gene ontology enrichments were performed with Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 (Huang et al., 2009), using all the proteins identified in the bulk EEC proteome as reference list.

QUANTIFICATION AND STATISTICAL ANALYSIS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to the sample allocation during experiments and outcome assessment. All data are presented as mean ± standard error of the mean (SEM), unless stated otherwise. Value of n is always displayed in the figure as individual data points, and in the legends. Statistical tests included unpaired two-tailed t test for Figures 5F and Figure S1G.

Supplemental Figures

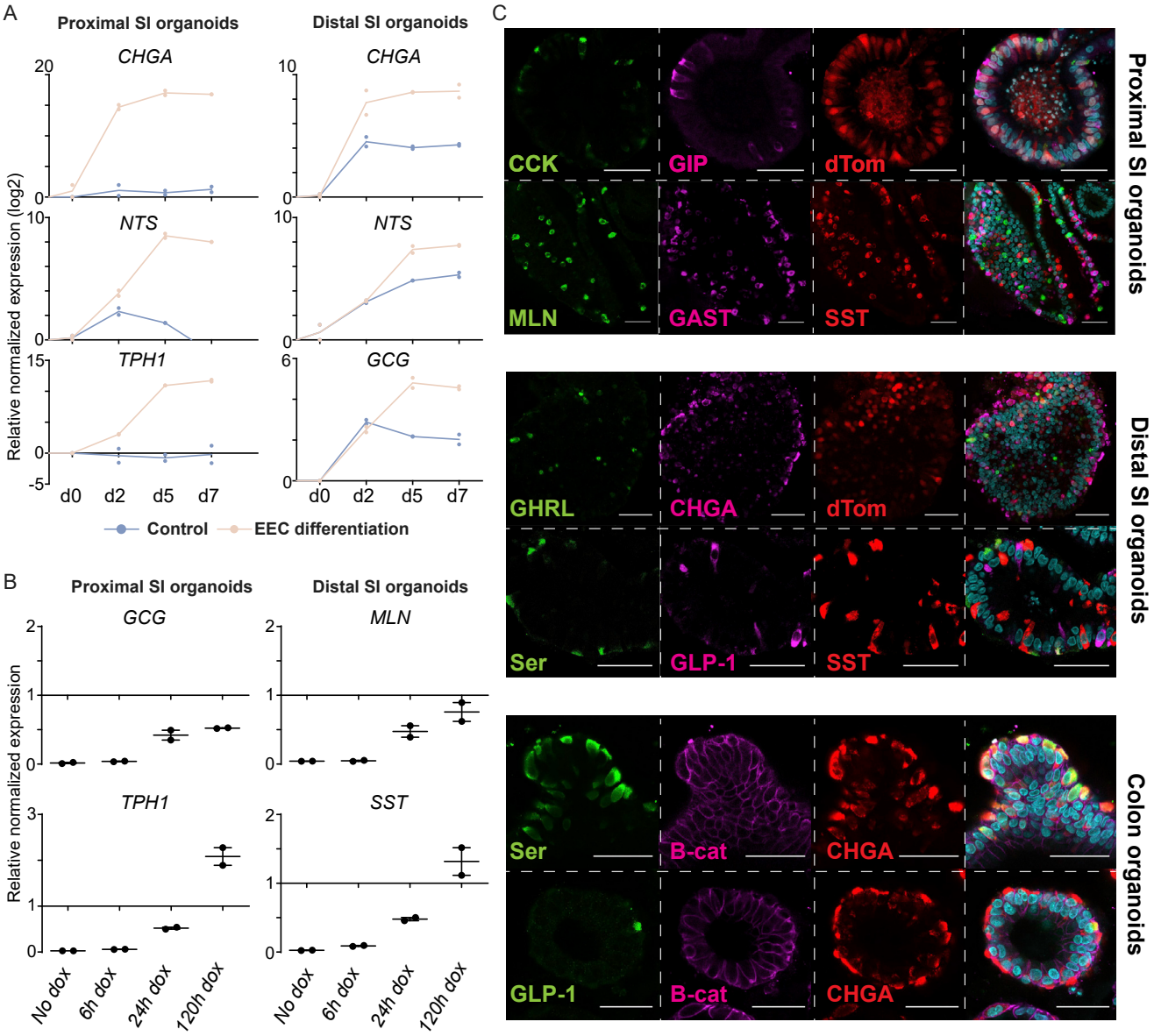


Figure S1. Enteroendocrine Cells in Human Intestinal Organoids Display Normal Co-expression Profiles, Related to Figure 1

(A) qPCR analysis showing expression of hormones over the course of EEC differentiation. Expressions levels are normalized to GADPH and relative to day 0. The experiment was performed in n = 2 independent experiments, and the individual datapoints are depicted.

(B) qPCR analysis showing expression of hormones after different durations of doxycycline (Dox) challenge. Organoids were differentiated for 5 days, and treated without dox, for 6 h, 24 h, 48 h or 120 h (the full differentiation time) dox. Expressions levels are normalized to GADPH and relative to 48 h dox treatment. The experiment was performed in n = 2 independent experiments, and the individual datapoints are depicted.

(C) Immunofluorescent staining of EEC-enriched organoids. Multiple hormones are expressed mutually exclusive. Scale bar is 50 μm.

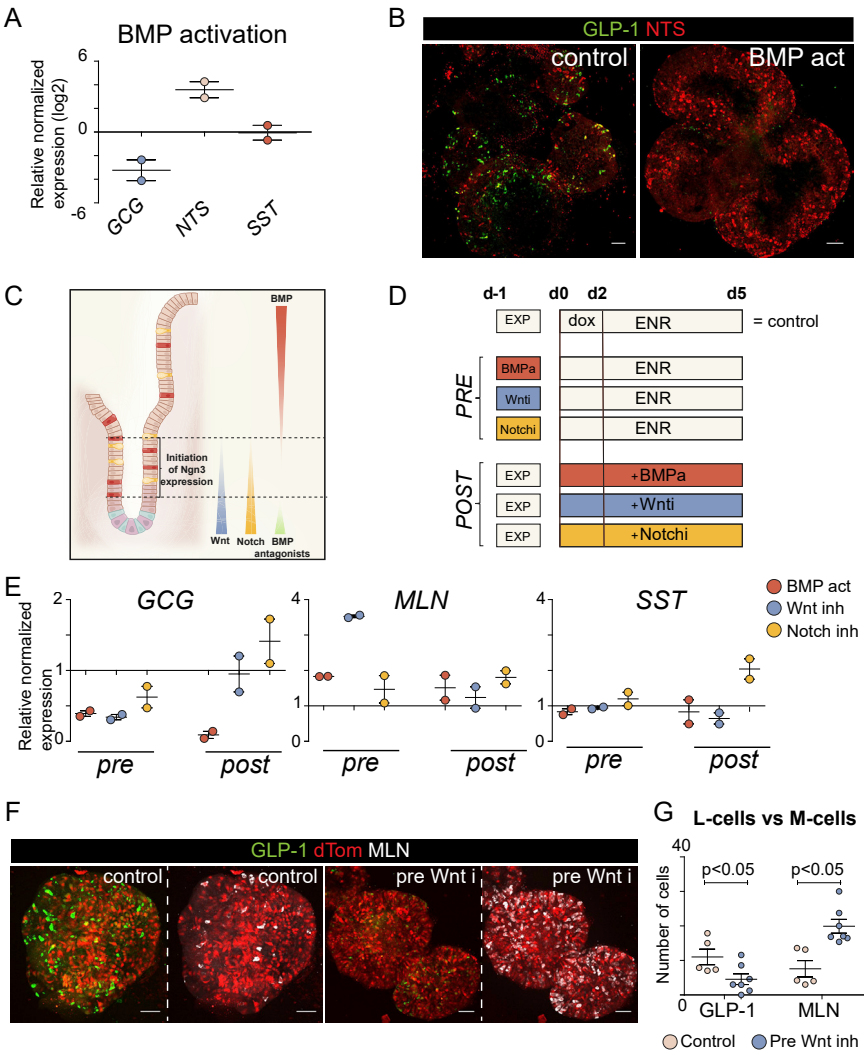


Figure S2. Manipulation of Wnt and BMP Signaling Allows Controlling Subspecification of EECs, Related to Figure 1

(A) qPCR analysis showing expression of hormones after BMP treatment. Expression levels are normalized to GADPH, and relative to a non-treated control. The experiment was performed in n = 2 independent experiments, and the mean expression and SEM are depicted.

(B) Immunofluorescent staining of BMP treated organoids. Scale bar is 50 μm.

(C) Schematic representation of morphogen gradients in the intestinal crypt related to sites of initiation of Neurogenin-3 (Ngn3) expression.

(D) Experimental paradigm. Different signaling pathways were modulated (BMP activation; BMPa, Wnt inhibition; Wnti, Notch inhibition; Notchi) either 24 h before (pre) or at the start (post) of *NEUROG3* expression mediated by doxycycline (dox) treatment. Control organoids were kept in standard expansion conditions (EXP) before dox treatment, and in standard differentiation conditions (ENR) after initiation of dox treatment.

(E) qPCR analysis showing expression of hormones after different treatments shown in d. Expression levels are normalized to GADPH, and relative to a non-treated control. The experiment was performed in n = 2 independent experiments, and the mean expression and SEM are depicted.

(F) Immunofluorescent staining of organoids differentiated toward EECs after a 24 h inhibition of Wnt (pre Wnt i). Organoids are shown as a maximum projection. Scale bar is 50 μm.

(G) Quantification of (F). Number of positive cells were counted on n = 5 organoid sections. Pre-inhibition of Wnt signaling caused a shift of L-cell to M-cell differentiation.

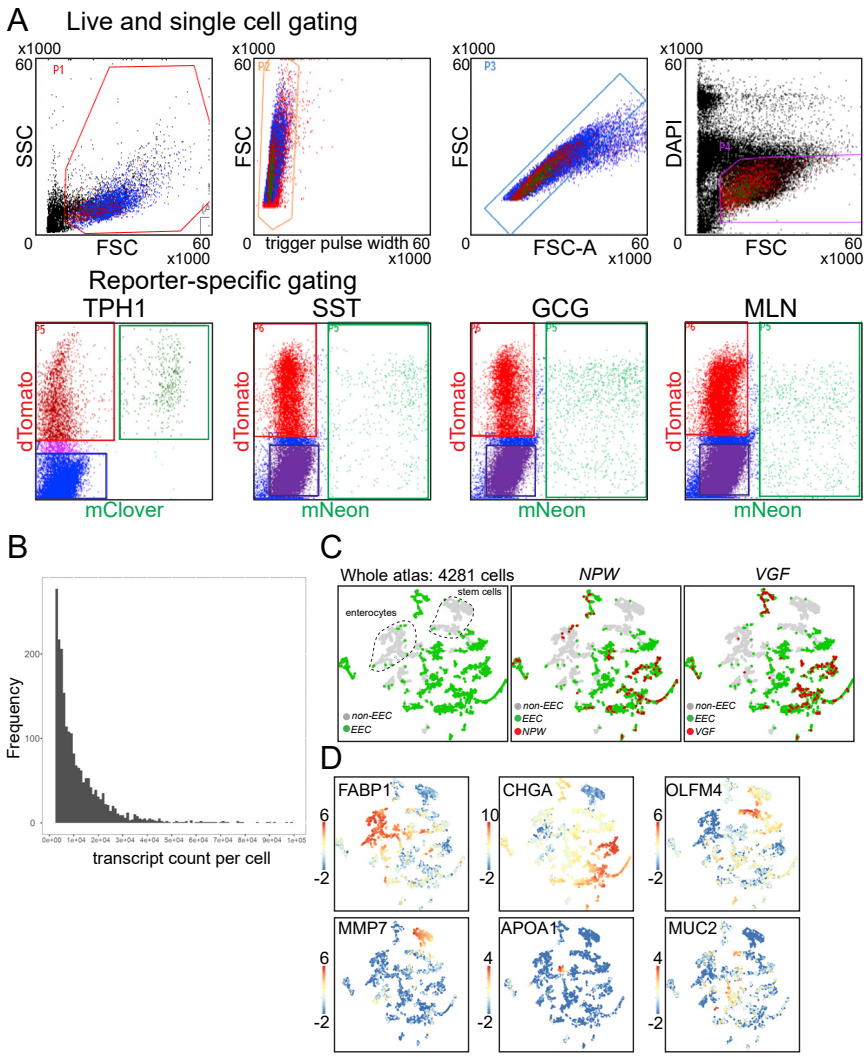


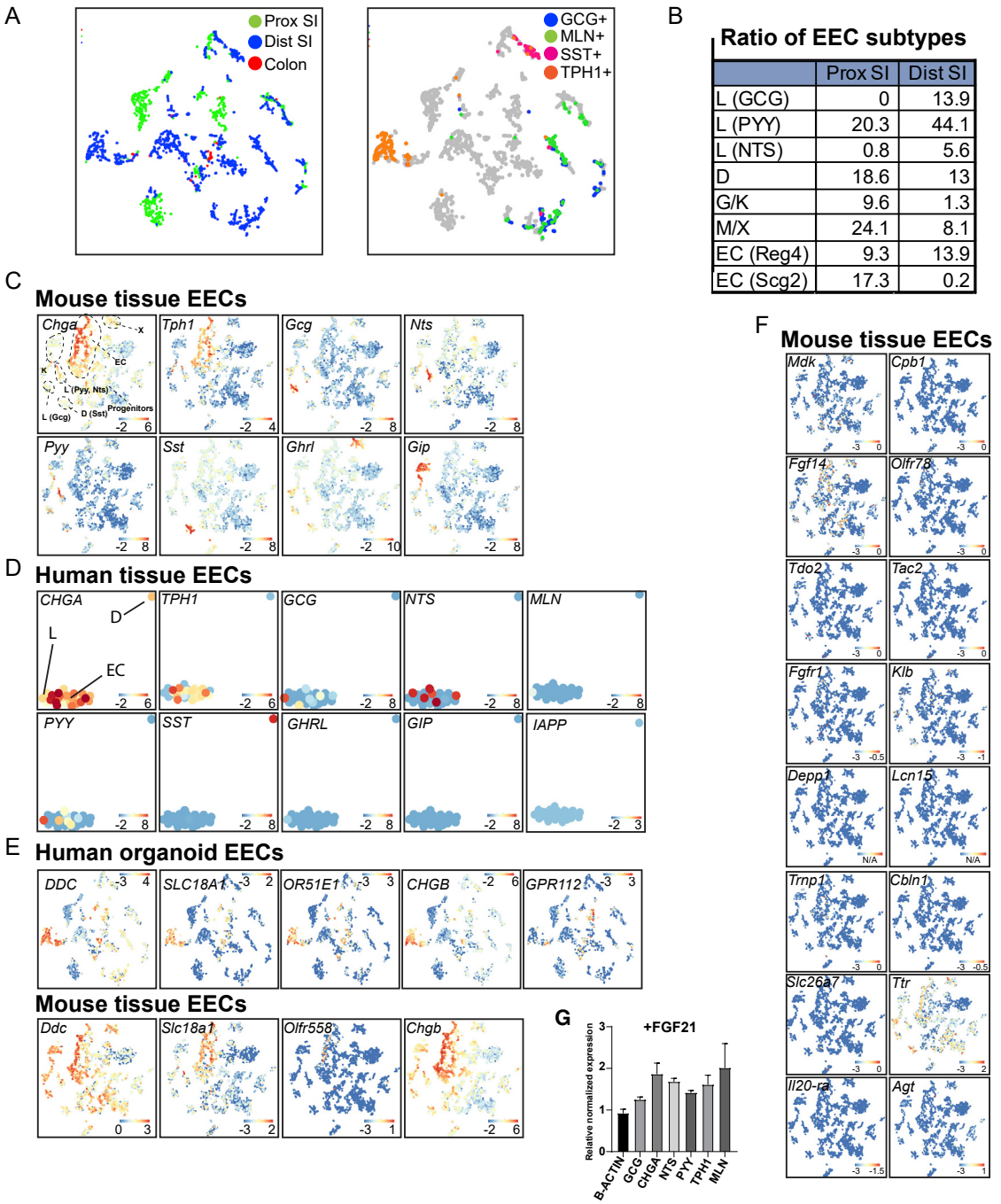
Figure S3. Generation of a Human Single-Cell RNA Sequencing Atlas of the Intestinal Tract Using Organoids, Related to Figure 3

(A) Fluorescence-activated cell sorting (FACS) gating parameters for sorting of different EEC subtypes from reporter organoids.

(B) Histogram displaying the total number of unique transcripts per cell (median number per cell is 7288 transcripts).

(C) A broad human intestinal organoid atlas (n = 4281 cells) generated by single cell RNA sequencing and displayed using a *t*-distributed stochastic neighbor-embedding (*t*-SNE) map. Cells defined as EECs (see methods) are shown in green. Cells expressing NPW or VGF (> 1 transcripts, respectively) are highlighted in red and are found exclusively in EEC clusters.

(D) t-SNE maps displaying lineage markers in the whole human intestinal organoid cell atlas (n = 8448 cells). Bars display color-coded unique transcript expression (logarithmic scale).



	Prox SI	Dist SI
L (GCG)	0	13.9
L (PYY)	20.3	44.1
L (NTS)	0.8	5.6
D	18.6	13
G/K	9.6	1.3
M/X	24.1	8.1
EC (Reg4)	9.3	13.9
EC (Scg2)	17.3	0.2

Figure S4. Single-Cell RNA Sequencing of Human EECs from Organoids and Tissue and Mouse EECs from Tissue, Related to Figure 3

(A) t-SNE maps displaying the origin (left, tissue; right, reporter organoid) of cells from the human EEC atlas (n = 2255 cells).

(B) The percentages of EECs corresponding to the different subtypes in proximal and distal SI organoids.

(C) t-SNE maps displaying the expression levels of hormones in the different murine EEC subtypes from intestinal tissue. Bars display color-coded unique transcript expression (logarithmic scale).

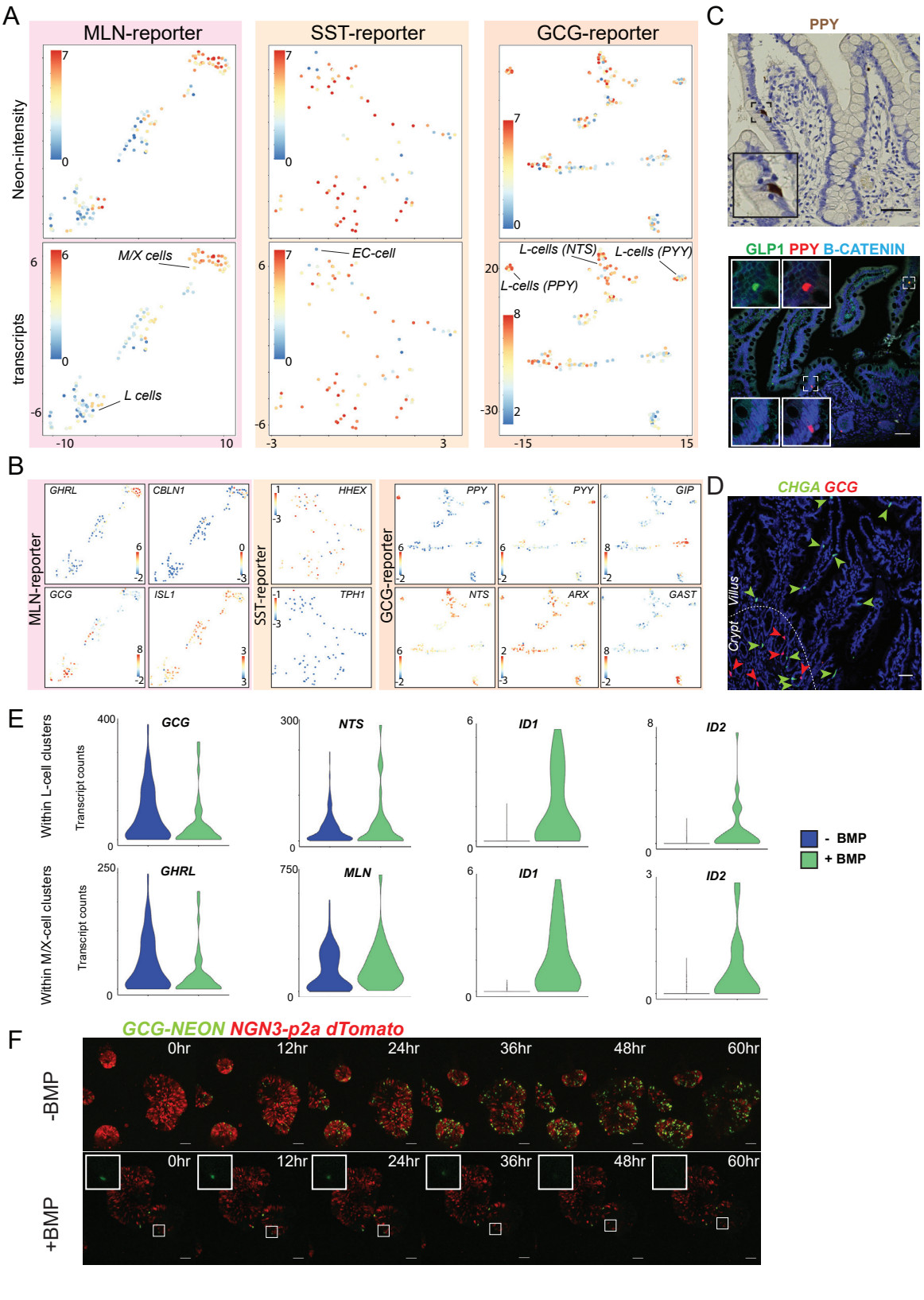
(D) t-SNE maps displaying the expression levels of hormones in the different human EEC subtypes from intestinal tissue. Bars display color-coded unique transcript expression (logarithmic scale).

(E) t-SNE maps displaying conserved expression of different EC markers in human and mouse EECs. Bars display color-coded unique transcript expression (logarithmic scale).

(F) t-SNE maps displaying the levels of hormone and marker gene expression of human M/X cells in the different murine EEC subtypes from intestinal tissue. Bars display color-coded unique transcript expression (logarithmic scale).

(legend continued on next page)

(G) qPCR analysis showing expression of hormones after FGF-21 treatment during the 5 day EEC differentiation. Expressions levels are normalized to GADPH and relative to control organoids that are EEC differentiated without FGF-21 treatment. B-ACTIN is displayed as second housekeeping gene. The experiment was performed in n = 2 independent experiments, and the mean expression and SEM are depicted.



(legend on next page)

Figure S5. Subclustering of Human EEC Subtypes and Responses to BMP Activation, Related to Figure 3

(A) Subclustering was performed on EECs sorted from different reporter organoids. t-SNE maps displaying the correlation between transcript levels and reporter intensity

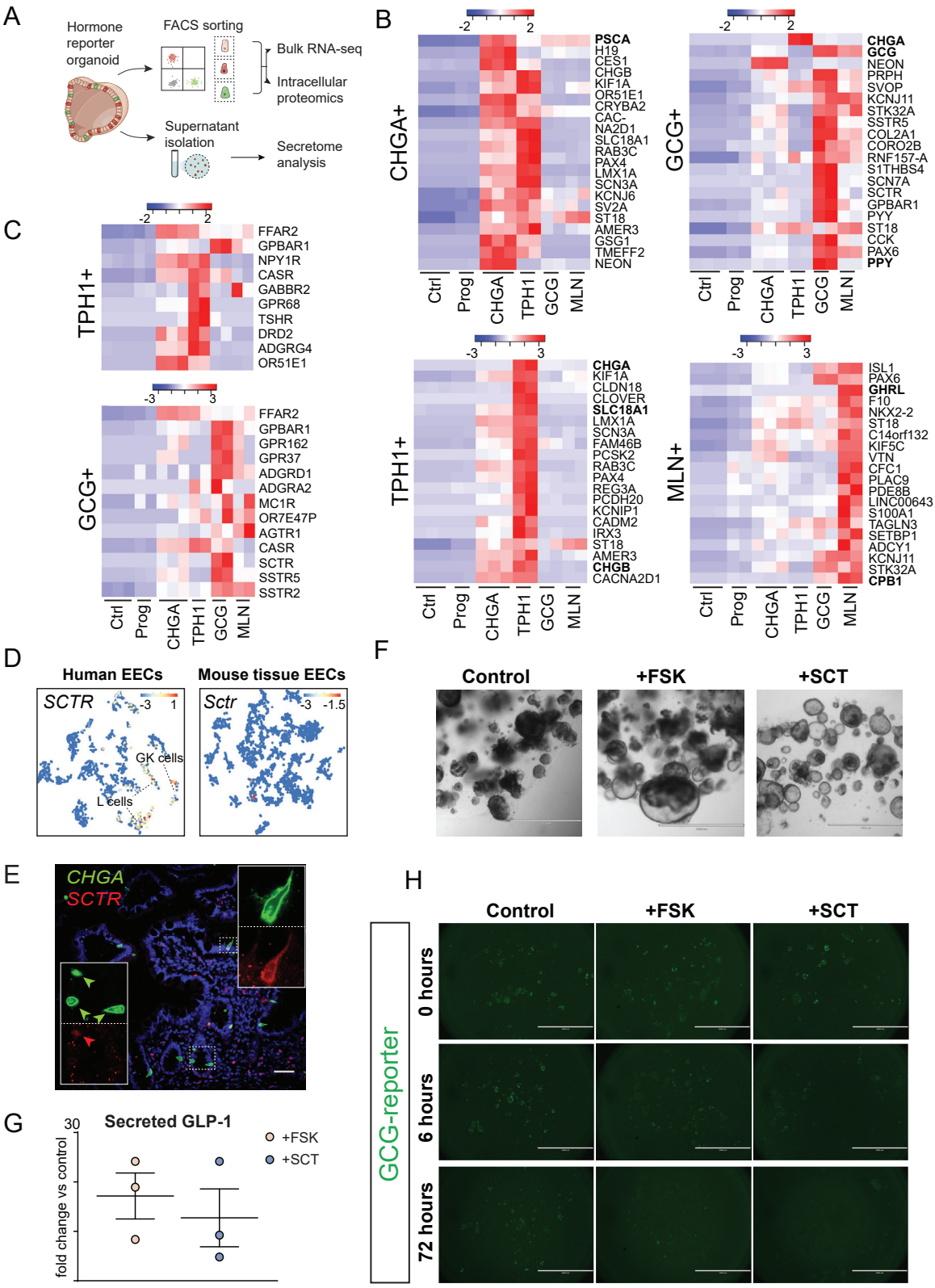
(B) t-SNE maps displaying different hormones in EECs from the different reporter organoids. *PPY*-expressing cell form a distinct cluster of *GCG*$^{+}$ cells.

(C) Immunohistochemistry of human ileal sections confirms PPY expression *in vivo*. Scale bar is 50 μm.

(D) Fluorescent *in situ* hybridization on human ileal section shows crypt-restricted expression of *GCG* (red arrowheads), whereas *CHGA* expression (green arrowheads) is expressed also in the villus. Scale bar is 50 μm.

(E) Violin plots depict the expression levels of selected hormones in single BMP-treated cells versus untreated cells in the EEC single cell RNA sequencing atlas. *ID1* and *ID2* are BMP target genes that confirm specific pathway activation of BMP agonist-treated cells.

(F) Snapshots are shown of *GCG-neon* reporter organoids that were treated with BMP after 2 days of dox treatment to induce NEUROG3-dTomato expression (= 0 h time point). BMP treatment blocks the appearance of GCG^{+} cells, while pre-existing L-cells downregulate GCG expression. No cell death is observed. Scale bar is 50 μm.



(legend on next page)

Figure S6. Bulk Transcriptomic Profiling of Sorted Enteroendocrine Cell Subtypes, Related to Figure 6

(A) Experimental paradigm. Hormone reporter organoids are differentiated, after which subpopulations of EECs are sorted using FACS and processed for bulk RNA-sequencing or intracellular proteomics. In a separate experiment, the supernatant of organoids is collected after 24 h forskolin stimulation and processed for proteomic analyses to determine the EEC secretome.

(B) Heatmaps showing the 20 most significant RNA markers enriched in purified reporter populations. In bold genes are highlighted that are also among the 20 most significant markers on protein level. Colored bars represent Z-scores.

(C) Heatmaps showing receptor expression most unique to $TPH1^+$ or GCG^+ cells. The receptor for EEC hormone Secretin (SCTR) is expressed highly in L-cells but not ECs, while ECs display unique expression of the PYY receptor (NPY1R).

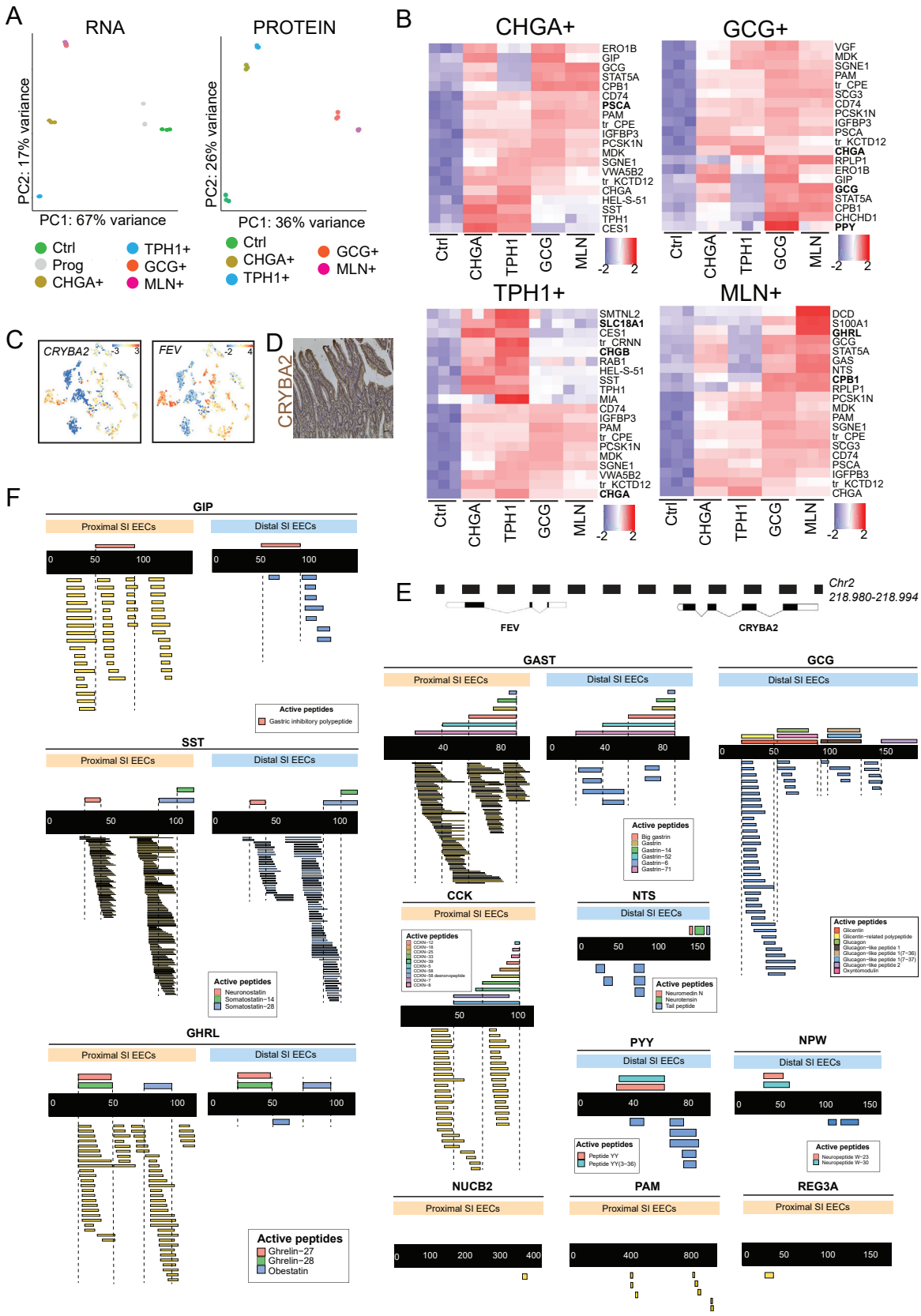
(D) t-SNE map displaying the expression level of the Secretin receptor (SCTR) in the EEC single cell sequencing atlas. Bars display color-coded unique transcript expression (logarithmic scale).

(E) Fluorescent *in situ* hybridization on human ileal sections shows rare *SCTR*-expressing cells (red arrowhead) that sometimes co-express *CHGA* (green arrowhead). Scale bar is 50 μm.

(F) Representative bright-field images of EEC-differentiated organoids after 24 h forskolin (FSK) or Secretin (SCT) treatment. Both FSK and SCT treatment causes swelling of organoids, indicative of cAMP activation. Scale bar is 1 mm.

(G) ELISA showing the fold increase in GLP-1 concentrations of EEC-enriched organoids after treatment with FSK or SCT. The experiment was performed in n = 3 independent experiments, and the mean fold change and SEM are depicted.

(H) GCG-reporter organoids were differentiated toward EECs and treated with FSK or SCT. Intracellular levels of GCG-neon reduce over the course of FSK and SCT treatment. Scale bar is 2 mm.



(legend on next page)

Figure S7. Proteomic and Secretomic Profiling of EECs, Related to Figure 7

(A) Principle component analysis (PCA) of RNA and protein data from different EEC populations. CHGA$^-$ cells were used as control. CHGA$^-$ cells that were positive for dTomato (induced NEUROG3-expression) were defined as EEC progenitors (Prog).

(B) Heatmaps showing the 20 most significant markers on protein level defining each EEC populations. Colored bars represent Z-scores. In bold genes are highlighted that are also among the 20 most significant markers on RNA level.

(C) t-SNE maps displaying the expressions level of *CRYBA2* and *FEV* in the EEC single cell sequencing atlas, illustrating a high degree of overlap in expression. Bars display color-coded unique transcript expression (logarithmic scale).

(D) Immunohistochemistry on human duodenal sections shows a lack of CRYBA2 expression *in vivo*. Scale bar is 50 μm.

(E) The location of the *CRYBA2* and *FEV* genes on Chromosome 2.

(F) Measured peptides (< 10 kDa) in the secretome mapping to different secreted prehormones are shown below the black bar. Data from proximal (yellow background) and distal SI organoid (blue background) supernatants are displayed.