

A proximal-to-distal survey of healthy adult human small intestine and colon epithelium by single-cell transcriptomics

Joseph Burclaff^{1,2,9}, R. Jarrett Bliton^{3,9}, Keith A Breau⁴, Meryem T Ok³, Ismael Gomez-Martinez⁴, Jolene S Ranek⁵, Aadra P Bhatt^{1,2,6}, Jeremy E Purvis^{5,7}, John T Woosley⁸, Scott T Magness^{1,2,3,4,6*}

¹Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

²Center for Gastrointestinal Biology and Disease, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

³Joint Department of Biomedical Engineering, University of North Carolina at Chapel Hill/North Carolina State University, Chapel Hill, North Carolina

⁴Department of Cell Biology and Physiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

⁵Curriculum in Bioinformatics and Computational Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

⁶Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

⁷Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

⁸Department of Pathology and Laboratory Medicine, University of North Carolina School of Medicine, Chapel Hill, NC

⁹These authors contributed equally

*Correspondence: magness@med.unc.edu

All authors declare no conflicts of interest

Author Contributions:

STM designed and supervised the project

JB, RJB, KAB, MTO, IG, STM obtained and analyzed data

RJB and KAB performed computational analysis

JB, RJB, STM wrote the manuscript

JTW analyzed histology slides

APB and JSR provided intellectual contributions

JSR and JEP reviewed computational analyses

All authors reviewed the final manuscript

All data will be available in the NCBI Gene Expression Omnibus: accession number GSE185224

Python Scripts allowing for main steps of our analysis to be performed will be available on GitHub

Abbreviations Used:

SI: Small intestine; scRNAseq: single-cell RNA sequencing; AC: Ascending Colon; TC: Transverse Colon; DC: Descending Colon; PC: Paneth Cell; FAE: Follicle Associated Epithelium; EEC: Enteroendocrine Cell; GC: Goblet Cell; DEG: Differentially Expressed Genes; ISC: Intestinal Stem Cell; TA: Transit Amplifying; ACC: Absorptive Colonocyte; AE: Absorptive Enterocyte; PGA: Partition-based Graph Abstraction; M-cell: Microfold Cell; GI: Gastrointestinal; crGC: Crypt-Resident Goblet Cell; icGC: Inter-Crypt Goblet Cell

Abstract

Background and Aims: Single-cell transcriptomics offer unprecedented resolution of tissue function at the cellular level, yet studies analyzing healthy adult human small intestine and colon are sparse. Here, we present single-cell transcriptomics covering the duodenum, jejunum, ileum, and ascending, transverse, and descending colon from 3 humans.

Methods: 12,590 single epithelial cells from three independently processed organ donors were evaluated for organ-specific lineage biomarkers, differentially regulated genes, receptors, and drug targets. Analyses focused on intrinsic cell properties and capacity for response to extrinsic signals along the gut axis across different humans.

Results: Cells were assigned to 25 epithelial lineage clusters. Human intestinal stem cells (ISCs) are not specifically marked by many murine ISC markers. Lysozyme expression is not unique to human Paneth cells (PCs), and PCs lack expression of expected niche-factors. BEST4⁺ cells express NPY and show maturational differences between SI and colon. Tuft cells possess a broad ability to interact with the innate and adaptive immune systems through previously unreported receptors. Some classes of mucins, hormones, cell-junction, and nutrient absorption genes show unappreciated regional expression differences across lineages. Differential expression of receptors and drug targets across lineages reveals biological variation and potential for variegated responses.

Conclusions: Our study identifies novel lineage marker genes; covers regional differences; shows important differences between mouse and human gut epithelium; and reveals insight into how the epithelium responds to the environment and drugs. This comprehensive cell atlas of the healthy adult human intestinal epithelium resolves likely functional differences across anatomical regions along the gastrointestinal tract and advances our understanding of human intestinal physiology.

Keywords: scRNAseq; Cell Atlas; Intestinal Stem Cell; Paneth cell; BEST4

Introduction

Colloquially called the 'gut', the small intestine (SI) and colon are distinct organs with overlapping and unique roles in maintaining health. A monolayer of epithelium lines the gut lumen, comprised of stem and differentiated cells that renew the epithelium each week¹. Broad health conditions develop at the intestinal epithelium, caused by pathological mucosal immunity², dysregulation of carefully orchestrated cell-cell signaling, or disrupted synergy between stem cell-driven self-renewal and production of differentiated lineages. This complexity is little understood at the cellular level.

The generalized function of the gut epithelium is maintaining barrier function, absorbing nutrients, and regulating water. Cellular roles include ion balance, hormone production, mucus production, and signaling through the luminal-epithelial-immune axis. While physiological functions differ across the gut length, how lineages differ along the SI-colon axis is poorly understood. Whether adult gut epithelial lineages adopt regional fates and functions is a central question of human gut physiology and disease.

Single-cell RNA sequencing (scRNAseq) approaches have provided unprecedented transcriptomic resolution of cells and revealed unappreciated cellular heterogeneity. Studies in mouse intestines³⁻⁵ led to human scRNAseq studies analyzing fetal gut development⁶⁻⁸ and adult colonic⁹⁻¹², ileal¹³⁻¹⁵, and duodenal¹⁶ epithelium. To date, one study compares adult human ileum and regionally-unspecified colon¹³, and one recent report compiles samples from across the gut yet has limited epithelial analysis¹⁷. Several human gut regions have sparse scRNAseq analysis available, with no studies analyzing regional differences within SI or colon. Addressing these gaps requires technically and logistically challenging approaches.

Here we comprehensively survey adult human gut epithelium using transplant-grade organs. scRNAseq libraries were prepared from epithelial cells from duodenum, jejunum, ileum, and ascending- (AC), transverse- (TC), and descending- (DC) colon from three donors. This robust cellular library avoids intra-donor batch effects and allows for observations between individual patients. Using this dataset, we probe understudied human lineages including Paneth cells (PCs), SI BEST4⁺ cells, and Follicle Associated Epithelium (FAE). We define comprehensive transcriptional signatures for lineages across the entire gut, highlight differences between human and mouse markers, and generate regional atlases of functional gene

families across the proximal-to-distal axis. We further probe how lineages might be affected by extrinsic signaling by mapping receptor families and analyzing primary gene targets of approved drugs.

Methods

Donor Selection

Human donor intestines were received from HonorBridge (Durham, NC) with acceptance criteria: age ≤ 65 years, brain-dead only, HIV(-), hepatitis(-), syphilis(-), tuberculosis(-), COVID-19(-), and no bowel surgery, severe abdominal injury, cancer, or chemotherapy. Pancreas donors were excluded to avoid duodenum loss. UNC IRB determined this study does not constitute human subjects research.

Tissue processing

Intestines were transported on ice in University of Wisconsin Solution. Tissue dissection began within eight hours of cross-clamping. Fat/connective tissue were trimmed and intestinal regions separated: duodenum (most-proximal 20 cm); jejunum/ileum splitting remaining SI; colon split into thirds for AC/TC/DC. Two 3x3 cm mucosectomies were isolated from the center of each region for dissociation.

Mucosectomies were incubated in 10 mM NAC in dPBS at room temperature for 30 min to remove mucus, then washed in ice-cold Chelating Buffer¹⁸ + 100 μ mol/L Y-27632. Tissues were incubated in Chelating Buffer with 2 mmol/L EDTA and 0.5mmol/L DTT, then shaken to remove crypts. High-yield colon shakes were pooled, with SI shakes pooled to approximate 1:1 villus to crypt tissue by cell mass. Crypts and villi were dissociated to single cells using 4 mg/mL Protease VIII in dPBS + Y-27632 on ice for ~45min with trituration via a P1000 micropipette every 10 min. Cells were checked under a light microscope then filtered.

Sample preparation

Single cells were washed with dPBS + Y-27632, resuspended in Advanced DMEM/F12 + 1% Bovine Serum Albumin + Y-27632, then stained with AnnexinV-APC (1:100) and one TotalSeq Anti-Human Hashtag Antibody (1:100) per region to track all six regions with a single library preparation¹⁹. Cells were washed and resuspended in AdvDMEM + 1% BSA + Y-27632 for sorting on a Sony Cell Sorter SH800Z to enrich for live single epithelial cells. 25,000 cells were collected for each region, then regions were combined pre-sequencing. Library prep was performed with the Chromium Next GEM Single Cell 3' GEM, Library & Gel Bead Kit v3.1 (PN-100012). Sequencing was performed on an Illumina NextSeq 500.

Data preparation and Hashtag Calling

Harmony (v0.0.5) was used to integrate the top 40 principal components from each dataset for clustering and visualization²⁰. Leiden clustering was initialized with a kNN graph (k=10 neighbors) and a Leiden resolution of 0.92²¹ to resolve most expected physiological lineages. UMAPs were initialized with PAGA of identified Leiden clusters^{21,22}, then non-epithelial EPCAM-negative lineages were eliminated. Regional hashtag deconvolution followed published methods: raw hashtag read counts were normalized using centered log ratio transformation followed by k-medoid clustering of hashtag read counts, with k=6 medoids for donor 1 and k=7 medoids for donors 2 and 3¹⁹. Hashtag noise distributions were determined by removing the highest-expressing cluster, then fitting a negative binomial distribution to the remaining cells. Cells were considered positive for a hashtag with counts above the distribution's 99th percentile (p<0.01) threshold. Cells positive for multiple hashtags were excluded as likely doublets.

Results

Sample processing

We define SI and colon as 'organs' and duodenum, jejunum, ileum, AC, TC, and DC as 'regions'. Intestinal tracts were obtained from three organ donors (Fig. 1A,S1) with no history of cancer or intestinal disease, and healthy mucosa was verified by a pathologist (Fig. S2). Tissue was resected from each donor (Fig. 1A, Fig. S2) then epithelium was dissociated to single cells using cold protease to preserve RNA integrity. Each region's cells were stained with Cell Hashtag antibody-oligo conjugates^{8,19} to multiplex regions for library

preparation and sequencing, avoiding intra-donor batch effects and reducing cost (Fig. S4). FACS excluded dead cells and doublets (Fig. S3) prior to sequencing. After filtering for quality, transcriptional readouts for 12,590 total cells (4,330, 4,465, and 3,795 cells/donor) were obtained (Fig S1,S5).

Donor sequencing datasets were individually processed then combined, with principal components integrated with Harmony²⁰ to minimize donor-specific differences prior to dimensional reduction and Leiden clustering²¹. Most lineages formed distinct SI/colon-specific clusters, suggesting functional differences by location. No organ-specific clustering occurred for enteroendocrine cells (EECs) and secretory progenitors (Fig. S6). One cluster expressed PC and goblet cell (GC) markers, so sub-clustering resolved these lineages (Fig. S6). Our final dataset identifies all lineages by organ along with rare FAE (Fig. 1B). The integrated dataset shows overlapping cell distributions from each donor and region within all major lineages, demonstrating that post-sequencing hashtag deconvolution preserves transcriptomic features across samples and batches (Fig. 1C-E).

To define significant marker genes across the gut, we calculated differentially expressed genes (DEGs) in each lineage versus all other cells from each organ. We identified DEGs consistently enriched across all three donors and the combined dataset. Through this rigorous statistical evaluation, we developed a unique signature for all major lineages across the human SI and colon epithelium, a previously unavailable resource (Fig. 1F,G, Table S1,S2).

Proliferative Cells

We found human Intestinal Stem Cells (ISCs) differentially expressed classical markers *LGR5*, *ASCL2*, *SLC12A2*, and *RGMB* (Fig. 2A,B)²³⁻²⁵. Though enriched, *SMOC2*²⁶ was not a DEG in SI ISCs, as PCs express higher levels (see PCs). While *in situ* hybridization showed *OLFM4* marks human SI and colonic ISCs²⁷, our data indicate significantly higher levels of *OLFM4* in SI ISCs, with colonic *OLFM4* higher in transit amplifying (TA) cells and early absorptive colonocytes (ACCs) (Fig. 2A,B). This concurs with mouse, where *Olfm4* transcripts marked SI but not colon ISCs²⁴. Notably, *RARRES2* was enriched in colon ISCs, with low expression in SI ISCs (Fig. 2B). We found no gut-related literature on this retinoid response gene, providing an intriguing target for future studies.

We constructed a comprehensive human ISC signature. SI ISCs had 68 DEGs compared to other clusters across SI regions for all three donors, whereas colon ISCs displayed 109 DEGs (Table S1, Fig. S7). To define a human ISC transcriptional signature spanning SI and colon, we identified 46 DEGs enriched in ISCs from both organs (Table S2). This signature includes classical ISC markers along with 30 ribosomal genes. While ribosomal genes are less abundant in murine ISC signatures^{26,28,29}, this is consistent with literature describing ribosomal control of transcriptional dynamics in other stem cells³⁰⁻³². To identify ISC DEGs conserved between human and mouse, we compared our 68-gene SI ISC signature with a mouse signature with 344 human homologs²⁶. Surprisingly, only 11 genes overlapped between the human and mouse signatures (Fig. 2C), although it is unclear whether this reflects species differences or the higher resolution and stringency of our computational approach. Conserved genes included classical markers: *LGR5*, *OLFM4*, *ASCL2*, *RGMB*, *SLC12A2*, and *MYC*; genes with known ISC function: *RNF43*, *ZBTB38*, *VDR*, and *CDK6*; and one gene not described in ISC literature: *TRIM24*, a gene involved in p53 degradation. These SI, colon, and full-gut ISC signatures underline key similarities and differences in proximal-distal human ISCs.

TA cells are classically defined by high proliferation and location above ISCs in the intestinal crypt^{9,33}. Leiden clustering separated SI TA cells undergoing S/G2 cell-cycle phases (TA) and M-phase (TA2) (File S1,S2). We evaluated DEGs between SI TA, TA2, and colon TA cells to define shared TA markers (Fig. 2D). DEGs were involved in cell cycle as well as in mitochondrial biogenesis and rRNA processing, consistent with increased mitochondrial load and translation levels seen as stem cells differentiate in various systems^{30,34-36}. Several organ-specific markers of differentiated lineages (Fig. S8) were unexpectedly enriched in their respective SI or colon ISC and TA populations (Fig. 2E), hinting that ISCs are already transcriptionally primed for organ-specificity, instead of existing in a naïve, pan-intestinal state. This is consistent with rodent studies showing adult SI ISCs produce daughter cells specific to their originating organ when engrafted into alternative SI or colon sites^{37,38}. Studies defining region/organ-specific chromatin or active transcriptomic differences in

ISCs were not found; thus, these genes may prove useful for studying differentiation and chromatin dynamics in early fate-determination.

Trajectory analyses computationally investigate lineage transitions, with intestinal analyses primarily using mouse data³⁹⁻⁴². We used Partition-based Graph Abstraction (PAGA), which estimates connectivity between clusters, to infer connections between proliferative crypt-based lineages and differentiated populations²². Absorptive enterocytes (AEs) and ACCs arise nearly exclusively from ISCs and TA cells¹⁵. The secretory progenitor population arises from ISCs and TA cells and gives rise to PCs, GCs, and EECs (Fig. 2F,G). Interestingly, tuft cells appear to derive from secretory progenitors in colon but not SI, consistent with murine findings³⁹. Conversely, SI BEST4⁺ cells apparently arise from secretory progenitors while colon BEST4⁺ cells connect to TA cells, agreeing with absorptive and secretory functions ascribed to this novel lineage^{12,16}.

Predicted regional cell cycle phase distributions⁴³ were analyzed in ISCs, TA cells, and secretory progenitors (Fig. 2H-J). ISCs showed expected high G1 and S phase representation across regions^{13,44}, while highly-proliferative TA cells largely existed in S and G2/M. TC showed a significantly lower proportion of TA cells in G2/M than jejunum, with similar regional differences in rodents^{45,46}, but biological implications are unknown. Secretory progenitors had notably different distributions, with S phase proportions increasing proximally-to-distally and higher G1 proportion than TA cells (Fig. S1). Since secretory progenitors differentiate into specialized lineages, elongated G1 may allow for additional reception of differentiation factors, as stem cells are most receptive to such cues during G1⁴⁷.

Paneth Cells

Murine PCs play important niche-supporting and antimicrobial roles⁴⁸, yet little scRNAseq analysis covers human PCs. Our dataset includes 49 PCs representing all SI regions across three donors, 10-times more than analyzed in recent literature¹⁶. PCs were defined using *DEFA5*, *DEFA6*, *ITLN2*, and *PLA2G2A* (Fig. 3A). Surprisingly, Lysozyme (*LYZ*), an important murine PC marker, was expressed higher in SI BEST4⁺ cells and FAE, with measurable tuft cell expression, making *LYZ* an imprecise human PC marker (Fig. 3B). This is consistent with *LYZ* expression in organoids derived from a fetal human stage too young to form PCs⁶. Since PCs cluster alongside GCs and share *LYZ* expression with BEST4⁺ cells, classical PC markers were plotted to confirm PC identify (Fig. 3C). Importantly, our data indicate the cells described as PCs in a recent scRNAseq publication¹³ are BEST4⁺ cells, with high *LYZ*, *SPIB*, *BEST4*, and *CA7*. Similarly, the reported colonic 'Paneth-Like Cells' in the study are likely BEST4⁺ cells unrelated to PCs other than *LYZ* expression. The rarity of PCs (<1% of our dataset and others¹⁶) and divergence of PC and BEST4 markers highlights the precise lineage attribution needed when defining human PCs.

We next probed whether human PCs express ISC niche factors. Murine PCs express *Wnt3*, *Wnt11*, *Tgfa*, *Egf*, *Dll1*, *Rspo1*, and *Dll4*^{40,48-50}, but one report shows human PCs express no *WNT3/11*¹⁶. Our data confirm this and demonstrate no measurable *EGF* or *RSPO1* and minimal *TGFA*. While human PCs express *DLL1* and *DLL4*, both are higher in secretory progenitors (Fig. 3D). We found no members of any major intestinal growth factor family broadly enriched in PCs (Fig. S9), suggesting human PCs are not major niche-supporting cells. This notion is consistent with non-epithelial sources of WNTs and growth factors in the human niche^{6,51-53}, and echoes mouse biology, where PCs are sufficient to support ISCs⁴⁸ yet unnecessary for niche maintenance^{48,54-56}.

Unexpectedly, *SMOC2*, a murine ISC marker,²⁶ was expressed highest in PCs, with other mouse ISC-restricted markers (*LGR5*, *ASCL2*, *RGMB*) higher in human PCs than mouse PCs⁵ (Fig. 3E). ISC-PC doublet artifacts could explain this, however lack of ISC markers (e.g., *OLFM4*) does not support this hypothesis. *LGR5*, *SMOC2*, and *ASCL2* are involved in WNT reception^{23,57-59}, suggesting human PCs may receive WNT signals instead of providing WNT signals as in mice⁴⁸. PC-unique expression of *FZD9* supports a WNT-receptive PC role⁶⁰ (Fig. S9), with its expected ligand *WNT2*⁶¹ absent in our database yet induced in intestinal inflammation and cancer⁶²⁻⁶⁴. Mature PCs can dedifferentiate following injury in mice⁶⁵⁻⁶⁷, and *Ascl2* is required for dedifferentiation in mouse crypts⁶⁸. Thus, expression of these ISC genes may render human PCs more responsive to dedifferentiation cues; however, this needs functional validation.

A murine subset of colonic GCs termed Paneth-Like Cells or Deep-Crypt Secretory Cells are defined by MUC2, C-KIT, REG4, CD24, EGF, and FZD5^{69,70}. We found only two cells positive for both *REG4* and *KIT* across the 1,252 colon GCs and secretory progenitors (Fig. S9D). Colonic tuft cells expressed higher *KIT* and *CD24* but no *REG4*. Based on murine-defined markers, we conclude there is no human equivalent to this population. Since human PCs likely perform little niche-supporting activity, namely not producing *EGF*, human cells with murine PC-like functions may be unnecessary.

Despite striking differences between mouse and human PCs, both supply antimicrobial gene products. Six of the 10 highest-expressed SI antimicrobial peptides are PC-enriched (Fig. 3F). As antimicrobial genes comprise half of human PC DEGs (Table S1), they likely function primarily to protect the ISC-niche from bacterial invasion and regulate microbiota composition⁷¹.

Follicle-Associated Epithelium

Rare FAE cells, important for epithelial-immune crosstalk, reside in small puncta throughout the intestines⁷². FAE includes microfold (M)-cells, which transport luminal antigens to immune cells residing within their microfolds⁷³. M-cells have almost exclusively been explored in mice⁷⁴⁻⁷⁶ or using directed differentiation in vitro⁷⁷, with only one study reporting scRNAseq data from healthy human intestinal M-cells¹⁷. Our dataset includes a cluster of 19 cells from Donor 2 enriched for M-cell markers⁷⁸⁻⁸⁰ and immune cross-talk genes (Fig. S10). Strikingly, several murine M-cell-specific markers were either widely expressed (*MARCKSL1*, *ANXA5*, *CXCL16*)^{79,81} or absent (*CCL36*, *SCG5*, *TNFRSF9*, *CCL9*, *CCL6*, *PGLYRP1*)^{73,79,82}, suggesting species functional differences. With the caveat of including 19 cells from one donor, we defined 145 DEGs (Table S1,S2), finding many FAE-unique genes (Fig. S10). Pathway enrichment analysis implicates these DEGs in immune cell interactions, verifying expected M-cell function (File S3). Differences in chemokines between mouse and human M-cells, with *CCL20* the only major shared chemokine, call for enriching for FAE using newly described methods⁷² to probe how human M-cells interact with immune cells.

BEST4⁺ cells

Recent human single-cell studies describe a novel intestinal lineage, absent in mice, expressing high *BEST4*, *SPIB*, and *CAT*¹², with *CFTR* in SI^{16,17} and *OTOP2* in colon. Several papers describe colonic BEST4⁺ cells^{11,12}, so we analyze SI BEST4⁺ cells. With BEST4⁺ cell functions largely unknown, DEGs were used to predict physiological roles (Fig. 3G). DEG analysis revealed secreted peptides including *GUCA2A* and *GUCA2B*¹², which can act as pro-hormones effecting a satiety response⁸³⁻⁸⁵. A previous report showed these genes in SI and colon BEST4⁺ cells¹⁶, yet we find both expressed higher in SI than colon BEST4⁺ cells (Fig. 3H), indicating a role in satiety signaling in the SI.

We identified two unreported secreted peptides, *NPY* and *BMP3*, specifically in SI BEST4⁺ cells. *NPY* expression was unexpected in intestinal epithelium⁸⁶, and gut *BMP3* is largely studied for antitumor roles^{87,88}. Diffusion pseudotime, a computational method to define single-cell differentiation trajectories⁸⁹, indicated that *NPY*, *GUCA2A*, and *GUCA2B* expression increased with BEST4⁺ cell maturation, while *BMP3* expression appeared independent of maturation (Fig. 3I). Interestingly, receptors for all four genes are enriched in EECs, suggesting cross-talk between these lineages (Fig. 3H).

Since *NPY* is proposed to affect gastrointestinal (GI) motility⁹⁰ and energy homeostasis⁹¹, we probed if *NPY* correlated with genes induced following meals. Looking across SI regions for each individual donor, we found a strong positive correlation between SI BEST4⁺ cell *NPY* and AE expression of *SI* (R=0.82) and *APOA4* (R=0.86), which are induced in mice by dietary sugar⁹² and fat⁹³ (Fig. 3J). We found further positive correlations with AE genes involved in dietary metabolism (Table S3), and negligible correlation with housekeeping genes *ACTB* (R=-0.22) or *GAPDH* (R=-0.08). Thus, *NPY* expression in SI BEST4⁺ cells may be induced by luminal contents. Further DEGs from SI BEST4⁺ cells included adrenergic and cholinergic neurotransmitter receptors involved in intestinal motility⁹⁴, *ADRA2A* and *CHRM3* (Table S1), supporting that SI BEST4⁺ cells may regulate intestinal motility following meals.

BEST4⁺ cells likely absorb dietary heavy metals. Metallothionein expression has been implicated in colonic BEST4⁺ cells¹², yet we find seven members of this family, known to bind heavy metals and protect

against toxicity⁹⁵⁻⁹⁷, specifically enriched in SI BEST4⁺ cells (Fig 3K). Unique expression of *STEAP2*, a metalloredutase for copper and iron⁹⁸, supports a role for SI BEST4⁺ cells in maintaining SI homeostasis for many metal ions⁹⁷⁻⁹⁹. SI BEST4⁺ cells express endocytosis effector genes *DNM1* and *DNM2* (Fig. 3K), supporting an absorptive role for these cells. Our data indicate BEST4⁺ cells perform diverse roles within the intestinal epithelium, laying the groundwork for functional studies.

Tuft Cells

Tuft cells are chemosensory epithelial cells which regulate type-2 immune reactions in the intestinal epithelium through pathogenic metabolite detection and classical taste signal transduction pathways¹⁰⁰⁻¹⁰². Together, these initiate tuft cell IL-25 release¹⁰³⁻¹⁰⁵. SI and colon tuft cells share many classical markers^{16,106} (Fig. 4A, Table S1,S2). *DCLK1*, a key murine marker¹⁰⁵, was not observed. Interestingly, *SUCNR1*, a G-protein coupled receptor mediating SI IL-25 release¹⁰⁷, was negligible in colon, suggesting SI and colon tuft cells differentially detect and respond to luminal succinate-producing pathogens (Fig. 4B,C). Instead, colonic tuft cells likely respond to umami-chemosensory cues, such as microbe-derived free amino acids, as they express heterodimeric umami taste receptor subunits *TAS1R1* and *TAS1R3* (Fig. 4B,C)¹⁰⁸. We further analyzed downstream taste signal transduction genes, finding SI and colon tuft cells enriched in critical pathway components (*GNB1*, *GNG13*, *ITPR2*, *TRPM5*)^{100,101}, with SI-specific *GNAT3*, a G-protein alpha subunit, which likely activates *PDE4D* to decrease intracellular cAMP/cGMP¹⁰⁹ (Fig. 4B). This suggests human SI tuft cells have varied responses to succinate-producing microbes (e.g., *N. brasiliensis*), whereas colonic tuft cells may respond more broadly to other microbial taxa.

Beyond triggering type-2 immunity, tuft cell DEGs allow broad interaction with the adaptive and innate immune systems. Tuft cells express DEGs involving ubiquitin-mediated proteasome degradation, including SCF complex components (*SKP1*, *CUL3*, *FBXO32*, *RBX1*) which initiate processing of exogenous antigens for presentation^{110,111}, as well as genes for the MHC1 antigen presentation complex (Fig. 4D). This suggests tuft cells may interact with the adaptive immune system following luminal stimuli. Human tuft cells also uniquely express previously unappreciated Toll-Like Receptors –*TLR9*, *TLR5*, and *TLR4* – which bind bacterial and viral DNA, bacterial flagellin, and lipopolysaccharide (LPS), respectively¹¹²⁻¹¹⁴ (Fig. 4D). Expression of the LPS coreceptor *CD14* across tuft cells (Fig. 4D) supports a novel role in bacterial-related immune responses¹¹².

Tuft cells exhibit possible auto-regulatory mechanisms for these pathogen-response pathways. Tuft cells express heterodimeric IL-25-specific receptor components *IL17RA* and *IL17RB* (Fig. 4D) which may create a positive feedback loop to amplify IL-25 signaling, as in keratinocytes¹¹⁵. Tuft cells may also autoregulate their LPS response through *SIGIRR*, which negatively regulates TLR4-LPS signaling¹¹⁶⁻¹¹⁸. These implicate tuft cells as dynamic sentinels linking luminal contents to the immune system.

Consistent with a role regulating gut pathogens, tuft cells produce *LYZ*, *PRSS3*, and *DEFB1* antimicrobial peptides in the SI (Fig. 3F) and six of the top ten antimicrobial peptides in the colon, without PCs present (Fig. 4E). Finally, human and murine tuft cells both produce neuro- and immunomodulatory compounds. We find genes necessary for acetylcholine synthesis (*CHAT*, *SLC18A3/VACHT*), communication with neurons (*NCAM1*)¹¹⁹, and enzymes involved in eicosanoid production, namely cysteinyl leukotrienes (*ALOX5* and *ALOX5AP*) and Prostaglandin D₂ (*PTGS1* and *HPGDS*), which broadly regulate inflammation¹²⁰ (Fig. 4F). Altogether, these genetic analyses indicate tuft cells regulate luminal microbes, communicate with the nervous system, and effect systemic immune responses.

Goblet Cells

GCs produce membrane-bound and secreted mucin glycoproteins that lubricate the gut, act in signaling, support commensal bacteria, and form the protective mucus barrier^{103,121-123}. DEGs include classical markers *CLCA1*, *MUC2*, and *TFF3*, with colon GCs expressing higher mucins (*MUC5B*, *MUC4*, *MUC1*) and the antiprotease *WFDC2*¹² (Fig. 5A). Pathway enrichment analysis of DEGs confirms GCs principally act in mucus secretion, with 15 of 20 top enriched pathways involved in mucus production, including glycosylation, Golgi/ER vesicle transport, and unfolded protein response (File S3)¹²⁴⁻¹²⁷. Thus, we mapped regional GC mucin expression (Fig. 5B), finding secreted *MUC2* and transmembrane *MUC13* expressed across all regions and colon-enriched *MUC1*, *MUC4*, and *MUC5B*. While GCs are the major mucus-secreting intestinal lineage, AE

and ACC express transmembrane mucins that form a glycocalyx to protect against pathogenic bacteria^{128,129}. Regional mucin expression across these lineages showed high *MUC13*, *MUC17*, and *MUC3A* in AEs and several enriched in ACCs (Fig. S11), informing studies regarding mucus composition and function across GCs, AEs, and ACCs.

GCs are commonly considered fairly homogenous, yet recent work in mouse colon reported transcriptional signatures of early GCs, crypt-resident (crGC), and inter-crypt goblet cells (icGC), with icGCs producing more permeable mucus than crGCs¹³⁰. Human colonic secretory progenitors and GCs subclustered into similar groups marked by genes implicated in mouse GC heterogeneity: early GCs (*STMN1*, *HMGB2*, *MKI67*), crGCs (*PDIA3*, *AGR2*, *PDIA5*), and icGCs (*MXD1*, *RAB27A*, *FER1L6*) (Fig. 5C,F). Some canonical GC markers of mucus secretion (*MUC2*, *ZG16*) were expressed highest in icGCs, consistent with icGCs constitutively secreting mucus¹³¹. Diffusion pseudotime confirmed increasing maturity across these subpopulations (Fig. 5D-E). Notably, crGCs expressed higher *MUC5B* while icGCs expressed higher *MUC2*, *MUC13*, *MUC1*, and *MUC4* (Fig. S11C), consistent with distinct mucus production in human icGCs shown via lectin staining¹³⁰. Similar sub-clusters were observed in the SI, although mucin differences were less obvious (Fig. S11D-F). This demonstrates human GCs are more heterogeneous than appreciated, necessitating studies to determine functional differences.

Enteroendocrine Cells

EECs secrete hormones to communicate between the intestine and the body. EEC hormone expression profiles are well characterized at the single-cell level in mice, since EEC reporter models enable enrichment of this rare lineage (<1% of intestinal epithelium)^{132,133}. However, transcriptomic differences exist between mouse and human EECs^{132,134}. An approach in human organoids with an EEC reporter gene yielded sufficient EECs for detailed scRNAseq analysis, though it is unclear how these cells may differ from primary human EECs¹³². While several human scRNAseq studies show EEC data^{11-13,16,17}, our 154 EECs (Fig. 5G) represent the largest dataset of primary human EECs to our knowledge.

First, regional expression of hormones and other signaling machinery in EECs was surveyed. We found *SCT*, *CHGA*, *TPH1*, and *DDC* span regions with a SI bias and *GCG* and *PYY* span regions with a colonic bias (Fig. 5H). *GAST* and *GIP* express in proximal SI, while *TAC1*, *CCK*, *GHRL*, *MLN*, *SST*, and *NTS* express from duodenum through AC/TC, and *INSL5* is colon-specific. Rare *NPY* expression was detected in jejunum and AC. These results expand on an early study using immunohistochemistry in regional biopsies to find *CCK*, *GAST*, *GIP*, *NTS*, *MLN*, and *SCT* segregated to SI regions¹³⁵. Our data confirm the SI bias but show low colonic expression of *CCK*, *NTS*, *MLN*, and *SCT*, demonstrating the higher sensitivity in scRNAseq. *NTS* and *CCK* were also absent in a study analyzing region-unspecified colon¹³, showing the importance of analyzing all colon regions. Fatty-acid receptors *FFAR1* and *FFAR2* were enriched in SI EECs and *FFAR4* expression was colon-specific (Fig. 5I). The amino acid transporter *SLC38A2* was well-expressed, with lower expression of other amino acid transporters. EECs also express several hormone receptors, indicating crosstalk amongst EECs. An additional form of gut-brain crosstalk was recently discovered in mice, with EECs found to form synapses with the vagus nerve¹³⁶⁻¹³⁸. DEGs from SI and colon EECs (Fig S1,S12A) are consistent with a human equivalent of these mouse EECs, termed neuropods¹³⁷, with 33.7% of genes in the GOCC_Presynapse list expressed highest in EECs (Fig. S12B). Altogether, these patterns support roles for EECs in crosstalk within the gut and between the gut and brain, further illuminating their functional importance.

EECs are classified into subtypes based on hormone expression^{139,140}. A regional breakdown of individual EECs was constructed to visualize EEC subtypes across the gut (Fig. 5J). Enterochromaffin cells (*TPH1*, *DDC*, *CHGA*) appear in each region with fewer in colon, and ileum unexpectedly lacks L cells (*GCG*, *PYY*). Multiple EECs express 8-10 hormones, expanding on studies identifying poly-hormonal EECs^{141,142}. Despite low numbers, *GAST* and *GIP* largely segregate from duodenal L cells yet overlap in jejunum. We note rare *NPY* expression in *MLN*⁺ and *GHRL*⁺ EECs in jejunum and AC. Future studies combining our primary EECs with additional datasets from various regions would improve our understanding of human EECs.

Absorptive Enterocytes and Colonocytes

AEs and ACCs perform nearly all intestinal absorption¹⁴³. These lineages formed multiple Leiden clusters, with three AE clusters and two ACC clusters consistent with increasing maturity, reflecting other reports^{11,16}, and one cluster (AE2) separate from other AEs, largely from Donor 3 ileum (Fig. 1C,6B). A common DEG signature was defined by grouping all AEs and comparing their DEGs with all ACC DEGs. Surprisingly, only five DEGs were shared between organs (Fig. 6A), indicating stark organ differences. Interestingly, the novel AE2 population expressed mature AE markers (Fig. 6B), while uniquely expressing genes for bile acid uptake and processing¹⁴⁴⁻¹⁴⁶ (Fig. 6C). It is unclear why ileal AEs of Donor 3 clustered separately from other donors. Possible donor-specific factors contributing to this novel cluster include unique demographics (lowest BMI, African American, Type II diabetic) or differences in luminal contents between donors possibly inducing unique gene expression patterns, as described for certain genes^{92,93}. These donor-specific differences in transcriptomic signatures highlight the need for data from wide, diverse populations.

Macro- and micro-nutrient handling was mapped across all AEs and ACCs (Fig. 6D). Nearly all fatty acid, glucose, and cholesterol transporters enriched in SI, consistent with recent work¹³, but regional data revealed an undescribed trend of increasing expression from duodenum through ileum for most genes (Fig. 6D). Two notable exceptions were *FABP1* and *FABP5*, both with higher colon expression. *FABP1* is expected in the proximal SI¹⁴⁷ and *FABP5* is not implicated in healthy ACCs to our knowledge, suggesting a possible role for colonic fatty acid uptake. Digestive enzymes exhibited ileal enrichment except for the duodenum-specific serine protease *TMPRSS15/Enteropeptidase*¹⁴⁸. More regional variability was seen for amino acid transporters, with different genes peaking from duodenum through colon. The neutral amino acid transporter *SLC38A1* expressed in colon, in contrast to a study using colon biopsies¹⁴⁹. Ion transporters showed the most regional differences, with *SLC25A3* and *SLC4A4* spanning all regions, colon-enriched *SLC26A2*, and SI-enriched *SLC9A3R1*. Finally, *SCNN1* sodium transporter subunits were largely enriched in colon, consistent with its need for regulated water uptake. This regional map of nutrient handling genes expands upon previous organ-level analyses, emphasizing the importance of the ileum in digestion.

Intestinal barrier function, largely conferred by cell-junction proteins, is essential for regulated absorption and physical antimicrobial defense¹⁵⁰. The 20 highest-expressed cell junction genes were regionally evaluated (Fig. 6E). Several junction genes expressed equally across AEs and ACCs, while others exhibited regional enrichment. Claudins (*CLDN*) are the primary determinants of tight junction barrier function and intestinal epithelial integrity^{150,151}. *CLDN1* and *CLDN15* were SI-enriched and *CLDN3*, *CLDN4*, and *CLDN7* were highest in TC. Notably, no junction genes expressed highest in DC. This is intriguing, as ulcerative colitis often originates in the distal large intestine, raising the possibility that higher junction protein expression in AC and TC might protect against certain inflammatory conditions¹⁵²⁻¹⁵⁴. While our data cannot show pathological implications, the possibility that this differential expression pattern may protect against colitis in AC and TC is intriguing.

Aquaporins (AQPs) are the major transcellular transporters of water and small solutes in the intestine¹⁵⁵. Our data confirms a previous report showing elevated *AQP3*, *AQP7*, and *AQP11* in ileum relative to colon and *AQP8* elevated in colon (Fig. 6F)¹³, yet we find *AQP1* widely expressed. Aquaglyceroporins (*AQP3*, *AQP7*, *AQP10*) increase from duodenum to ileum, coincident with increasing lipid metabolism genes (Fig. 6D). Viewing *AQP* across lineages (Fig. 6G) shows enriched aquaglyceroporins in mature AEs, defining a likely role for AQP-mediated glycerol transfer in AE triglyceride processing. We note unappreciated specificity of *AQP1* expression in ISCs and TA cells across organs and uniquely restricted *AQP8* expression in late ACCs in the AC. Diffusion pseudotime demonstrates that *AQP8* is expressed in the most mature late ACCs, likely on the surface epithelium (Fig. 6H). Distinct expression of *AQP1* and *AQP8* at the crypt base and surface, respectively, suggests specific physiological roles that should be functionally interrogated.

Receptors/Drugs

We finish by examining how extrinsic signals may affect the intestinal epithelium. Two approaches were designed to show how associations with receptors, drug targets, and lineage states can be revealed. First, expression plots were compiled of major receptor families across lineages, separated into high-, intermediate-, and low-expressing genes for easier visualization (File S5, Table S4). The five highest-expressing genes from

each family were then grouped to visualize expression across lineages (Fig. 7A). Several patterns appear from these 60 receptors: 20 receptors are highest in tuft cells, 11 in EECs, 10 in AE/ACC, 9 in FAE; receptors bias towards villi (12) vs crypts (3); receptors bias toward SI (4) vs colon (0); and many uniquely express in certain lineages (12 in tuft, 3 in EECs, etc.), showing potential ways regions and lineages might be targeted by exogenous signals.

To test the novelty of these observations, we probed receptors with unique expression in tuft cells and searched for literature connecting these genes to tuft cells in mice or humans. We found direct connections to intestinal tuft cells for only five of the 12 (*TRPM5*, *ITPR2*, *HTR3E*, *IL13RA1*, *IL17RB*), with no connection to intestinal tuft cells found for the remaining seven (*GABRA4*, *ADGRG6/GPR126*, *SIGIRR*, *ITGB5*, *KIT*, *PTPRJ*, *TLR9*). These seven newly-defined lineage-specific receptors arose from just the five highest-expressed receptors from each family, and our full dataset includes 669 total receptors (Fig. 7B, File S5). This receptor expression atlas across lineages, organs, regions, and donors provides a powerful foundation to explore how extrinsic signals from local epithelial microenvironments, dietary and microbial influences, and pharmaceuticals affect intestinal epithelial lineages.

We explored how medicines might directly affect the intestinal epithelium. While many drugs have GI-related side effects, few are intended to target the intestinal epithelium itself¹⁵⁶⁻¹⁵⁹, and side effects such as diarrhea are often unexplained at the cell-lineage level¹⁶⁰. We searched for primary targets of all FDA-approved drugs and found 498 approved drugs had 232 primary gene targets expressed in our gut epithelial dataset (Fig. 7C, Table S5). As most of these drugs are prescribed for non-GI diseases, primary targets within the intestinal epithelium are likely often overlooked and may contribute to unexpected GI side effects.

While many drugs are metabolized by the liver, oral drugs can be altered by metabolism within the gut epithelium¹⁵⁷⁻¹⁵⁹. We show expression of genes for Phase I and Phase II drug metabolism proteins by lineage with highest expression in the intestinal epithelium (Fig. 7C) and quantified gene expression by lineage and by region (Table S6). *CES2*, which metabolizes the cancer drug irinotecan into its biologically active form SN-38¹⁶¹, is found to be the highest-expressed Phase I metabolism gene in the SI, with AE enrichment. Interestingly, *UGT1A1*, the Phase II gene which inactivates SN-38¹⁶², has low gut epithelial expression (Table S6). This suggests irinotecan might remain active in the gut, supporting the idea that orally administered irinotecan might be effective against intestinal cancers¹⁶³⁻¹⁶⁵. Our easily-searchable dataset provides expression values for genes important for further studies probing intestinal metabolism of endobiotics, environmental toxicants, and pharmaceuticals.

As an example of a disease-focused approach, we searched for primary targets of drugs prescribed for inflammatory bowel disease (IBD). Most are anti-inflammatory or immunomodulatory, so primary targets are often not expressed in the intestinal epithelium. Our database shows nine primary gene targets of eight IBD drugs with epithelial expression (Fig. 7D). As many of these drugs are intended to affect immune cells, we mapped epithelial expression of their primary target genes to see which lineages might be inadvertently affected (Fig. 7E). We find high *FKBP1A*, a tacrolimus (Prograf[®]) target, in the little-understood BEST4⁺ cells, making this important to follow as their functions become better understood. Mycophenolate mofetil (CellCept[®]) targets *IMPDH2* and *IMPDH1* expressed in proliferative crypt populations and EECs, respectively. The methotrexate target *DHFR* is highest in TA and progenitor cells, while the tofacitinib (Embrel[®]) target *JAK1* has broader expression. These drugs can be orally administered and have side effects including diarrhea, nausea, vomiting, or appetite loss¹⁶⁶⁻¹⁶⁸. Interrogating this small subset of drugs in our database highlights a spectrum of potential unintended epithelial effects on renewal by ISC or TA cells, EEC hormonal influences on appetite and gut motility, and unknown effects from other lineages.

Personalized precision medicine is an emerging field motivating new technologies¹⁶⁹. We used our drug-target atlas to evaluate regional variability of tacrolimus, mycophenolate mofetil, and tofacitinib target genes, *FKBP1A*, *IMPDH2*, and *JAK1*, across individual donors as an approach to inform potential patient-dependent effects (Fig. 7F). We find higher colonic expression of all three targets, suggesting that patients may experience effects of drugs targeting these genes in their lower GI tract. Comparing multiple donors may hint at susceptibility to drug side-effects, with Donors 2 and 3 generally expressing target genes higher than Donor 1.

While our donor number is insufficient for statistically significant conclusions, it provides a framework to generate observations to inform larger studies. We hope our lineage-, regional-, and donor-specific data on primary drug targets will be used across gastroenterology and pharmacology to better understand how drugs may affect the intestinal epithelium.

Discussion

We use transplant organ tissue to characterize intestinal epithelial cells from duodenum through DC for three adult human donors, allowing us to define comprehensive lineage markers and map regional functionality. Our experimental design has many strengths not found in other studies. DNA-oligo hashtag antibodies allowed for all six regions from each donor to be sequenced together to save cost and avoid intra-donor batch effects, then separated computationally to analyze regional differences. This allowed for comparing gene expression between multiple regions within SI or colon. The regional differences found highlight the importance of regional selection when studying the gut, yet many colonic scRNAseq studies do not specify sample region or mention if pooled samples are from consistent regions. We analyzed cells from donors with no known intestinal diseases or cancer, avoiding effects on non-inflamed adjacent tissues in patients with inflammatory disease⁹.

Our experimental design allowed us to study many specific phenomena. Analyzing cells across six regions allowed us to compile comprehensive transcriptional signatures of genes significantly enriched in each lineage versus other lineages in all six regions and across donors. We map mucins, hormones, transporters, and barrier function genes across all six regions. We analyze PCs, showing drastic differences in growth factor expression from mouse literature and highlighting the insufficiency of *LYZ* for marking human PCs. We also analyze rare FAE cells, showing mouse/human differences and defining DEGs. We use PAGA to infer the differentiation trajectory for each lineage and suggest organ-specific maturation for tuft and BEST4⁺ cells. We propose novel tuft cell interactions with pathogens and the immune system. Finally, we map receptor expression and primary drug targets across lineages and highlight the ease of using this database to find previously undescribed gene expression. We hope our database serves as a resource to understand how drugs affect the intestinal epithelium and as guidance for future precision medicine approaches.

SI and colon from three male donors varying in age, race, and BMI were used in our study. Optimally, future studies will add regional data from diverse donors to build upon this foundation. While multiplexing six regions in one library prep allowed us to make a wide-ranging atlas, this approach results in limited numbers of each lineage analyzed per region. Our trajectory analysis suggests which lineages mature through secretory progenitors, yet future studies including more progenitors may provide more sensitive and high-resolution analyses.

In this study, we provide a comprehensive cell-level transcriptomic view of the SI and colon epithelium with regional resolution across multiple humans. Our analyses independently confirm and advance prior studies, define important differences between mouse and human lineages, and highlight how lineages change along the proximal-distal axis. We include easy-to-search tables for DEGs, receptors, and drug targets that can be interrogated by most investigators and trainees. Overall, our database provides a foundation for understanding individual contributions of diverse epithelial cells across the length of the human intestine and colon to maintaining physiological function.

Figure Legends

Figure 1: Sample Processing. **A)** Schematic for isolating cells from six intestinal regions for three donors and using hashtag antibodies to sequence regions side-by-side. **B)** UMAP of analyzed cells in 25 lineage clusters. **C-E)** UMAP of all cells by donor (C) or region (D,E). **F-G)** Heatmaps showing unique markers for major lineages in SI (F) and colon (G).

Figure 2: Proliferative crypt populations. **A)** Heatmap of DEGs in ISCs vs. other lineages (top; red: classical markers), SI vs. colon ISCs (middle), colon vs. SI ISCs (bottom). **B)** UMAP of *LGR5*, *OLFM4*, and *RARRES2* expression. **C)** Venn Diagram showing overlap between murine and human ISC signatures. **D)** Heatmap of

DEGs in TA cells vs. other lineages (top), SI vs. colon TA cells (middle), colon vs. SI TA cells (bottom). **E**) Dotplot showing DEGs from SI or colon-specific lineages within organ-delineated ISCs and TA cells. **F,G**) Partition-based graph abstraction (PAGA) showing connectivity between major lineages in SI (F) and colon (G) to infer maturation trajectory. Line thickness represents connectivity strength. **H-J**) Regional cell cycle phase distribution in ISCs (H), TA cells (I), and secretory progenitors (J).

Figure 3: Paneth and BEST4⁺ cells. **A**) Heatmap of DEGs in PCs vs. other lineages (red: classical markers). **B**) Dotplot showing Lysozyme across FAE, BEST4, Paneth, and tuft cells. **C**) Dotplot showing PC, goblet, and BEST4 markers. **D**) Dotplot showing growth factors implicated in murine PCs across SI lineages. **E**) Heatmap showing PC and ISC markers across both lineages. **F**) Dotplot showing 10 highest-expressed antimicrobial peptides across SI lineages. **G**) Heatmap of DEGs in BEST4⁺ cells vs other lineages (top; red: classical markers), SI vs. colon BEST4⁺ cells (middle), colon vs. SI BEST4⁺ cells (bottom). **H**) Dotplot showing secreted genes and their receptors across lineages. **I**) UMAPs of BEST4⁺ cells showing predicted diffusion pseudotime and expression of secreted peptides. **J**) Expression of *NPY*, *SI*, and *APOA1* across regions by donor. **K**) Dotplot showing genes involved in metal-binding and endocytosis across lineages.

Figure 4: Tuft Cells. **A**) Heatmap of DEGs in tuft cells vs. other lineages (top; red: classical markers), SI vs. colon tuft cells (middle), colon vs. SI tuft cells (bottom). **B**) Dotplot showing tuft cell enrichment of genes specific to taste signal transduction. **C**) Organ-specific signal transduction in SI vs. colon tuft cells. **D**) Dotplot showing tuft cell-enriched genes enabling interactions with innate and adaptive immune system. **E**) Dotplot showing 10 highest-expressed antimicrobial peptides across colon lineages. **F**) Dotplot showing tuft cell-specific genes for producing acetylcholine, eicosanoids, and prostaglandins.

Figure 5: Goblet and Enteroendocrine Cells. **A**) Heatmap of DEGs in GCs vs. other lineages (top; red: classical markers), SI vs. colon GCs (middle), colon vs. SI GCs (bottom). **B**) Dotplot showing 9 highest-expressed mucins across GCs by region (blue: gel-forming mucins). **C**) Leiden sub-clustering of colon GCs. **D**) Diffusion pseudotime of colon GCs. **E**) UMAP of *MUC2* expression in colon GCs. **F**) Dotplot showing mouse-implicated markers in human colon GC subclusters. **G**) Heatmap of DEGs in EECs vs. other lineages (top, red: classical markers), SI vs. colon EECs (middle), colon vs. SI EECs (bottom). **H**) Dotplot of EEC regional hormone gene expression. **I**) Dotplot of EEC expression of select receptors by region. **J**) Heatmap showing hormone expression in each EEC.

Figure 6: Absorptive cells. **A**) Heatmap of DEGs in absorptive cells vs. other lineages (top) AEs vs. ACCs (middle), ACCs vs. SI AEs (bottom). **B**) UMAPs showing AE2 Leiden cluster (top) and cells by region (bottom). **C**) Dotplot of classical Mature AE markers and top 10 DEGs for AE2 cluster. **D**) Dotplots showing regional expression of genes involved in digestion and absorption in all AEs and ACCs. **E**) Dotplots showing 20 highest-expressed cell junction genes in AEs and ACCs by region. **F**) Dotplots showing regional aquaporin expression in AEs and ACCs. **G**) Aquaporin expression across lineages. **H**) UMAPs of late ACCs showing predicted diffusion pseudotime (left) and *AQP8* expression (right).

Figure 7: Extrinsic receptors and drug targets. **A**) Dotplot showing five highest-expressing members of major receptor families by lineage. **B**) Pie chart showing receptor genes expressed in the intestinal epithelium by lineage with highest expression. **C**) Pie chart showing primary targets of all approved drugs and Phase I and Phase II drug metabolism genes expressed in the intestinal epithelium by lineage with highest expression. **D-E**) Primary targets of drugs used to treat IBD by expression across lineages. Note scaling changes. **F**) Dotplot showing expression of three targets of IBD drugs across regions split by donor.

Supplemental Figures and Files

Figure S1: Patient characteristics and cell counts. **A**) Donor information. **B**) Cells collected per donor region. **C**) Small intestinal lineages collected per donor. **D**) Colonic lineages collected per donor. **E**) Small intestinal lineages per donor region. **F**) Colonic lineages per donor region.

Figure S2: Tissue histology. Hemotoxylin and eosin stained tissues from each region for all three donors. All scale bars = 200 μ m

Figure S3: FACS strategy. FACS strategy for gating out cell fragments, likely doublets, and dead cells. 'APC-A' channel is detecting AnnexinV-APC.

Figure S4: Hashtag deconvolution. **A)** Per donor hashtag noise distributions. Blue dotted lines indicate 99th percentile values for noise. Values above this line were called positive for a specific hashtag. **B)** (left) K-medoid clustering for each donor based only on hashtag reads. Cells positive ($p < 0.01$) for multiple hashtags are removed as likely multiplets. Cells are called as negative if they do not surpass the noise threshold for all hashtags, (right) k-medoid clustering with final hashtag labelling for non-multiplet cells

Figure S5: Filtering for cell quality. Total counts, N genes, and mitochondrial gene percentages shown for each donor before and after filtering; (top) pre-filtering and (bottom) post filtering.

Figure S6: Determining final lineage clusters. **A)** Original Leiden clustering for all cells. **B)** Separating EEC and secretory progenitors by organ. **C-E)** An *ITLN1*-high cluster, all from SI, contains cells expressing PC markers (*DEFA5*, *DEFA6*, *ITLN2*, *LYZ*) along with cells expressing GC marker *MUC2*. (C) cluster defined by *ITLN1* and dotplot showing expression of markers; (D) UMAP expression of PC and GC markers across all cells; (E) UMAP expression of PC and GC markers within the *ITLN1*-high cluster. **F)** Subclustering to define Paneth and goblet cells. **G)** Final lineage clusters used for the rest of the analyses in our study.

Figure S7: DEG dotplots for each lineage. Dotplots showing expression of top DEGs (max: 20) for each lineage, as sorted by expression fold-change above the cluster with the next highest expression. DEGs included are genes significantly enriched in both the SI and colon (if applicable).

Figure S8: Organ-specific lineage markers. Relating to Fig. 2E, UMAPs showing expression of DEGs from mature lineages found to be higher enriched in SI or colon.

Figure S9: Additional Paneth cell data. **A)** Dotplot showing members of major intestinal growth factor families with expression in PCs across SI lineages. **B)** Dotplot showing expression of all Frizzled family receptors across SI lineages. **C)** Heatmap of GC markers (top), mouse-defined Paneth-Like Cell markers (middle), and PC markers (bottom) plotted across colon GCs, secretory progenitors, and tuft cells. Cells in each lineage are sorted by increasing *REG4* expression (top row of middle third) to more easily visualize potential overlap of marker expression.

Figure S10: Follicle Associated Epithelium. **A)** Dotplot showing expression of conserved M-cell markers and other genes known to interact with the immune system across lineages. Bottom third shows genes implicated in mouse M-cells that are not specific to human FAE. **B)** Dotplot showing expression of top 20 FAE DEGs across lineages.

Figure S11: Additional mucin and goblet cell data. **A)** Dotplot showing expression of top nine expressing mucins across GCs and proliferative and absorptive lineages of the SI and colon. **B)** Dotplot showing expression of top 9 expressed mucins in all AE and ACC lineages by intestinal region. **C)** Dotplot showing expression of mucins in colonic intercrypt goblet cells (icGC), crypt-resident goblet cells (crGCs), and early goblet cells. **D)** (Left) Leiden subclustering of SI goblet cells, with subclusters named according to genes with high expression. (Middle) UMAP of SI goblet cells marked by diffusion pseudotime. (Right) UMAP of SI goblet cells marked by *MUC2* expression. **E)** Dotplot showing expression of mucins in SI GC subpopulations. **F)** Dotplot showing expression of mouse-implicated markers in human SI GC subclusters.

Figure S12: Additional EEC Data. **A)** Dotplot showing expression of DEGs found for SI or colon EECs that are present in the GOCC_Presynapse gene list. **B)** Pie chart showing all genes within the GOCC_Presynapse gene list marked by lineage in which they have highest expression (SI and colon lineages are combined when applicable).

File S1: REACTOME pathway enrichment analysis for TA vs TA2

File S2: REACTOME pathway enrichment analysis for TA2 vs TA

File S3: REACTOME pathway enrichment analysis for FAE DEGs

File S4: REACTOME pathway enrichment analysis for goblet cell DEGs

File S5: Receptor Family dotplots by lineage

Table S1: Differentially expressed gene signatures for lineages by organ

Table S2: Lineage signatures across full gut

Table S3: Correlations for BEST4⁺ NPY and AE gene expression

Table S4: Receptor family expression by lineage

Table S5: Primary targets of approved drugs expressed by lineage

Table S6: Phase I and Phase II drug metabolism genes expressed by lineage

Acknowledgements

We first thank the organ donors without whom our study would be impossible. We thank HonorBridge (formerly Carolina Donor Services) for assistance in coordinating organ donations. We thank Gabrielle Cannon at the Advanced Analytics core at UNC Chapel Hill and the Histology Research Core at UNC Chapel Hill. This work was supported by NIH R01DK115806, R01DK109559, and P30-DK034987 to STM, F32DK124929 to JRB, and F31-HL156433 and 5T32-GM067553 for JSR. APB is supported by a Career Development Award from the Crohn's and Colitis Foundation, and the University Cancer Research Fund.

Works cited

1. Barker N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nat Rev Mol Cell Biol* 2014;15:19-33.
2. Mowat AM, Agace WW. Regional specialization within the intestinal immune system. *Nat Rev Immunol* 2014;14:667-85.
3. Grün D, Lyubimova A, Kester L, et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* 2015;525:251-255.
4. Kim TH, Saadatpour A, Guo G, et al. Single-Cell Transcript Profiles Reveal Multilineage Priming in Early Progenitors Derived from Lgr5⁺ Intestinal Stem Cells. *Cell Reports* 2016;16:2053-2060.
5. Haber AL, Biton M, Rogel N, et al. A single-cell survey of the small intestinal epithelium. *Nature* 2017;551:333-339.
6. Holloway EM, Czerwinski M, Tsai YH, et al. Mapping Development of the Human Intestinal Niche at Single-Cell Resolution. *Cell Stem Cell* 2020;28:568-580.e4.
7. Fawcner-Corbett D, Antanaviciute A, Parikh K, et al. Spatiotemporal analysis of human intestinal development at single-cell resolution. *Cell* 2021;184:810-826.e23.
8. Elmentaite R, Ross ADB, Roberts K, et al. Single-Cell Sequencing of Developing Human Gut Reveals Transcriptional Links to Childhood Crohn's Disease. *Dev Cell* 2020;55:771-783 e5.
9. Smillie CS, Biton M, Ordovas-Montanes J, et al. Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* 2019;178:714-730.e22.
10. Huang B, Chen Z, Geng L, et al. Mucosal Profiling of Pediatric-Onset Colitis and IBD Reveals Common Pathogenics and Therapeutic Pathways. *Cell* 2019;179:1160-1176.e24.
11. Kanke M, Kennedy MM, Connelly S, et al. Single-cell analysis of colonic epithelium reveals unexpected shifts in cellular composition and molecular phenotype in treatment-naïve adult Crohn's disease. *bioRxiv* 2021:2021.01.13.426602.
12. Parikh K, Antanaviciute A, Fawcner-Corbett D, et al. Colonic epithelial cell diversity in health and inflammatory bowel disease. *Nature* 2019;567:49-55.
13. Wang Y, Song W, Wang J, et al. Single-cell transcriptome analysis reveals differential nutrient absorption functions in human intestine. *J Exp Med* 2020;217.
14. Fujii M, Matano M, Toshimitsu K, et al. Human Intestinal Organoids Maintain Self-Renewal Capacity and Cellular Diversity in Niche-Inspired Culture Condition. *Cell Stem Cell* 2018;23:787-793.e6.

15. Triana S, Stanifer ML, Metz-Zumaran C, et al. Single-cell transcriptomics reveals immune response of intestinal cell types to viral infection. *Mol Syst Biol* 2021;17:e9833.
16. Busslinger GA, Weusten BLA, Bogte A, et al. Human gastrointestinal epithelia of the esophagus, stomach, and duodenum resolved at single-cell resolution. *Cell Reports* 2021;34:108819.
17. Elmentaite R, Kumasaka N, Roberts K, et al. Cells of the human intestinal tract mapped across space and time. *Nature* 2021;597:250-255.
18. Wang Y, Gunasekara DB, Reed MI, et al. A microengineered collagen scaffold for generating a polarized crypt-villus architecture of human small intestinal epithelium. *Biomaterials* 2017;128:44-55.
19. Stoeckius M, Zheng S, Houck-Loomis B, et al. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol* 2018;19:224.
20. Korsunsky I, Millard N, Fan J, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* 2019;16:1289-1296.
21. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* 2019;9:5233.
22. Wolf FA, Hamey FK, Plass M, et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol* 2019;20:59.
23. Barker N, van Es JH, Kuipers J, et al. Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* 2007;449:1003-7.
24. van der Flier LG, van Gijn ME, Hatzis P, et al. Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell* 2009;136:903-12.
25. Whissell G, Montagni E, Martinelli P, et al. The transcription factor GATA6 enables self-renewal of colon adenoma stem cells by repressing BMP gene expression. *Nat Cell Biol* 2014;16:695-707.
26. Munoz J, Stange DE, Schepers AG, et al. The *Lgr5* intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers. *EMBO J* 2012;31:3079-91.
27. van der Flier LG, Haegerbarth A, Stange DE, et al. OLFM4 is a robust marker for stem cells in human intestine and marks a subset of colorectal cancer cells. *Gastroenterology* 2009;137:15-7.
28. Habowski AN, Flesher JL, Bates JM, et al. Transcriptomic and proteomic signatures of stemness and differentiation in the colon crypt. *Commun Biol* 2020;3:453.
29. Merlos-Suarez A, Barriga FM, Jung P, et al. The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell* 2011;8:511-24.
30. Gabut M, Bourdelais F, Durand S. Ribosome and Translational Control in Stem Cells. *Cells* 2020;9.
31. Woolnough JL, Atwood BL, Liu Z, et al. The Regulation of rRNA Gene Transcription during Directed Differentiation of Human Embryonic Stem Cells. *PLoS One* 2016;11:e0157276.
32. Zhang Q, Shalaby NA, Buszczak M. Changes in rRNA transcription influence proliferation and cell fate within a stem cell lineage. *Science* 2014;343:298-301.
33. Marshman E, Booth C, Potten CS. The intestinal epithelial stem cell. *Bioessays* 2002;24:91-8.
34. Kristensen AR, Gsponer J, Foster LJ. Protein synthesis rate is the predominant regulator of protein expression during differentiation. *Mol Syst Biol* 2013;9:689.
35. Fu W, Liu Y, Yin H. Mitochondrial Dynamics: Biogenesis, Fission, Fusion, and Mitophagy in the Regulation of Stem Cell Behaviors. *Stem Cells Int* 2019;2019:9757201.
36. Lu R, Markowitz F, Unwin RD, et al. Systems-level dynamic analyses of fate change in murine embryonic stem cells. *Nature* 2009;462:358-62.
37. Avansino JR, Chen DC, Hoagland VD, et al. Orthotopic transplantation of intestinal mucosal organoids in rodents. *Surgery* 2006;140:423-34.
38. Fukuda M, Mizutani T, Mochizuki W, et al. Small intestinal stem cell identity is maintained with functional Paneth cells in heterotopically grafted epithelium onto the colon. *Genes Dev* 2014;28:1752-7.
39. Herring CA, Banerjee A, McKinley ET, et al. Unsupervised Trajectory Analysis of Single-Cell RNA-Seq and Imaging Data Reveals Alternative Tuft Cell Origins in the Gut. *Cell Syst* 2018;6:37-51 e9.
40. van Es JH, Sato T, van de Wetering M, et al. *Dll1*+ secretory progenitor cells revert to stem cells upon crypt damage. *Nat Cell Biol* 2012;14:1099-1104.
41. Yang Q, Bermingham NA, Finegold MJ, et al. Requirement of *Math1* for secretory cell lineage commitment in the mouse intestine. *Science* 2001;294:2155-8.

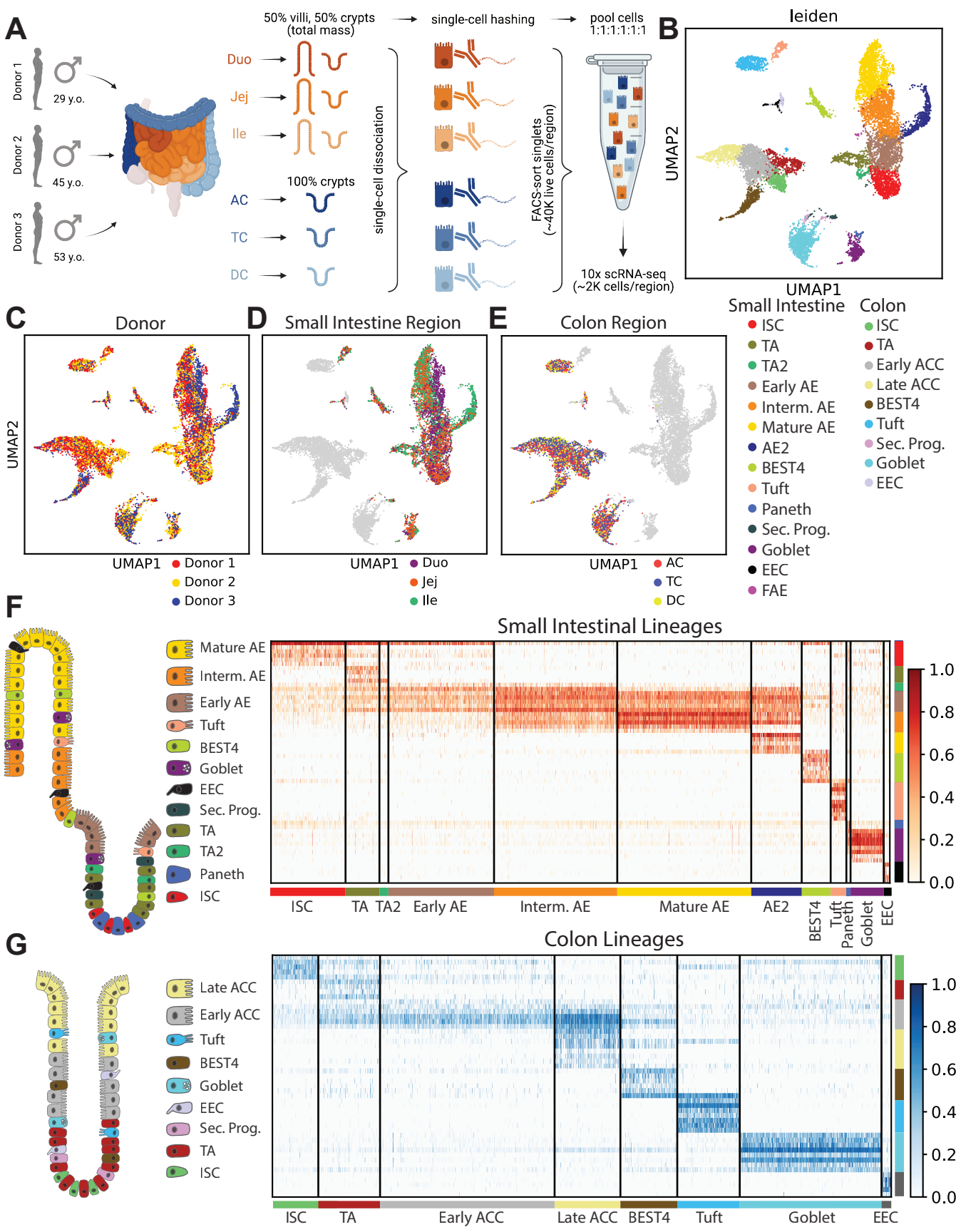
42. Sancho R, Cremona CA, Behrens A. Stem cell and progenitor fate in the mammalian intestine: Notch and lateral inhibition in homeostasis and disease. *EMBO Rep* 2015;16:571-81.
43. Kowalczyk MS, Tirosh I, Heckl D, et al. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. *Genome Res* 2015;25:1860-72.
44. Wang Y, Song W, Wang J, et al. Single-cell transcriptome analysis reveals differential nutrient absorption functions in human intestine. *Journal of Experimental Medicine* 2020;217.
45. Wille AC, Oliveira FA, Soares MA, et al. Cell cycle time and rate of entry of cells into mitosis in the small intestine of young rats. *Cell Prolif* 2004;37:189-94.
46. Gandara RM, Mahida YR, Potten CS. Regional differences in stem and transit cell proliferation and apoptosis in the terminal ileum and colon of mice after 12 Gy. *Int J Radiat Oncol Biol Phys* 2012;82:e521-8.
47. Dalton S. Linking the Cell Cycle to Cell Fate Decisions. *Trends Cell Biol* 2015;25:592-600.
48. Sato T, van Es JH, Snippert HJ, et al. Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* 2011;469:415-8.
49. Poulsen SS, Nexø E, Olsen PS, et al. Immunohistochemical localization of epidermal growth factor in rat and man. *Histochemistry* 1986;85:389-94.
50. Zhao J, de Vera J, Narushima S, et al. R-spondin1, a novel intestinotrophic mitogen, ameliorates experimental colitis in mice. *Gastroenterology* 2007;132:1331-43.
51. Shoshkes-Carmel M, Wang YJ, Wangenstein KJ, et al. Subepithelial telocytes are an important source of Wnts that supports intestinal crypts. *Nature* 2018;557:242-246.
52. Degirmenci B, Valenta T, Dimitrieva S, et al. GLI1-expressing mesenchymal cells form the essential Wnt-secreting niche for colon stem cells. *Nature* 2018;558:449-453.
53. Kabiri Z, Greicius G, Madan B, et al. Stroma provides an intestinal stem cell niche in the absence of epithelial Wnts. *Development* 2014;141:2206-15.
54. Durand A, Donahue B, Peignon G, et al. Functional intestinal stem cells after Paneth cell ablation induced by the loss of transcription factor Math1 (Atoh1). *Proc Natl Acad Sci U S A* 2012;109:8965-70.
55. Kim TH, Escudero S, Shivdasani RA. Intact function of Lgr5 receptor-expressing intestinal stem cells in the absence of Paneth cells. *Proc Natl Acad Sci U S A* 2012;109:3932-7.
56. Farin HF, Van Es JH, Clevers H. Redundant sources of Wnt regulate intestinal stem cells and promote formation of Paneth cells. *Gastroenterology* 2012;143:1518-1529 e7.
57. Lu H, Ju DD, Yang GD, et al. Targeting cancer stem cell signature gene SMOC-2 Overcomes chemoresistance and inhibits cell proliferation of endometrial carcinoma. *EBioMedicine* 2019;40:276-289.
58. Jubb AM, Chalasani S, Frantz GD, et al. Achaete-scute like 2 (ascl2) is a target of Wnt signalling and is upregulated in intestinal neoplasia. *Oncogene* 2006;25:3445-57.
59. Schuijers J, Junker JP, Mokry M, et al. Ascl2 acts as an R-spondin/Wnt-responsive switch to control stemness in intestinal crypts. *Cell Stem Cell* 2015;16:158-70.
60. Flanagan DJ, Pheasant TJ, Barker N, et al. Frizzled7 functions as a Wnt receptor in intestinal epithelial Lgr5(+) stem cells. *Stem Cell Reports* 2015;4:759-67.
61. Karasawa T, Yokokura H, Kitajewski J, et al. Frizzled-9 is activated by Wnt-2 and functions in Wnt/beta -catenin signaling. *J Biol Chem* 2002;277:37479-86.
62. Liu X, Lu R, Wu S, et al. Wnt2 inhibits enteric bacterial-induced inflammation in intestinal epithelial cells. *Inflamm Bowel Dis* 2012;18:418-29.
63. Kramer N, Schmollerl J, Unger C, et al. Autocrine WNT2 signaling in fibroblasts promotes colorectal cancer progression. *Oncogene* 2017;36:5460-5472.
64. Huang TX, Tan XY, Huang HS, et al. Targeting cancer-associated fibroblast-secreted WNT2 restores dendritic cell-mediated antitumour immunity. *Gut* 2021.
65. Jones JC, Brindley CD, Elder NH, et al. Cellular Plasticity of Defa4-Expressing Paneth Cells in Response to Notch Activation and Intestinal Injury. *Cellular and Molecular Gastroenterology and Hepatology* 2018;1-21.
66. Schmitt M, Schewe M, Sacchetti A, et al. Paneth Cells Respond to Inflammation and Contribute to Tissue Regeneration by Acquiring Stem-like Features through SCF/c-Kit Signaling. *Cell Reports* 2018;24:2312-2328. e7.
67. Yu S, Tong K, Zhao Y, et al. Paneth Cell Multipotency Induced by Notch Activation following Injury. *Cell Stem Cell* 2018;1-14.
68. Murata K, Jadhav U, Madha S, et al. Ascl2-Dependent Cell Dedifferentiation Drives Regeneration of Ablated Intestinal Stem Cells. *Cell Stem Cell* 2020;26:377-390 e6.

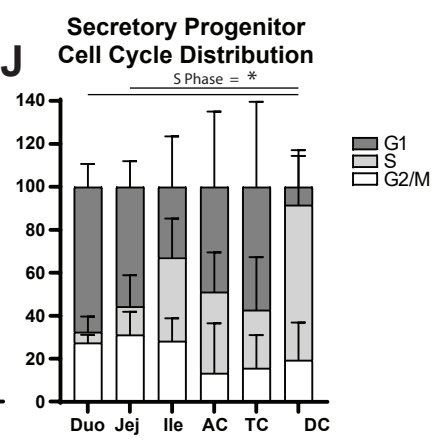
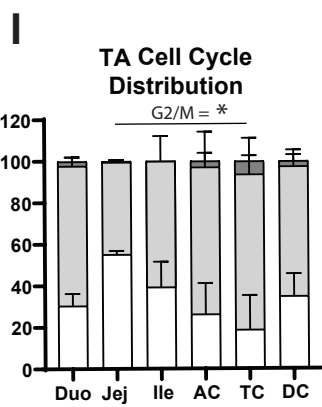
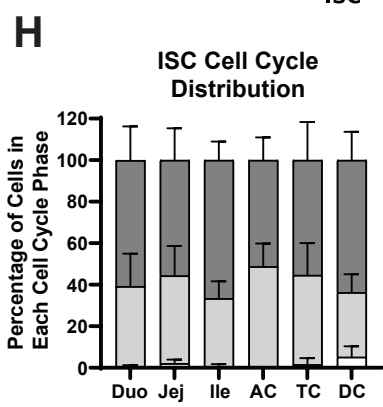
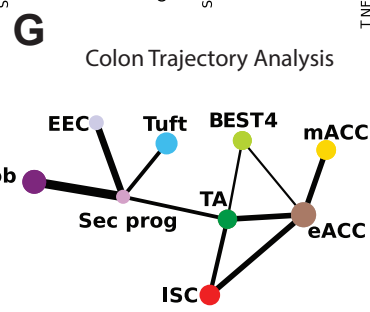
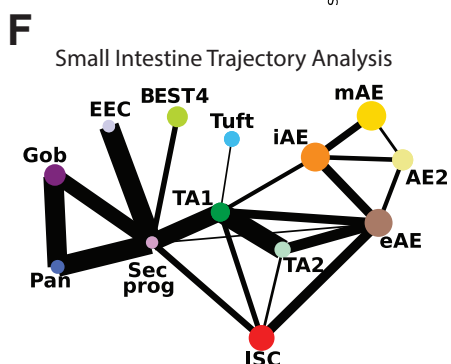
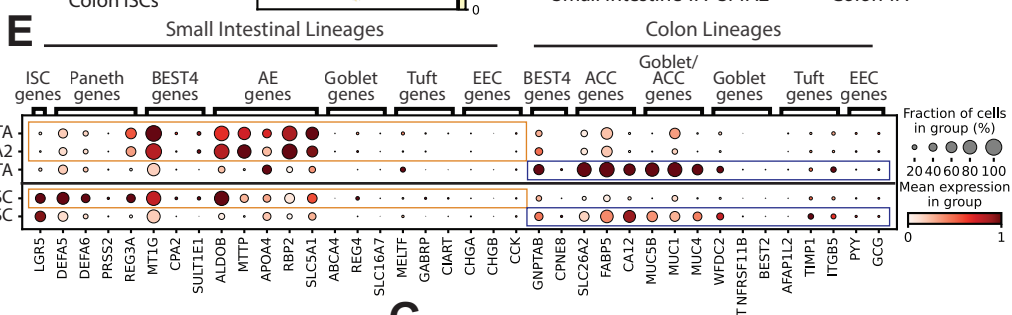
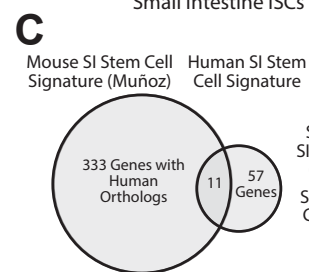
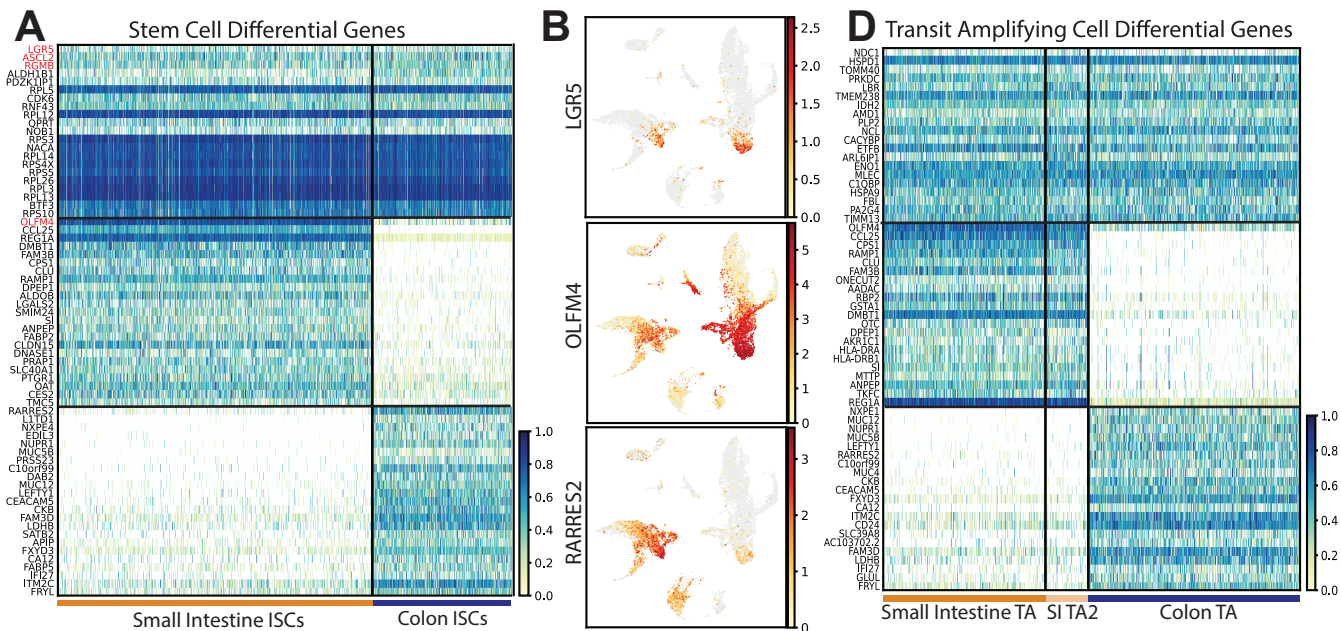
69. Rothenberg ME, Nusse Y, Kalisky T, et al. Identification of a cKit+ colonic crypt base secretory cell that supports Lgr5+ stem cells in mice. *Gastroenterology* 2012;142:1195-1205.e6.
70. Sasaki N, Sachs N, Wiebrands K, et al. Reg4+ deep crypt secretory cells function as epithelial niche for Lgr5+ stem cells in colon. *Proceedings of the National Academy of Sciences of the United States of America* 2016;113:E5399-E5407.
71. Salzman NH. Paneth cell defensins and the regulation of the microbiome: detente at mucosal surfaces. *Gut Microbes* 2010;1:401-6.
72. Jorgensen PB, Fenton TM, Morbe UM, et al. Identification, isolation and analysis of human gut-associated lymphoid tissues. *Nat Protoc* 2021;16:2051-2067.
73. Dillon A, Lo DD. M Cells: Intelligent Engineering of Mucosal Immune Surveillance. *Front Immunol* 2019;10:1499.
74. Hase K, Ohshima S, Kawano K, et al. Distinct gene expression profiles characterize cellular phenotypes of follicle-associated epithelium and M cells. *DNA Res* 2005;12:127-37.
75. Terahara K, Yoshida M, Igarashi O, et al. Comprehensive gene expression profiling of Peyer's patch M cells, villous M-like cells, and intestinal epithelial cells. *J Immunol* 2008;180:7840-6.
76. Nakato G, Fukuda S, Hase K, et al. New approach for m-cell-specific molecules screening by comprehensive transcriptome analysis. *DNA Res* 2009;16:227-35.
77. Lo D, Tynan W, Dickerson J, et al. Cell culture modeling of specialized tissue: identification of genes expressed specifically by follicle-associated epithelium of Peyer's patch by expression profiling of Caco-2/Raji co-cultures. *Int Immunol* 2004;16:91-9.
78. Kimura S, Kobayashi N, Nakamura Y, et al. Sox8 is essential for M cell maturation to accelerate IgA response at the early stage after weaning in mice. *J Exp Med* 2019;216:831-846.
79. Nakamura Y, Kimura S, Hase K. M cell-dependent antigen uptake on follicle-associated epithelium for mucosal immune surveillance. *Inflamm Regen* 2018;38:15.
80. Mercado-Lubo R, McCormick BA. A unique subset of Peyer's patches express lysozyme. *Gastroenterology* 2010;138:36-9.
81. Hase K, Murakami T, Takatsu H, et al. The membrane-bound chemokine CXCL16 expressed on follicle-associated epithelium and M cells mediates lympho-epithelial interaction in GALT. *J Immunol* 2006;176:43-51.
82. Zhao X, Sato A, Dela Cruz CS, et al. CCL9 is secreted by the follicle-associated epithelium and recruits dome region Peyer's patch CD11b+ dendritic cells. *J Immunol* 2003;171:2797-803.
83. Begg DP, Steinbrecher KA, Mul JD, et al. Effect of guanylate cyclase-C activity on energy and glucose homeostasis. *Diabetes* 2014;63:3798-804.
84. Sindic A. Current understanding of guanylin peptides actions. *ISRN Nephrol* 2013;2013:813648.
85. Valentino MA, Lin JE, Snook AE, et al. A uroguanylin-GUCY2C endocrine axis regulates feeding in mice. *J Clin Invest* 2011;121:3578-88.
86. Cox HM. Neuropeptide Y receptors; antisecretory control of intestinal epithelial function. *Auton Neurosci* 2007;133:76-85.
87. Wen J, Liu X, Qi Y, et al. BMP3 suppresses colon tumorigenesis via ActRIIB/SMAD2-dependent and TAK1/JNK signaling pathways. *J Exp Clin Cancer Res* 2019;38:428.
88. Loh K, Chia JA, Greco S, et al. Bone morphogenetic protein 3 inactivation is an early and frequent event in colorectal cancer development. *Genes Chromosomes Cancer* 2008;47:449-60.
89. Haghverdi L, Buttner M, Wolf FA, et al. Diffusion pseudotime robustly reconstructs lineage branching. *Nat Methods* 2016;13:845-8.
90. Holzer P, Reichmann F, Farzi A. Neuropeptide Y, peptide YY and pancreatic polypeptide in the gut-brain axis. *Neuropeptides* 2012;46:261-74.
91. Loh K, Herzog H, Shi YC. Regulation of energy homeostasis by the NPY system. *Trends Endocrinol Metab* 2015;26:125-35.
92. Kishi K, Takase S, Goda T. Enhancement of sucrase-isomaltase gene expression induced by luminally administered fructose in rat jejunum. *J Nutr Biochem* 1999;10:8-12.
93. Kalogeris TJ, Rodriguez MD, Tso P. Control of synthesis and secretion of intestinal apolipoprotein A-IV by lipid. *J Nutr* 1997;127:537S-543S.
94. Scheibner J, Trendelenburg AU, Hein L, et al. Alpha 2-adrenoceptors in the enteric nervous system: a study in alpha 2A-adrenoceptor-deficient mice. *Br J Pharmacol* 2002;135:697-704.
95. Si M, Lang J. The roles of metallothioneins in carcinogenesis. *J Hematol Oncol* 2018;11:107.

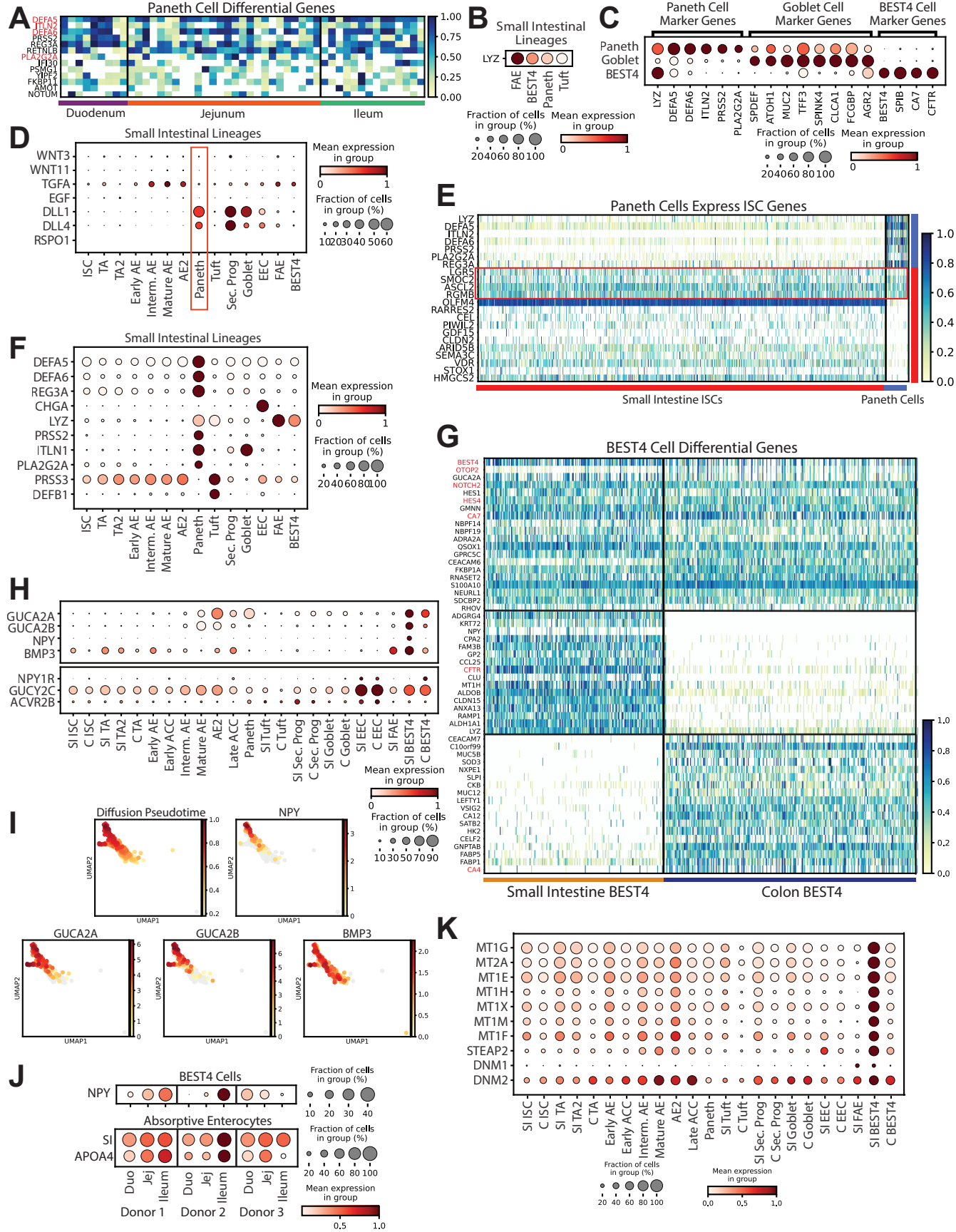
96. Klaassen CD, Liu J, Diwan BA. Metallothionein protection of cadmium toxicity. *Toxicol Appl Pharmacol* 2009;238:215-20.
97. Coyle P, Philcox JC, Carey LC, et al. Metallothionein: the multipurpose protein. *Cell Mol Life Sci* 2002;59:627-47.
98. Ohgami RS, Campagna DR, McDonald A, et al. The Steap proteins are metallo-reductases. *Blood* 2006;108:1388-94.
99. Anderson GJ, Frazer DM. Current understanding of iron homeostasis. *Am J Clin Nutr* 2017;106:1559S-1566S.
100. Schneider C, O'Leary CE, Locksley RM. Regulation of immune responses by tuft cells. *Nature Reviews Immunology* 2019;19:584-593.
101. O'Leary CE, Schneider C, Locksley RM. Tuft Cells-Systemically Dispersed Sensory Epithelia Integrating Immune and Neural Circuitry. *Annual Review of Immunology* 2019;37:47-72.
102. Nadsombati MS, McGinty JW, Lyons-Cohen MR, et al. Detection of Succinate by Intestinal Tuft Cells Triggers a Type 2 Innate Immune Circuit. *Immunity* 2018;49:33-41 e7.
103. von Moltke J, Ji M, Liang HE, et al. Tuft-cell-derived IL-25 regulates an intestinal ILC2-epithelial response circuit. *Nature* 2016;529:221-5.
104. Howitt MR, Lavoie S, Michaud M, et al. Tuft cells, taste-chemosensory cells, orchestrate parasite type 2 immunity in the gut. *Science* 2016;351:1329-33.
105. Gerbe F, van Es JH, Makrini L, et al. Distinct ATOH1 and Neurog3 requirements define tuft cells as a new secretory cell type in the intestinal epithelium. *J Cell Biol* 2011;192:767-80.
106. Bezençon C, Fürholz A, Raymond F, et al. Murine intestinal cells expressing Trpm5 are mostly brush cells and express markers of neuronal and inflammatory cells. *The Journal of Comparative Neurology* 2008;509:514-525.
107. Schneider C, O'Leary CE, von Moltke J, et al. A Metabolite-Triggered Tuft Cell-ILC2 Circuit Drives Small Intestinal Remodeling. *Cell* 2018;174:271-284 e14.
108. Roper SD, Chaudhari N. Taste buds: cells, signals and synapses. *Nature Reviews Neuroscience* 2017;18:485-497.
109. Kinnamon SC. Taste receptor signalling - from tongues to lungs. *Acta Physiol (Oxf)* 2012;204:158-68.
110. Vyas JM, Van der Veen AG, Ploegh HL. The known unknowns of antigen processing and presentation. *Nat Rev Immunol* 2008;8:607-18.
111. Kloetzel PM, Ossendorp F. Proteasome and peptidase function in MHC-class-I-mediated antigen presentation. *Curr Opin Immunol* 2004;16:76-81.
112. Lu Y-C, Yeh W-C, Ohashi PS. LPS/TLR4 signal transduction pathway. *Cytokine* 2008;42:145-151.
113. Hayashi F, Smith KD, Ozinsky A, et al. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* 2001;410:1099-1103.
114. Bauer S, Kirschning CJ, Häcker H, et al. Human TLR9 confers responsiveness to bacterial DNA via species-specific CpG motif recognition. *Proceedings of the National Academy of Sciences of the United States of America* 2001;98:9237-9242.
115. Xu M, Lu H, Lee YH, et al. An Interleukin-25-Mediated Autoregulatory Circuit in Keratinocytes Plays a Pivotal Role in Psoriatic Skin Inflammation. *Immunity* 2018;48:787-798 e4.
116. Wald D, Qin J, Zhao Z, et al. SIGIRR, a negative regulator of Toll-like receptor-interleukin 1 receptor signaling. *Nat Immunol* 2003;4:920-7.
117. Zhang C, Wu X, Zhao Y, et al. SIGIRR inhibits toll-like receptor 4, 5, 9-mediated immune responses in human airway epithelial cells. *Mol Biol Rep* 2011;38:601-9.
118. Qin J, Qian Y, Yao J, et al. SIGIRR inhibits interleukin-1 receptor- and toll-like receptor 4-mediated signaling through different mechanisms. *J Biol Chem* 2005;280:25233-41.
119. Schütz B, Ruppert A-L, Strobel O, et al. Distribution pattern and molecular signature of cholinergic tuft cells in human gastro-intestinal and pancreatic-biliary tract. *Scientific Reports* 2019;9:17466.
120. McGinty JW, Ting H-A, Billipp TE, et al. Tuft-Cell-Derived Leukotrienes Drive Rapid Anti-helminth Immunity in the Small Intestine but Are Dispensable for Anti-protist Immunity. *Immunity* 2020;52:528-541.e7.
121. Grondin JA, Kwon YH, Far PM, et al. Mucins in Intestinal Mucosal Defense and Inflammation: Learning From Clinical and Experimental Studies. *Front Immunol* 2020;11:2054.
122. Sicard JF, Le Bihan G, Vogeleeer P, et al. Interactions of Intestinal Bacteria with Components of the Intestinal Mucus. *Front Cell Infect Microbiol* 2017;7:387.
123. Pelaseyed T, Bergstrom JH, Gustafsson JK, et al. The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. *Immunol Rev* 2014;260:8-20.

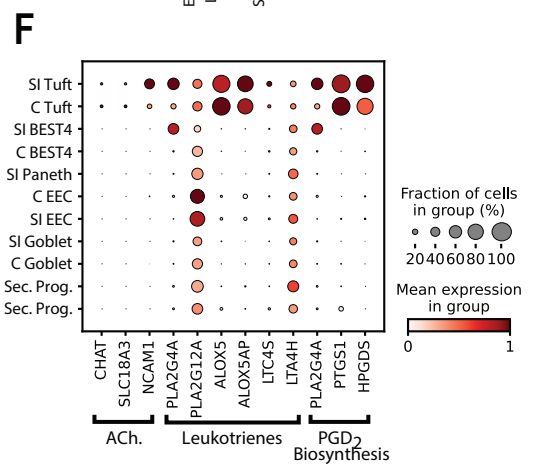
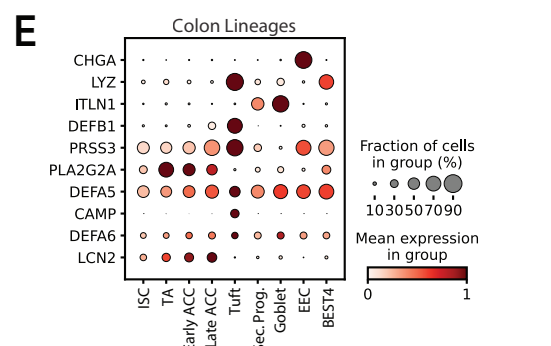
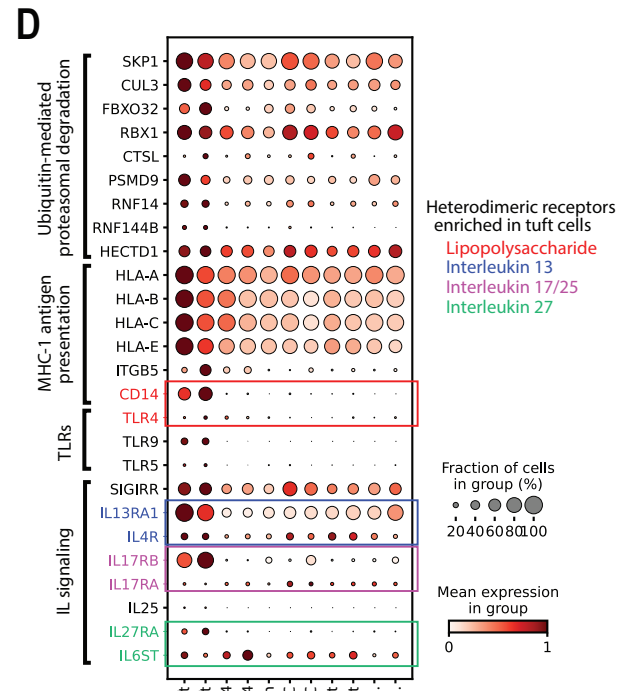
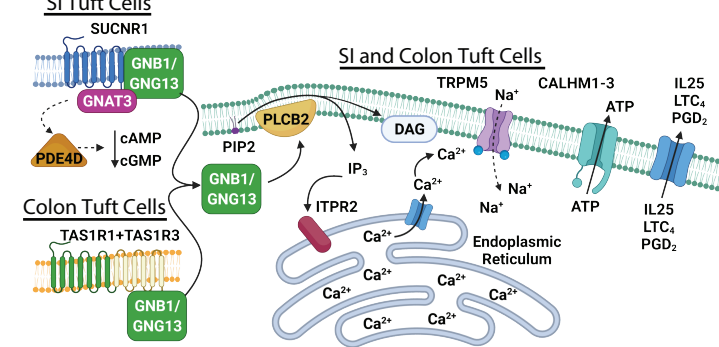
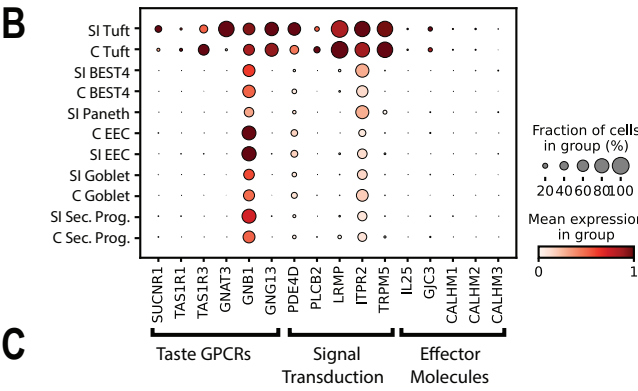
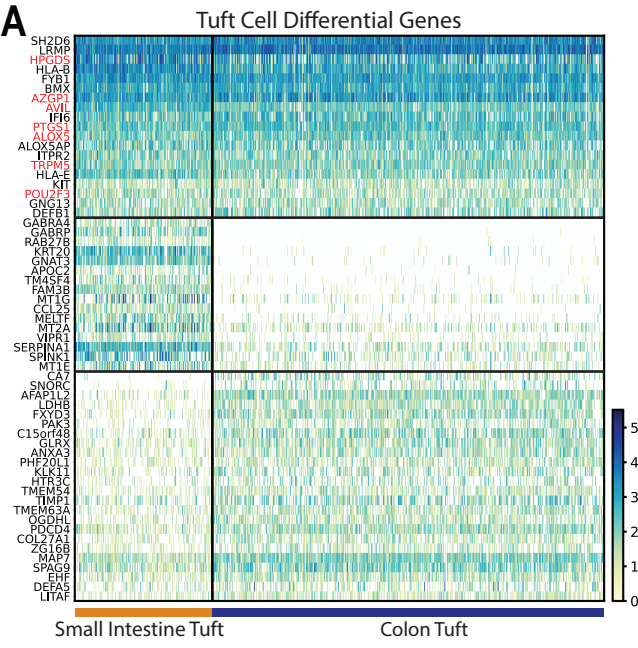
124. Pelaseyed T, Hansson GC. Membrane mucins of the intestine at a glance. *Journal of Cell Science* 2020;133.
125. Ma J, Rubin BK, Voynow JA. Mucins, mucus, and goblet cells. *Chest* 2018;154:169-176.
126. Jassal B, Matthews L, Viteri G, et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Research* 2020;48:D498-D503.
127. Cornick S, Tawiah A, Chadee K. Roles and regulation of the mucus barrier in the gut. *Tissue Barriers* 2015;3:e982426.
128. Layunta E, Javerfelt S, Dolan B, et al. IL-22 promotes the formation of a MUC17 glycocalyx barrier in the postnatal small intestine during weaning. *Cell Rep* 2021;34:108757.
129. Sun WW, Krystofiak ES, Leo-Macias A, et al. Nanoarchitecture and dynamics of the mouse enteric glycocalyx examined by freeze-etching electron tomography and intravital microscopy. *Commun Biol* 2020;3:5.
130. Nyström EEL, Martinez-Abad B, Arike L, et al. An intercrypt subpopulation of goblet cells is essential for colonic mucus barrier function. *Science* 2021;372.
131. Birchenough GM, Johansson ME, Gustafsson JK, et al. New developments in goblet cell mucus secretion and function. *Mucosal Immunol* 2015;8:712-9.
132. Beumer J, Puschhof J, Bauza-Martinez J, et al. High-Resolution mRNA and Secretome Atlas of Human Enteroendocrine Cells. *Cell* 2020;181:1291-1306 e19.
133. Billing LJ, Larraufie P, Lewis J, et al. Single cell transcriptomic profiling of large intestinal enteroendocrine cells in mice - Identification of selective stimuli for insulin-like peptide-5 and glucagon-like peptide-1 co-expressing cells. *Mol Metab* 2019;29:158-169.
134. Roberts GP, Larraufie P, Richards P, et al. Comparison of Human and Murine Enteroendocrine Cells by Transcriptomic and Peptidomic Profiling. *Diabetes* 2019;68:1062-1072.
135. Sjolund K, Sanden G, Hakanson R, et al. Endocrine cells in human intestine: an immunocytochemical study. *Gastroenterology* 1983;85:1120-30.
136. Liddle RA. Neuropods. *Cell Mol Gastroenterol Hepatol* 2019;7:739-747.
137. Kaelberer MM, Buchanan KL, Klein ME, et al. A gut-brain neural circuit for nutrient sensory transduction. *Science* 2018;361.
138. Bellono NW, Bayrer JR, Leitch DB, et al. Enterochromaffin Cells Are Gut Chemosensors that Couple to Sensory Neural Pathways. *Cell* 2017;170:185-198 e16.
139. Worthington JJ, Reimann F, Gribble FM. Enteroendocrine cells-sensory sentinels of the intestinal environment and orchestrators of mucosal immunity. *Mucosal Immunol* 2018;11:3-20.
140. Gribble FM, Reimann F. Signalling in the gut endocrine axis. *Physiol Behav* 2017;176:183-188.
141. Egerod KL, Engelstoft MS, Grunddal KV, et al. A major lineage of enteroendocrine cells coexpress CCK, secretin, GIP, GLP-1, PYY, and neurotensin but not somatostatin. *Endocrinology* 2012;153:5782-95.
142. Habib AM, Richards P, Cairns LS, et al. Overlap of endocrine hormone expression in the mouse intestine revealed by transcriptional profiling and flow cytometry. *Endocrinology* 2012;153:3054-65.
143. Kiela PR, Ghishan FK. Physiology of Intestinal Absorption and Secretion. *Best Pract Res Clin Gastroenterol* 2016;30:145-59.
144. Gajda AM, Storch J. Enterocyte fatty acid-binding proteins (FABPs): different functions of liver and intestinal FABPs in the intestine. *Prostaglandins Leukot Essent Fatty Acids* 2015;93:9-16.
145. Dawson PA. Role of the intestinal bile acid transporters in bile acid and drug disposition. *Handb Exp Pharmacol* 2011:169-203.
146. Dawson PA, Karpen SJ. Intestinal transport and metabolism of bile acids. *J Lipid Res* 2015;56:1085-99.
147. Agellon LB, Toth MJ, Thomson AB. Intracellular lipid binding proteins of the small intestine. *Mol Cell Biochem* 2002;239:79-82.
148. Imamura T, Kitamoto Y. Expression of enteropeptidase in differentiated enterocytes, goblet cells, and the tumor cells in human duodenum. *Am J Physiol Gastrointest Liver Physiol* 2003;285:G1235-41.
149. Zhou FF, Xie W, Chen SQ, et al. SLC38A1 promotes proliferation and migration of human colorectal cancer cells. *J Huazhong Univ Sci Technolog Med Sci* 2017;37:30-36.
150. Lu Z, Ding L, Lu Q, et al. Claudins in intestines: Distribution and functional significance in health and diseases. *Tissue barriers* 2013;1:e24978.
151. Kiela PR, Ghishan FK. Physiology of intestinal absorption and secretion. *Best Practice & Research. Clinical Gastroenterology* 2016;30:145-159.
152. Camilleri M. Leaky gut: mechanisms, measurement and clinical implications in humans. *Gut* 2019;68:1516-1526.

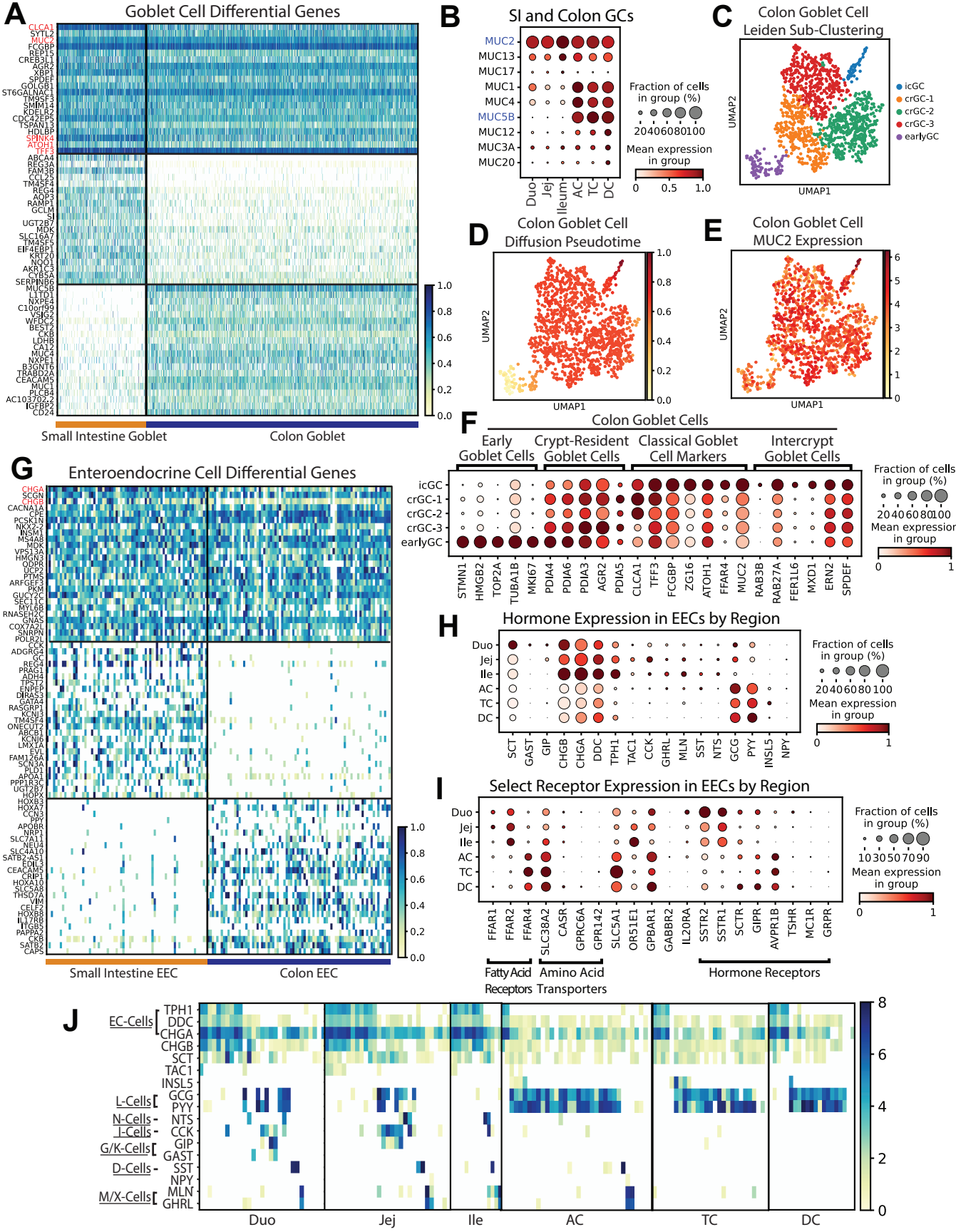
153. Koutroubakis IE. Recent advances in the management of distal ulcerative colitis. *World J Gastrointest Pharmacol Ther* 2010;1:43-50.
154. Ahmad R, Rah B, Bastola D, et al. Obesity-induces Organ and Tissue Specific Tight Junction Restructuring and Barrier Deregulation by Claudin Switching. *Sci Rep* 2017;7:5125.
155. Zhu C, Chen Z, Jiang Z. Expression, distribution and role of aquaporin water channels in human and animal stomach and intestines. *International Journal of Molecular Sciences* 2016;17.
156. Hua S. Advances in Oral Drug Delivery for Regional Targeting in the Gastrointestinal Tract - Influence of Physiological, Pathophysiological and Pharmaceutical Factors. *Front Pharmacol* 2020;11:524.
157. Thelen K, Dressman JB. Cytochrome P450-mediated metabolism in the human gut wall. *J Pharm Pharmacol* 2009;61:541-58.
158. Speer JE, Wang Y, Fallon JK, et al. Evaluation of human primary intestinal monolayers for drug metabolizing capabilities. *J Biol Eng* 2019;13:82.
159. Iswandana R, Irianti MI, Oosterhuis D, et al. Regional Differences in Human Intestinal Drug Metabolism. *Drug Metab Dispos* 2018;46:1879-1885.
160. Philpott HL, Nandurkar S, Lubel J, et al. Drug-induced gastrointestinal disorders. *Frontline Gastroenterol* 2014;5:49-57.
161. Wang D, Zou L, Jin Q, et al. Human carboxylesterases: a comprehensive review. *Acta Pharm Sin B* 2018;8:699-712.
162. Ando Y, Saka H, Asai G, et al. UGT1A1 genotypes and glucuronidation of SN-38, the active metabolite of irinotecan. *Ann Oncol* 1998;9:845-7.
163. Pitot HC, Adjei AA, Reid JM, et al. A phase I and pharmacokinetic study of a powder-filled capsule formulation of oral irinotecan (CPT-11) given daily for 5 days every 3 weeks in patients with advanced solid tumors. *Cancer Chemother Pharmacol* 2006;58:165-72.
164. Rothenberg ML, Kuhn JG, Schaaf LJ, et al. Alternative dosing schedules for irinotecan. *Oncology (Williston Park)* 1998;12:68-71.
165. Kumler I, Sorensen PG, Palshof J, et al. Oral administration of irinotecan in patients with solid tumors: an open-label, phase I, dose escalating study evaluating safety, tolerability and pharmacokinetics. *Cancer Chemother Pharmacol* 2019;83:169-178.
166. Panes J, Vermeire S, Lindsay JO, et al. Tofacitinib in Patients with Ulcerative Colitis: Health-Related Quality of Life in Phase 3 Randomised Controlled Induction and Maintenance Studies. *J Crohns Colitis* 2018;12:145-156.
167. Sanders DB, Hart IK, Mantegazza R, et al. An international, phase III, randomized trial of mycophenolate mofetil in myasthenia gravis. *Neurology* 2008;71:400-6.
168. Kremer JM, Cohen S, Wilkinson BE, et al. A phase IIb dose-ranging study of the oral JAK inhibitor tofacitinib (CP-690,550) versus placebo in combination with background methotrexate in patients with active rheumatoid arthritis and an inadequate response to methotrexate alone. *Arthritis Rheum* 2012;64:970-81.
169. Seyhan AA, Carini C. Are innovation and new technologies in precision medicine paving a new era in patients centric care? *J Transl Med* 2019;17:114.

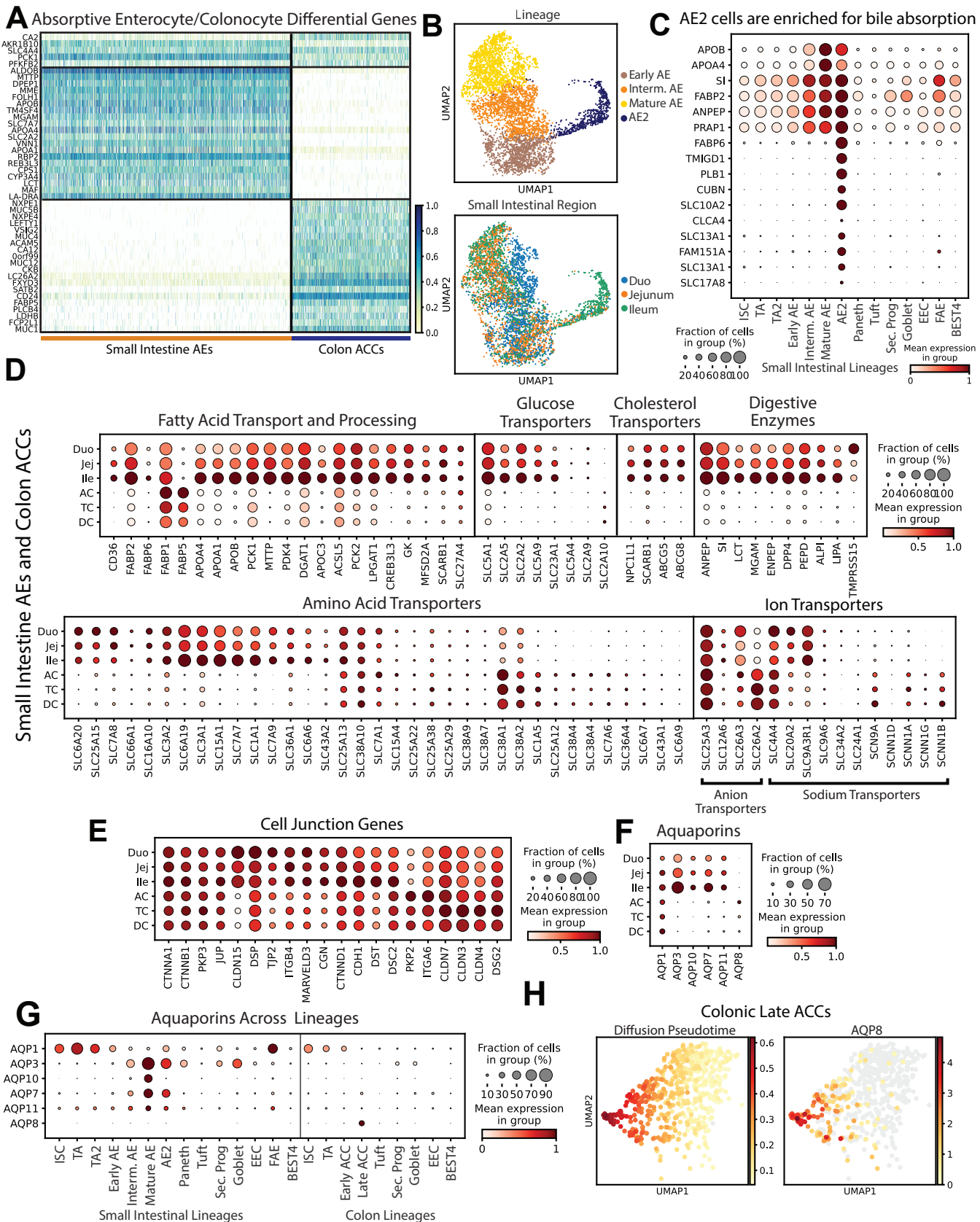


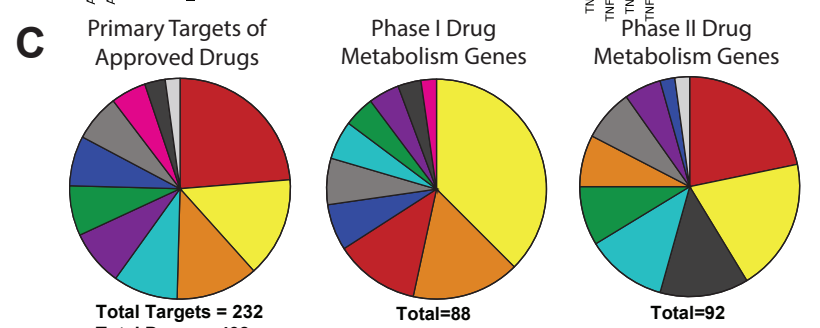
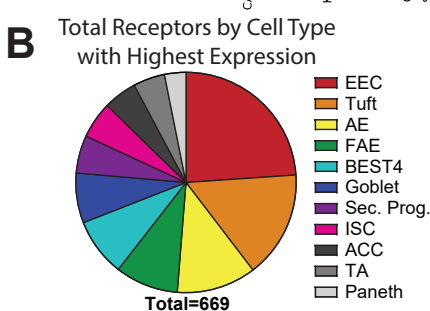
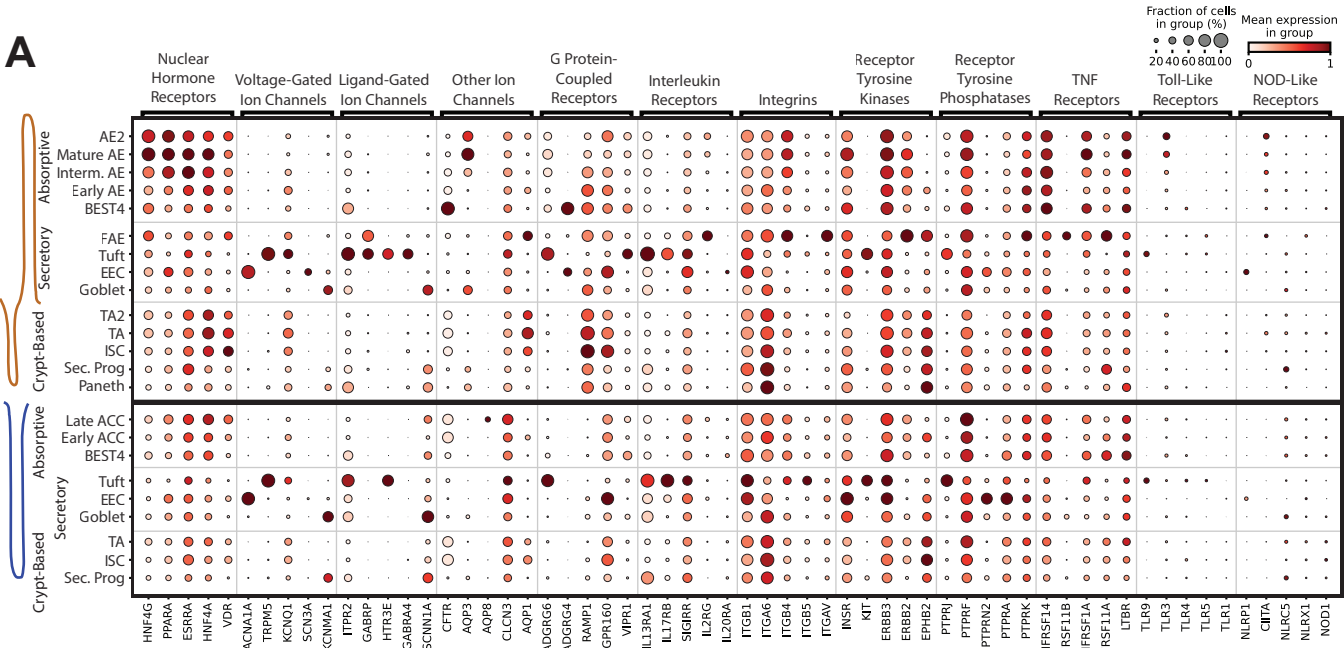












D

Drug	Function	Target
Tacrolimus	Immunosuppressant	FKBP1A
Mycophenolate Mofetil	Immunosuppressant	IMPDH1, IMPDH2
Tofacitinib	JAK Inhibitor	JAK1, JAK2, JAK3
Methotrexate	Immunosuppressant	DHFR
Prednisolone	Corticosteroid	NR3C1
Prednisone	Corticosteroid	NR3C1
Natalizumab	Immune Cell Blocker	ITGA4
Vedolizumab	Anti-Inflammatory	ITGA4

