

The cellular states and fates of shed intestinal cells

Keren Bahar Halpern^{1*}, Yael Korem Kohanim², Adi Biram³, Adi Egozi¹, Ziv Shulman³, Shalev Itzkovitz^{1*}

¹Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel

²Department of Immunobiology, Yale University School of Medicine, New Haven, CT

³Department of Systems Immunology, Weizmann Institute of Science, Rehovot, Israel

*To whom correspondence should be addressed. Email: kerenb@weizmann.ac.il, shalev.itzkovitz@weizmann.ac.il

Abstract

The intestinal epithelium is replaced every few days¹. Enterocytes are shed into the gut lumen predominantly from the tips of villi^{3,4}, and are believed to rapidly die upon their dissociation from the tissue. However, technical limitations prohibited studying the cellular states and fates of shed intestinal cells. Here, we used bulk and single cell RNA sequencing of mouse intestinal fecal washes to demonstrate that shed epithelial cells remain viable and up-regulate distinct anti-microbial programs upon shedding. We further identify abundant shedding of immune cells, a process that is elevated in DSS-induced colitis. We find that fecal host transcriptomics mirrors changes in the intestinal tissue following perturbations. Our study suggests potential functions of shed cells in the intestinal lumen and demonstrates that host cell transcriptomes in intestinal washes can be used to probe tissue states.

Introduction

The gut epithelium is one of the cell populations with the highest turnover rates in mammals. The mouse epithelium is replaced every 3-5 days¹ and the human gut sheds more than 30 grams of cells every day². Gut cells are shed into the lumen predominantly from the tips of villi in the small intestine and from the inter-crypt epithelium in the large intestine^{3,4}. This shedding is a result of replacement by cells that are continuously produced at the crypts and driven by mitotic pressure⁵⁻⁷. It is believed that shed cells enter anoikis, a process in which the cell loses contact with adjacent cells as well as with the extracellular matrix⁸⁻¹². Notably, some studies showed shedding of cells without apoptotic markers in the human colon¹³, suggesting that shed cells may continue to be viable and perhaps even functional. This hypothesis has thus far not been systematically examined. Understanding the biology of intestinal shed cells has important clinical implications. Several studies analyzed stool host transcriptomes to investigate gut development¹⁴, disease progression¹⁵ or impacts of therapeutic regimens¹⁶. Fecal wash host transcriptomes collected during

colonoscopies have been shown to be statistically powerful classifiers of histological remission in patients with Inflammatory Bowel Disease (IBD)¹⁷. These findings highlighted the potential prognostic value of shed cell transcriptomics in the gut. However, it remains unclear whether these informative fecal transcriptomes are remnants of cells that have died immediately after shedding, or whether shed cells may still be viable and even functionally distinct upon their dissociation from the tissue.

Here, we present a comprehensive analysis of mouse fecal-wash host-transcriptomes throughout the mouse intestinal tract. We performed bulk and single cell RNA sequencing (scRNAseq) as well as single molecule FISH (smFISH) and intra-vital imaging to study the identities and cellular states of shed cells. We show that substantial amounts of immune cells are regularly shed during homeostasis, and that immune cell shedding increases in DSS-induced colitis. Our work suggests that shed cells may continue to carry functions in the lumen, and establishes the power of fecal wash host transcriptomics in mirroring mucosal physiology and pathology.

Results

Fecal wash host transcriptomes have distinct segment-dependent profiles

To study the molecular identities of shed intestinal cells in mice, we divided the gut into 5 segments, 3 equal-length segments spanning the small intestine (duodenum, jejunum and ileum) and 2 equal-length segments spanning the large intestine (proximal and distal colon, Fig. 1a). We flushed each segment with PBS into a tube (i.e. Fecal wash). In addition, we collected a small piece of tissue into a separate tube. We performed bulk RNA sequencing using the mcSCRbseq protocol¹⁸ (Supplementary Table 1). Principal Component Analysis (PCA, Fig. 1b) of the transcriptomic data revealed two axes of variations. The first PC correlated with the intestinal segments, highlighting large variation between the small and large intestine and smaller variations among sub-segments (Fig. 1b). The second PC separated the tissue and fecal wash samples.

Variations among segments in fecal wash samples from the small intestine were less pronounced compared to the tissue samples (between segment /within segment ratio of 1.18 for tissues vs. 1.13 for fecal washes, $p < 3e-196$). We reasoned that fecal washes at each segment might still contain transcripts from cells that were shed in up-stream segments. Genes with clear segmentally varying profiles, such as the proximal Fabp1 and the distal Fabp6¹⁹ showed broadly similar segmental profiles in the fecal washes (Fig. 1c). Notably, fecal wash segmental profiles often declined less abruptly compared to the tissues (e.g. Gstm3 and Clec2h, Fig. 1c), or even exhibited a shift in the segmental peak (e.g. Itln1 and Fabp2, Fig. 1c), consistent with a picture of elongated lifetime of the shed cells or their mRNA content.

Using multi-linear regression (Fig. 1d), we found that fecal wash transcriptomes predominantly constituted weighted averages of the transcriptomes of the matching and immediate up-stream tissue segments. The proximal colon fecal wash (pCol-F) had lower representation of the upstream ileum samples, potentially indicating increased degradation of mRNA at the cecum. Since the intestinal segment correlates with transit time, we formulated a mathematical model of cell shedding (Methods, Fig. 1e-f) that facilitates inferences of shed cell transcriptome lifetimes. The model incorporated the contribution of shed cell transcriptomes at each segment (shedding rate α) and the luminal transcript degradation rate (δ). Estimated mRNA lifetimes were on the order of 1.2-1.6 hours in the small intestine and 2.5-3.7 hours in the colon (Fig. 1g), assuming a total transit time of 8 hrs throughout the mouse gastrointestinal tract with faster transit in the small intestine compared to the large intestine²⁰⁻²² (Methods). Shed cells, or their mRNA content therefore seem to survive for several hours in the gut lumen.

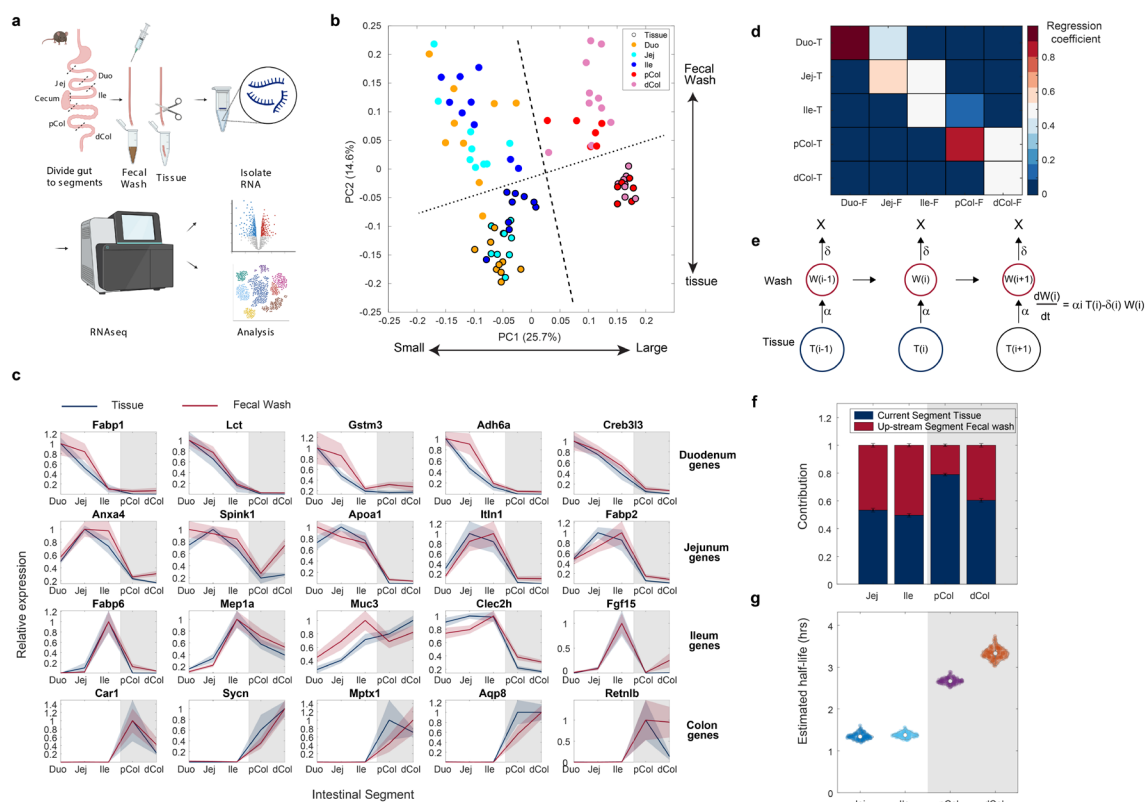


Fig. 1 – Fecal wash host transcriptomes have distinct segment-dependent profiles. **a.** Illustration of fecal wash and tissue sample collection and experimental layout. Duo - Duodenum, Jej – Jejunum, Ile – Ileum, pCol- Proximal colon, dCol - Distal colon. Created with BioRender.com. **b.** Principal component analysis (PCA) of all control samples. Samples are color coded by intestinal segment source. Black circles denote tissue samples. Numbers in parentheses show percent of explained variance by each PC. **c.** Gene expression segmental profiles of tissue samples (blue) or fecal wash samples (red) along the gastrointestinal tract. Gray

shades denote the colon. **d.** Multi-linear regression of the fecal wash transcriptomes based on the segmental tissue transcriptomes. Shown are the regression coefficients. **e.** Mathematical model of shed cell turnover. Intestinal segment correlates with time, shed cells are produced at rate α and removed at rate δ . **f.** Fecal wash transcriptome predominantly constituted a weighted average of the matching tissue contribution (blue) and previous segment fecal wash (red, Methods). **g.** Estimated half-life of shed cells throughout the gastrointestinal tract (Methods). Data was bootstrapped to obtain statistics, circles denote the medians, vertical bars denote the Inter-Quartile Range (IQR).

Differences in gene expression between shed and tissue cells

To explore the differences in expression between the tissue and the shed cells, we performed differential gene expression analysis (Fig. 2, Extended Data Fig. 1, and Supplementary Table 2). Genes that were upregulated in fecal washes included immune-related genes such as *Adrgre1* encoding the macrophage marker F4/80, the T-cell marker gene *Ccl5*, as well as bacterial defense genes such as *Nfkbia* and *Cdkn1a* (Fig. 2a). Genes enriched in the tissue included the proximal and distal small intestinal genes *Fabp1* and *Fabp6*¹⁹ respectively, as well as mesenchymal genes such as *Colla1* (Fig. 2a). Gene Set Enrichment Analysis (GSEA)²³ analysis indicated that fecal washes from both small and large intestine were further enriched in genes associated with TNF α signaling, IFN γ/α signaling, and hypoxia pathways, and depleted in pathways such as oxidative phosphorylation and cell cycle pathways G2M checkpoint and E2F targets (Fig. 2b). In addition, many villus tip markers were elevated in fecal washes, including *Fos*, *Ada* and *Nt5e* (Fig. 2a), indicating an increase cell shedding from the villus tip²⁴. To infer cell type composition of both fecal washes and tissues we used the cellanneal computational deconvolution software²⁵. We found that fecal washes were enriched in villus tip cells. Notably, fecal washes were also enriched in neutrophils, monocytes and macrophages (Fig. 2c,d).

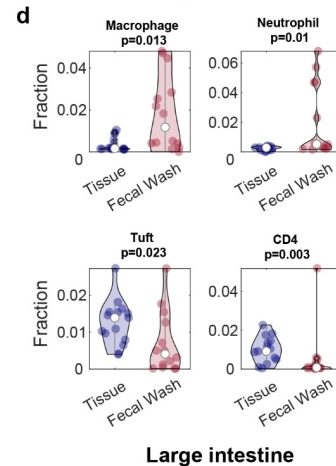
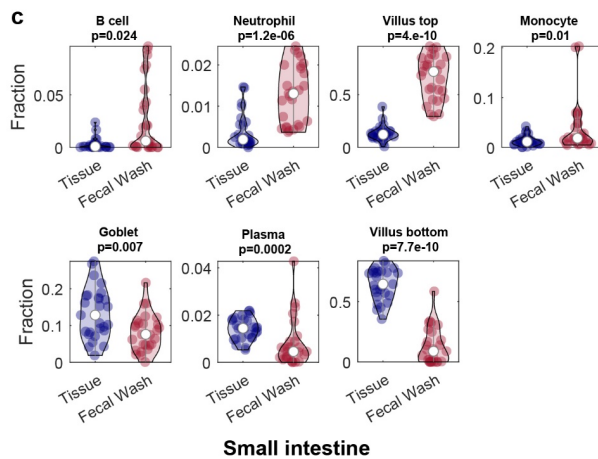
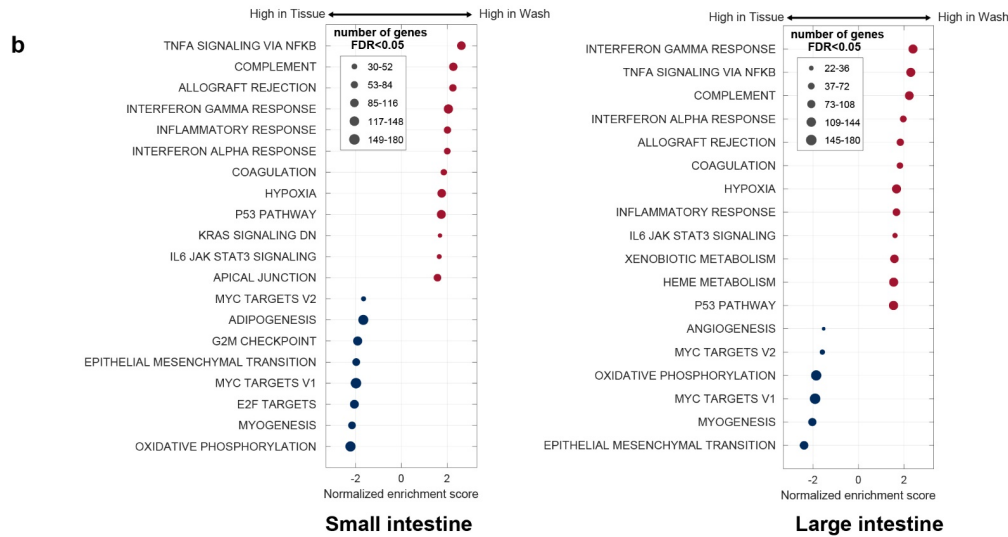
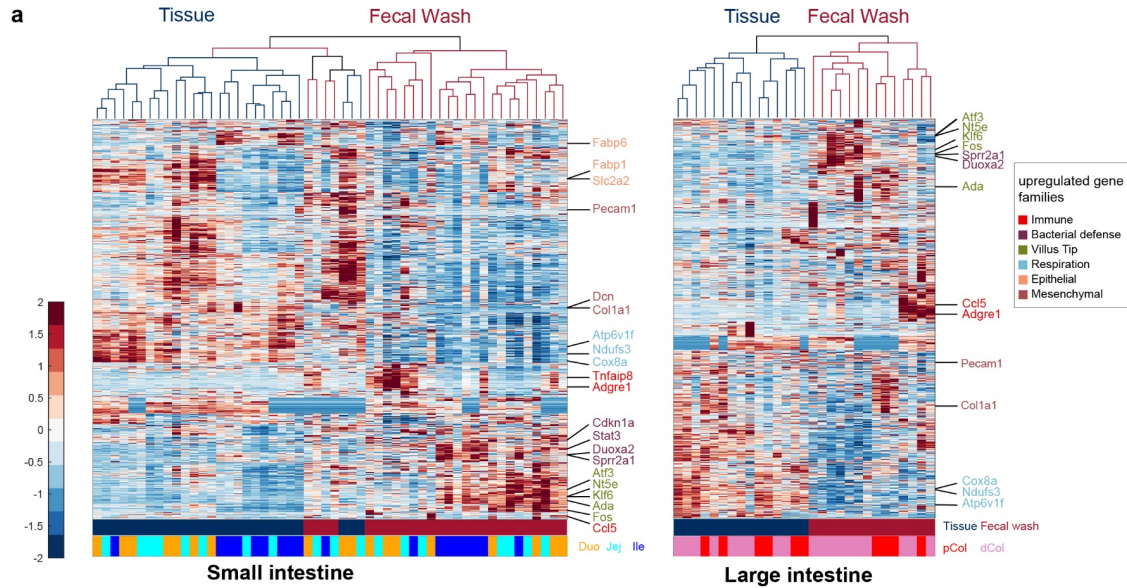


Fig. 2 – Differences in gene expression between the tissue and the shed cells. **a.** Hierarchical clustering of fecal wash samples (red branches) or Tissues (blue branches). Left, small intestine samples, right, large intestine samples. Samples naming nomenclature and color coding; Duo – Duodenum - Orange, Jej – Jejunum - Cyan, Ile – Ileum - Blue, pCol- Proximal colon - Red, dCol - Distal colon - Magenta. Representative genes are shown on the right. Expression values are Z scores (methods). **b.** GSEA over the Hallmark gene sets. Left - small intestine, right - large intestine, included are gene sets with q-value <0.05. Gene sets elevated in fecal washes (red circles) were associated with immune cell pathways, while gene sets elevated in the tissue samples (blue circles) included cell proliferation and respiratory pathways. FDR, false discovery rate; **c-d.** Estimated fractions of different cell types in fecal washes and tissue samples inferred by computational deconvolution²⁵. **c** – small intestine, **d** – large intestine. White circles are medians, vertical bars denote the Inter-Quartile Range (IQR). P-values computed using Kruskal-Wallis tests.

Fecal washes contain diverse intact viable cells

The inferred abundance of immune cells in fecal washes was unexpected, as cell shedding is thought to be dominated by epithelial cells in homeostasis⁷. To further establish this finding, we sorted fecal wash cells into 384 well capture plates and sequenced them using the MARSseq protocol (Fig. 3a, Extended Data Fig. 2)²⁶. Consistent with the computational deconvolution results, our scRNAseq atlas included abundant monocytes, B-cells, Cycling monocytes, Mast cells and T-cells, in addition to epithelial Enterocytes, Goblet cells and Tuft cells (Fig. 3a,b). Using smFISH of fecal wash cells, we validated the existence of intact enterocytes (marked by Ada), B-cells (marked by Cd19) and monocytes (marked by Lyz1, Fig. 3c-e). Notably, shed cells exhibited bright transcription sites, indicating active transcription (Fig. 3c, white arrow). Using intravital imaging (Methods), we identified monocyte shedding in a CX3CR1^{GFP} mouse model (Fig. 3f, supplementary movie 1, Methods). Shed cells therefore seem to be diverse, viable and transcriptionally active.

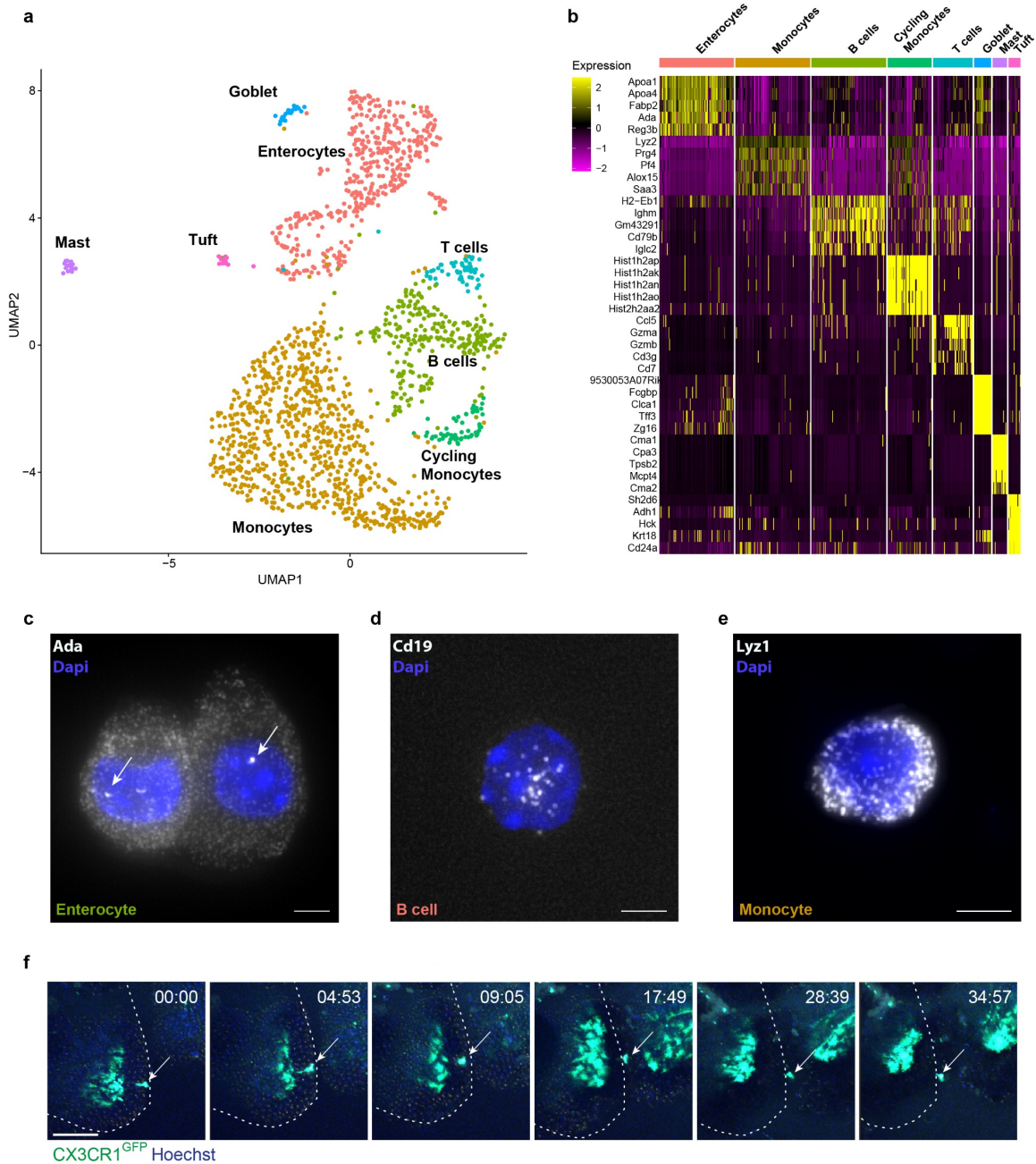


Fig. 3 – single cell analysis of fecal washes reveals diverse cell type presence. **a.** UMAP (Uniform Manifold Approximation and Projection) visualization of single cells obtained from small intestine fecal washes analyzed by single cell MARSseq. Cells clustered into 8 groups colored by cell type annotation. $n=1,717$ cells from 4 mice. **b.** Heatmap showing the top 5 differentially expressed genes in each cell type. Expression is log-normalized by Seurat²⁷, cells randomly down-sampled to a maximum of 100 per cell type. **c-e.** smFISH validation of selected cell type markers found in **b**, Ada mark enterocytes, white arrow marks an active transcription site, Cd19 marks B-cell and Lyz1 marks monocytes. Scale bar is $5\mu\text{m}$. **f.** Intra-vital imaging of CX3CR1^{GFP} mice injected with $10\mu\text{g}$ TNF- α intraperitoneally in PBS 20 min prior to the surgery (Methods). Capture movie was 35 minutes. Scale bar is $70\mu\text{m}$. Dashed lines mark villus border. Arrow highlights a shedding monocyte (supplementary movie 1).

Shed enterocytes upregulate anti-inflammatory modules

Our findings that shed cells continue to be viable and transcriptionally active, suggested that they might carry luminal functions that are distinct from their tissue roles. In order to understand the changes enterocytes undergo after shedding, we next combined our fecal wash scRNAseq measurements with scRNAseq of tissue enterocytes²⁸ (Fig. 4a). We used landmark genes²⁴, to infer the location of each cell along the crypt-villus axis, denoted as ‘zonation score’, ranging from 0 at the crypt to 1 at the villus tip (Fig. 4a). We found that the fecal wash cells had a clear villus tip signature (median zonation score 0.73 for shed enterocytes and 0.16 for tissue enterocytes, $p=8e-102$, Methods, Fig. 4b). Notably, differential gene expression between the shed enterocytes and tissue enterocytes with matched zonation distributions identified significant expression changes (1,589 genes out of 7,021 genes with expression above 10^{-5} were significantly different between the groups at false discovery rate (FDR) of 0.1). Genes that were upregulated in fecal wash enterocytes compared to tissue enterocytes included the anti-microbial genes *Reg3g* and *Reg3b*, the redox-signaling genes *Duoxa2* and *Duox2* and the pro-inflammatory response genes *Saa1* and *Cd74* genes (Fig. 4c). Shed enterocytes were enriched in genes associated with TNF α signaling, IFN γ/α signaling, and hypoxia pathways in addition to IL6 signaling, and the JAK STAT pathway, and depleted in oxidative phosphorylation related genes (Fig. 4d). Our data indicates that shed enterocytes up-regulate anti-microbial and immune-related processes.

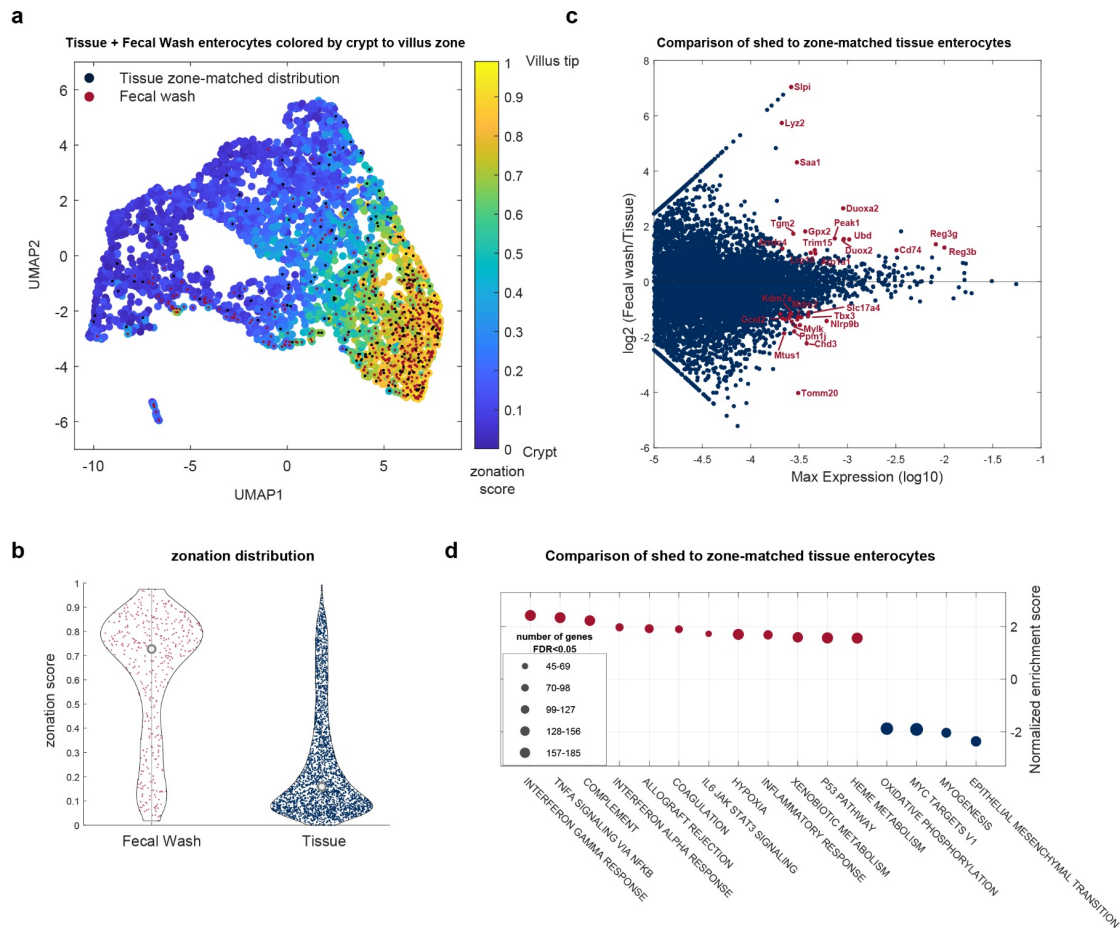


Fig. 4 – Enterocytes are predominantly shed from villi tips and elevate anti-microbial pathways. a. UMAP visualization of single enterocytes obtained from both fecal washes and jejunum tissues. Cells are colored by zonation score (Methods), which correlates with crypt-villus-tip axis location. Red dots mark fecal wash cells; Black dots mark tissue cells with distribution of zonal coordinates that matches that of shed cells zonal coordinates (Methods). **b.** zonation distribution of all fecal wash enterocytes (red dots) and tissue enterocytes (blue dots) present in **a**. Medians are marked with grey circles. **c.** MA plot of differential gene expression between fecal wash and matched distribution tissue cells. Each dot is a gene, x-axis is the log₁₀ of max expression, y-axis is log₂ ratio between mean expression in fecal wash enterocytes and tissue enterocytes, p-value is computed using Wilcoxon rank-sum tests. Names of representative up-regulated genes (q-values below 0.1) are marked in red (q-values computed using Benjamini- Hochberg FDR correction). **d.** GSEA over the Hallmark sets. Shown are all gene sets with q-value <0.05. Gene sets up regulated in shed enterocytes (red circles) included immune cell pathways, gene sets up regulated in tissue enterocytes (blue circles) included cell proliferation and respiratory pathways.

Gut microbial alteration changes the shed cell composition

We next asked if fecal wash host mRNA composition changes under perturbations that modulate microbial loads and challenge barrier function. To this end, we applied two perturbations: long term broad-spectrum antibiotics (ABX) exposure and DSS (dextran sulfate sodium) induced colitis (Fig. 5a). We compared the induction of gene expression in colon post DSS treatment (Fig. 5b) in both

tissue and fecal washes. We found that the changes in the fecal wash host transcriptomics mirrored the changes measured in the tissue (R spearman 0.29, $p=3.1e-308$). Elevated genes were consistent with ones previously reported to be induced in both mice and human colitis²⁹ (median fold change 7.9 in tissue, 4.9 in fecal washes). These genes included S100a8 and S100a9, encoding the subunits of the calprotectin protein, a known marker of gut inflammation³⁰. As well as the anti-microbial genes, Reg3g and Reg3b and immune cell markers Il1b and Cxcl2 (Extended Data Fig. 3a). ABX yielded a slightly lower, yet statistically significant correlation between the tissue and fecal wash transcriptomics (R spearman 0.11, $p=1.7e-39$, Fig. 5C). Changes including a strong decrease in the antimicrobial genes Reg3g and Reg3b in both tissue and fecal washes. Other downregulated genes were consistent with ones previously reported to be repressed in ABX treated mice³¹ (Extended Data Fig. 3b). These genes included Ang4, an antimicrobial factor involved in innate immunity³² as well as Retnlb encoding Relm β a goblet cell secreted factor that selectively target gram negative bacteria in the gut³³. Computational deconvolution indicated a significant increase in fecal wash immune cell fraction in DSS colitis and a decrease in ABX, both only seen in the colon (Fig. 5d,e, Supplementary Table 3). GSEA indicated that fecal washes from DSS mice compared to control mice were enriched in genes associated with Jak/Stat3 and Il2/Stat5 signaling, known pathways upregulated in colitis^{34,35}. These pathways were not elevated in the ABX mice fecal washes (Supplementary Table 4). Thus, fecal wash host transcriptomics mirrors pathological changes in the intestinal tissue.

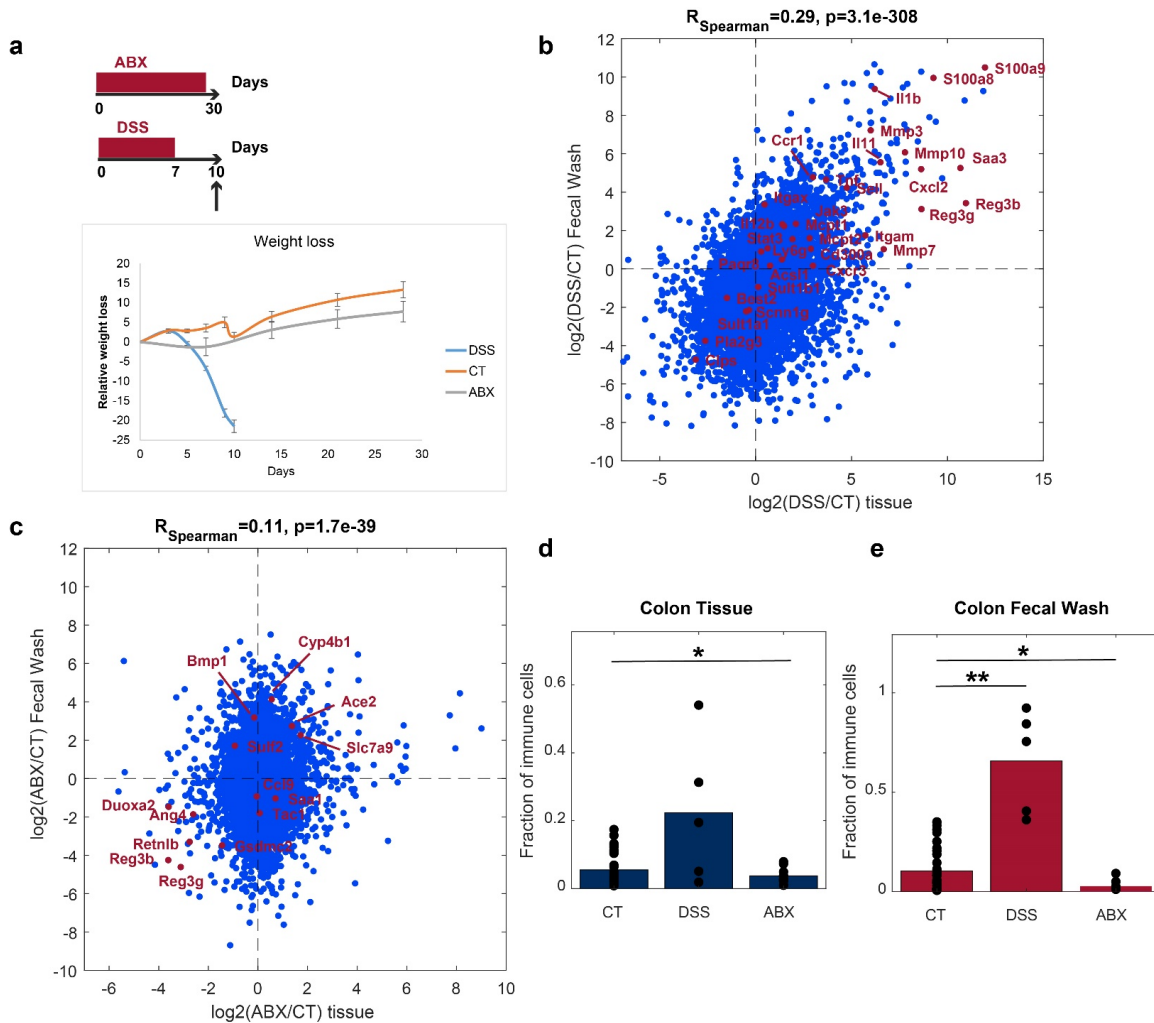


Fig. 5 – Fecal washes mirror perturbed tissue states. **a.** Time scale scheme of mice perturbations. ABX-antibiotics, DSS – dextran sulfate sodium induced colitis. Weight loss graph show changes in body weight throughout the drug administration. **b.** Scatter plot of \log_2 gene expression ratio between DSS and control mice in colon tissue (x-axis) or colon fecal washes (y-axis). Each dot is a gene. Red dots are inflammatory-related genes previously shown to be upregulated in colon tissues from DSS-colitis²⁹. Spearman correlation 0.29, p-value = $3.1e^{-308}$. **c.** Scatter plot of \log_2 gene expression ratio between ABX and control mice in colon tissue (x-axis) or colon fecal washes (y-axis). Red dots are genes previously shown to change in colon tissues from after ABX gavage treatment³⁰ Spearman correlation 0.11, p-value = $1.7e^{-39}$. **d-e.** Changes in fractions of immune cells between treatments calculated from computational-deconvolution cell compositions of colon tissue (**d**) or colon fecal wash (**e**). CT – control. *p-value<0.05 **p-value<0.01.

Discussion

The traditional view of intestinal cell turnover posits that enterocytes die immediately upon their dissociation from the tissue. Contrary to this view, our study demonstrated that shed epithelial cells remain viable for a few hours in mice. Using single cell RNA sequencing and comparison to the zoned programs of tissue enterocytes (Fig. 4), we found that enterocytes are shed predominantly from the villus tip, but are transcriptionally distinct from their tissue counterparts. Enterocytes have been shown to elevate the expression of dozens of genes at the very tip of the villi, shortly before their shedding²⁴. Our study demonstrates that they further up-regulate genes, including anti-microbial programs, after their shedding. Release of viable shed cells might serve to modulate the luminal microbial composition.

Our study demonstrated continuous shedding not only of epithelial cells, but also of diverse populations of immune cells. Trans-epithelial migration of neutrophils is a hallmark of gut inflammation³³. Using scRNAseq, smFISH and intra-vital imaging, we showed that immune cells that include both myeloid and lymphoid cells are normally shed in un-perturbed mice as well. It will be important to explore the potential functions of these cells within the lumen. We also found that immune cell shedding was modulated by the microbiome, increasing in DSS-colitis and decreasing upon broad-spectrum antibiotic treatment. The elevated immune-cell shedding in DSS-colitis, a mouse model for IBD, are in line with recent studies in human, that demonstrated elevated shedding of immune cells during IBD flares¹⁷.

Our controlled measurements of transcriptomes of tissue and shed cells along the gastrointestinal tract enabled estimates of their turnover dynamics (Fig. 1d-g). We found that segmental fecal washes represent a mixture of cells shed from the corresponding intestinal segments, as well as from up-stream segments. The cecum seems to form a barrier that potentially stalls the cells, increases their rate of death or the degradation rate of their transcriptomes, as representation of ileal segments in washes of the proximal colon was lower (Fig. 1d,f). Notably, we estimate that shed cells survive longer in the colon compared to the small intestine (Fig. 1g). Although the colon is substantially shorter than the small intestine, transit time along this segment is significantly longer²⁰⁻²², yet the distal colon had a substantial contribution from the proximal colon shed cells (Fig. 1f). The shorter lifetime of shed cells in the small intestine may be due to the high abundance of pancreatic enzymes that degrade biological material in this segment, or potentially due to the differential consistency of the luminal content between the small and large intestine.

Our study highlights an additional stage in the lifetime of intestinal cells that occurs outside of the tissue context. It will be important to use single cell approaches to study similar processes in the human gut in health and in diverse intestinal pathologies that may modulate shed cell dynamics.

These include IBD, Celiac and colorectal cancer. Our work demonstrates that shed cells in the gut mirror pathological processes, and therefore could be utilized to probe gastrointestinal biology. Our approaches for combined analysis of transcriptomes of tissue and shed cells can be applied to study turnover dynamics in other mammalian tissues³⁷.

Methods

Mice and sample collection

Mouse experiments were approved by the Institutional Animal Care and Use Committee of the Weizmann Institute of Science and were conducted in agreement with the institute guidelines. Male (C57BL/6J0laHsd) mice aged 8-12 weeks, housed under regular 12h light-dark cycle, were used in our experiments, all mice were sacrificed by cervical dislocation. For control bulk RNA experiments, the whole intestine was taken out and divided into 5 parts; the small intestine from the end of the stomach to the cecum was divided to 3 equal parts of about 8 cm each. The large intestine from the end of the cecum to the rectum was divide to 2 equal length parts of about 2 cm each. For collection of fecal wash; each of the 5 segments was flushed with 700 ul of DPBS and was snap frozen in liquid nitrogen. After flushing, a small tissue fragment of 0.5 cm was taken from the distal part of each segment and snap frozen in liquid nitrogen. For the DSS colitis model, mice were treated with 2% (w/v) DSS (M.W. = 36,000–50,000 Da; MP Biomedicals) in their drinking water for 7 days followed by regular access to water for 3 days. Mice were monitored daily for weight loss (fig. 5a) and rectal bleeding signs. Colon fecal washes and tissue samples were taken at the 10th day of experiment. Since the colon was severely inflamed, it was taken as a whole and not split into two parts. For the antibiotics model, mice were given a combination of the following antibiotics for 4 weeks, vancomycin (1 g/l), ampicillin (1 g/l), kanamycin (1 g/l), and metronidazole (1 g/l) in their drinking water for 30 days^{38–40}. All antibiotics were obtained from Sigma Aldrich. Colon fecal wash and tissue samples were taken at the 30th day of experiment.

RNA isolation

For tissue samples – snap frozen tissues were thawed in 300 µl Tri-reagent and mechanically homogenized with bead beating, followed by a short centrifugation step to pull down beads and any tissue left-overs. For fecal washes – Tri-reagent was added to sample at a ratio of 3:1, samples were allowed to thaw on ice followed by thorough mixing. A first centrifugation step was used (1 minute, 18,000 rpm) to eliminate fecal solids. Following this, ethanol was added in a ratio of 1:1 to the supernatant from the previous step and continued according to the manufacturer instructions of Direct-zol mini (for tissues) and micro (for fecal wash) prep kit (ZYMO research, R2052).

Bulk RNA sequencing of samples

RNA was processed by the mcSCRBseq protocol¹⁸ with minor modifications. For tissue RNA, RT reaction was applied on 20 ng of total RNA, for fecal wash RNA, RT reaction was started with one third of the total eluted volume with a final reaction volume of 20 μ l (1 \times Maxima H Buffer, 1 mM dNTPs, 2 μ M TSO* E5V6NEXT, 7.5% PEG8000, 20U Maxima H enzyme, 1 μ l barcoded RT primer). Subsequent steps were applied as mentioned in the protocol. Library preparation was done using Nextera XT kit (Illumina) on 0.6 ng amplified cDNA. Library final concentration of 2nM was loaded on NextSeq 500/550 (Illumina) sequencing machine aiming at 20 M reads per sample with the following setting: Read1 – 16bp, Index1 – 8bp, Read2 – 66bp.

Bioinformatics and computational analysis

Illumina output files were demultiplexed with bcl2fastq v.2.17. Resulting FASTQ files were analyzed using the zUMIs pipeline⁴¹. Reads were aligned with STAR (v.2.5.3a) to the GRCm38 genome (release 84; Ensembl) and exonic unique molecular identifier (UMI) counts were used for downstream analysis. Statistical analyses were performed with MATLAB R2021a. Mitochondrial genes and non-protein coding genes were removed from the analysis. Protein coding genes were extracted using the annotation in the Ensembl database (BioMart) for reference genome GRCm38 version 84, using the R package "biomaRt" (version 2.44.4). Gene expression for each sample was consequently normalized by the sum of the UMIs of the remaining genes that individually take up less than 5% of the sample sum. Samples with less than 100,000UMIs over the remaining genes were removed from the analyses. Clustering was performed with the MATLAB function clustergram over the Zscore-transformed expression matrix, using Spearman distances and included genes with maximal expression above 10^{-4} of summed UMIs. Differential gene expression was performed using Wilcoxon ranksum tests and Benjamini-Hochberg FDR corrections. Computational deconvolution was performed using cellanneal using signature tables obtained from a single cell atlas of small and large mouse intestines as follows (Supplementary Table 3). Villus bottom and top enterocytes were defined as the averages of the lower or upper 3 villus zones in Moor and colleagues²⁴, respectively. Goblet cells and enteroendocrine cells signatures were extracted based on Haber and colleagues¹⁹. Immune cell signatures were extracted based on Biton and colleagues⁴². All dendritic cell subtypes were averaged and coarse-grained into one group, as were monocytes and macrophages. Large intestine signature file was extracted based on Širvinskas and colleagues aged and young colon single cell atlas⁴³. Gene Set Enrichment Analysis (GSEA)²³ was performed over the Hallmark gene sets.

Mathematical model for cell shedding in the intestine

To estimate the contribution of each tissue segment to the fecal wash transcriptome (Fig. 1d) we used a multi-linear regression model:

$$W_i = \sum_{i=1}^5 a_i T_i, \quad a_i \geq 0$$

Where W_i is the i 'th fecal wash segment, T_i is the i 'th tissue segment, and a_i is a non-negative weight accounting for the contribution of the i 'th tissue segment.

To estimate shed cell life-times, a simple model was formulated, in which the wash mRNA in each segment W_i is shed from the corresponding tissue segment T_i at rate α_i , and degraded at rate δ_i :

$$\frac{dW_i}{dt} = \alpha_i T_i - \delta_i W_i$$

Where $i=1 \dots 5$ stands for duodenum, jejunum, ileum, proximal colon and distal colon respectively.

The equation was discretized, using the fact that the time axis in this problem corresponds to the spatial 1D axis along the intestinal tract:

$$\frac{W_{i+1} - W_i}{\Delta t_i} = \alpha_i T_i - \delta_i W_i$$

Where Δt_i is the time it takes for the wash to move from the i 'th segment to the $(i + 1)$ 'th segment. We took t_i to be $t_1, t_2 = 1h$, representing the transitions from duodenum to jejunum and from jejunum to ileum. $t_3, t_4 = 3h$, representing transitions from ileum to proximal colon and from proximal colon to distal colon following²⁰⁻²².

This leads to the final equation:

$$W_{i+1} = A_i T_i + B_i W_i$$

With $A_i = \alpha_i \Delta t_i$, $B_i = 1 - \delta_i \Delta t_i$.

The equation was fit for $i = 1 \dots 4$ by multi-linear regression with non-negative weights. The half-life time of the mRNA in each segment τ_i was computed from B_i by $\tau_i = \frac{\ln(2)}{\delta_i} = \frac{\ln(2)\Delta t_i}{1-B_i}$.

We excluded from the analysis genes that had expression level lower than 10^{-4} in all the tissue and wash samples. In Fig. 1f, the contribution of the current segment tissue and previous segment fecal wash were calculated as $\frac{A_i}{A_i+B_i}, \frac{B_i}{A_i+B_i}$ respectively. Violin plots in Fig. 1g and error bars in Fig. 1f were generated by bootstrapping the genes a hundred times (error bars represent standard errors of the mean).

Single-cell isolation and sorting

Single cells were obtained from fecal washes of the entire mouse small intestine from four mice in total. Collected cells were filtered through 70µm filter and washed once with FACS buffer. Cells were stained with Hoechst (15 µg/ml) in the presence of reserpine (5 µM) for elimination of bacterial populations. Cells were sorted with SORP-FACSAriaII machine (BD Biosciences) using a 100 µm nozzle. Dead cells were excluded on the basis of propidium iodide (PI) incorporation. Sorted cells were high on hoechst and low on PI (Extended Data Fig. 2). Cells were sorted into 384-well cell capture plates containing 2 µl of lysis solution and barcoded poly(T) reverse-transcription (RT) primers for single-cell RNA-seq²⁶. Barcoded single cell capture plates were prepared with a mantis automated liquid handling platform. Four empty wells were kept in each 384-well plate as a no-cell control during data analysis. Immediately after sorting, each plate was spun down to ensure cell immersion into the lysis solution, snap frozen on dry ice and stored at -80°C until processed.

MARS-Seq library preparation

Single cell libraries were prepared, as described in²⁶. Briefly, mRNA from cells sorted into MARS-Seq capture plates were barcoded and converted into cDNA and pooled using an automated pipeline. The pooled sample was then linearly amplified by T7 in vitro transcription and the resulting RNA was fragmented and converted into sequencing ready library by tagging the samples with pool barcodes and Illumina i7 barcode sequences during ligation, reverse transcription and PCR. Each pool of cells was tested for library quality and concentration was assessed as described in²⁶. Machine raw files were converted to fastq files using bcl2fastq package, to obtain the UMI counts, reads were aligned to the mouse reference genome (GRCm38.84) using zUMI package⁴¹ with the following flags that fit the barcode length and the library strandedness: -c 1-7, -m 8-15, -l 66, -B 1, -s 1, -p 16.

scRNAseq data processing

For each single cell and for each gene we first subtracted the estimated background expression. Background was calculated for each 384-well plate separately, as the mean gene expression in the four empty wells. After subtraction, negative values were set to zero. Next, cells with total UMI counts lower than 300 or higher than 10,000 and total gene counts lower than 200 were removed. We used Seurat v4.1.0 package in R⁴⁴ v4.1.2 to visualize and cluster the single cell RNAseq data (Fig. 3a, Fig. 4a). Gene expression measurements (UMIs per gene) were normalized for each cell by the summed UMI, multiplied by a scale factor 10,000, and then log-transformed. To avoid

undesired sources of variation in gene expression, we used Seurat to regress out cell-cell variation driven by total number of UMIs. Cell clustering was based on PCA dimensionality reduction using the first 15 PCs, and a resolution value of 0.8. Cell type-specific markers were used to interpret the clusters cell types. The enterocytes cell cluster was determined based on average expression of Epcam higher than 10^{-4} and of Muc2 lower than 10^{-6} fraction of UMI counts per cell. This cluster included 424 cells. This set of shed enterocytes was pooled with Epcam positive enterocytes isolated from intestinal tissue²⁸ and re-ran Seurat. For each cell in the new UMAP (Fig. 4a) zonation score was calculated along the crypt villus axis, as the fractional sum of villus tip landmark genes as in²⁴. To compare gene expression changes in fecal washes compared to tissue cells we sampled the tissue cells to match the zonation score distribution of the fecal wash enterocytes using the inverse transform sampling method⁴⁵ (Fig. 4a).

Single molecule Fluorescence in-situ Hybridization (smFISH)

Fecal wash cells from the entire small intestine were centrifuged at 500g for 5 minutes, pellet of cells was fixed with pre-chilled 4% FA for 15 minutes in 4°C followed by 2hr incubation 70% EtOH. SmFISH experiments were performed on suspended cells. Probe libraries were designed using the Stellaris FISH Probe Designer Software (Biosearch Technologies, Inc., Petaluma, CA). Hybridization was performed according to published protocol⁴⁶ with modifications for cell suspension. Briefly, cells were washed with 2× SSC (Ambion AM9765) for 5min. Cells were incubated in hybridization buffer (10% Dextran sulfate Sigma D8906, 20% Formamide, 1mg/ml E.coli tRNA Sigma R1753, 2× SSC, 0.02% BSA Ambion AM2616, 2 mM Vanadylribonucleoside complex NEB S1402S) mixed with probes overnight in a 30°C incubator. smFISH probe libraries were coupled to Cy5, TMR or Alexa594. Supplementary Table 5 includes a list of all the probe sequences used in this study. After the hybridization, cells were washed with wash buffer containing 50ng/ml DAPI for 30 min at 30°C. DAPI (Sigma-Aldrich, D9542) was used for nuclear staining. All images were performed on a Nikon-Ti-E inverted fluorescence microscope using the NIS element software AR 5.11.01.

Intravital imaging by two-photon laser scanning microscopy (TPLSM)

Mice were anesthetized with a mixture of 50 mg ketamine, 15 mg xylazine and 2.5 mg acepromazine per kg of body weight. Mice were mounted in a custom-designed heated stage, kept at 37°C and supplemented with half dose of the anesthesia above every hour. A 1 cm incision was made in the abdominal skin and the midline and proximal intestinal tissue was externalized and a 15mm incision using a cautery was done to expose the luminal surface. Tissue was immobilized to

a custom-made heated chamber using tissue adhesive (Vetbond, M3). Imaging area was hydrated with PBS using silica gel and a glass cover slip was placed at the region of interest. Intravital imaging was performed using a water-dipping objective (Zeiss 20X 1.05 NA plan objective). Zeiss LSM 880 upright microscope fitted with Coherent Chameleon Vision laser was used. Images were acquired with a femtosecond-pulsed two-photon laser tuned to 900 nm. The microscope was fitted with a filter cube containing 565 LPXR to split the emission to a PMT detector (with a 579-631 nm filter) and to an additional 505 LPXR mirror to further split the emission to 2 GaAsp detectors (with a 500-550nm filter for GFP fluorescence). Images were acquired as 33 μ m z-stacks with 3 μ m steps between each z-plane (12 frames) at minimal interval for 30-60 min. The zoom was set to 1.5, and images were acquired at 512x512 x-y resolution. CX3CR1^{GFP} mice were intravenously injected with 4 μ l Hoechst 33342 (Thermo Fisher Scientific) in 100 μ l PBS 10 min prior imaging. For TNF- α treated mice, mice were injected with 10 μ g TNF- α intraperitoneal in PBS 20 min prior to the surgery.

Acknowledgements

We thank Hagit Shapiro and Eran Elinav for help with experimental procedures of mice perturbations. S.I. is supported by the Wolfson Family Charitable Trust, the Edmond de Rothschild Foundations, the Fannie Sherr Fund, the Dr. Beth Rom-Rymer Stem Cell Research Fund, the Helen and Martin Kimmel Institute for Stem Cell Research, the Minerva Stiftung grant, the Israel Science Foundation grants No. 1486/16 and No. 3663/21, the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program grant No. 768956, the Chan Zuckerberg Initiative grant No. CZF2019-002434, the Bert L. and N. Kuggie Vallee Foundation and the Howard Hughes Medical Institute (HHMI) international research scholar award.

Data availability

All data generated in this study are available at the Zenodo repository: <https://doi.org/10.5281/zenodo.7137319>

References

1. Flier, L. G. & Clevers, H. Stem Cells, Self-Renewal, and Differentiation in the Intestinal Epithelium. *Annu. Rev. Physiol.* **71**, 241–260 (2009).
2. Sender, R. & Milo, R. The distribution of cellular turnover in the human body. *Nat. Med.* **27**, 45–48 (2021).
3. Eastwood, G. L. Gastrointestinal epithelial renewal. *Gastroenterology* **72**, 962–975 (1977).
4. Radtke, F. & Clevers, H. Self-renewal and cancer of the gut: two sides of a coin. *Science* **307**, 1904–1909 (2005).
5. Krndija, D. *et al.* Active cell migration is critical for steady-state epithelial turnover in the gut. *Science* **365**, 705–710 (2019).
6. Parker, A. *et al.* Cell proliferation within small intestinal crypts is the principal driving force for cell migration on villi. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **31**, 636–649 (2017).
7. Ngo, P. A., Neurath, M. F. & López-Posadas, R. Impact of Epithelial Cell Shedding on Intestinal Homeostasis. *Int. J. Mol. Sci.* **23**, 4160 (2022).
8. Madara, J. L. Maintenance of the macromolecular barrier at cell extrusion sites in intestinal epithelium: Physiological rearrangement of tight junctions. *J. Membr. Biol.* **116**, 177–184 (1990).
9. Guan, Y. *et al.* Redistribution of the tight junction protein ZO-1 during physiological shedding of mouse intestinal epithelial cells. *Am. J. Physiol.-Cell Physiol.* **300**, C1404–C1414 (2011).
10. Gudipaty, S. A. & Rosenblatt, J. Epithelial cell extrusion: Pathways and pathologies. *Semin. Cell Dev. Biol.* **67**, 132–140 (2017).
11. Beauséjour, M., Boutin, A. & Vachon, P. H. Anoikis and the Human Gut Epithelium in Health and Disease. in *Anoikis: How the Extracellular Matrix Regulates Life-or-Death*

Decisions (ed. Frisch, S. M.) 95–126 (Springer International Publishing, 2021).

doi:10.1007/978-3-030-73856-3_5.

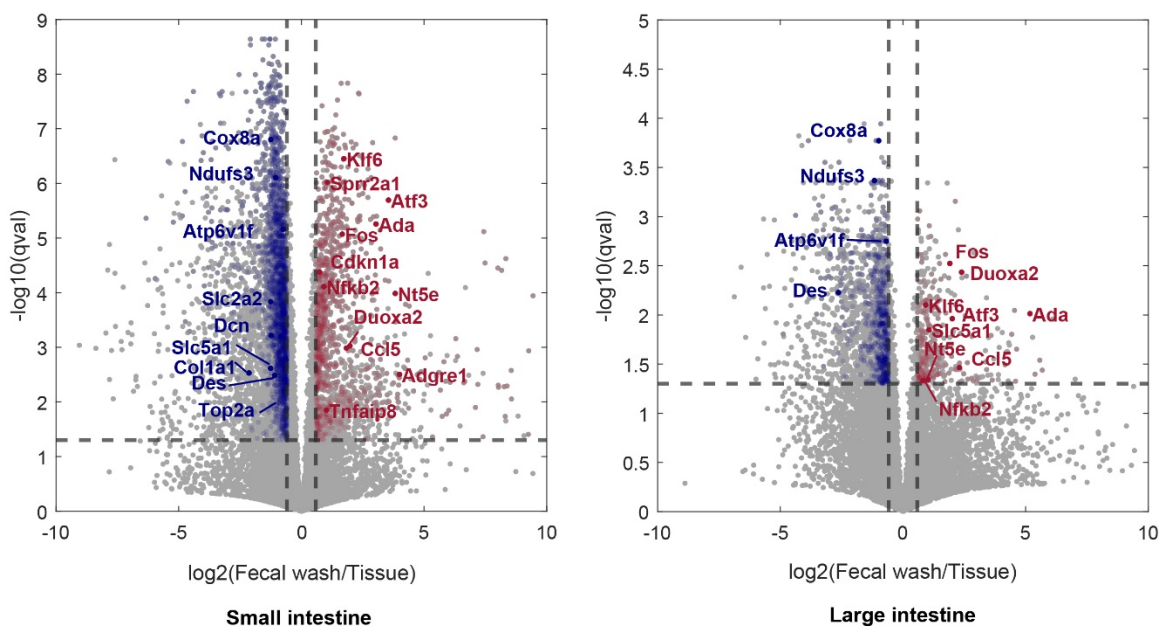
12. Patankar, J. V. & Becker, C. Cell death in the gut epithelium and implications for chronic inflammation. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 543–556 (2020).
13. Eisenhoffer, G. T. *et al.* Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia. *Nature* **484**, 546–549 (2012).
14. Chapkin, R. S. *et al.* Noninvasive stool-based detection of infant gastrointestinal development using gene expression profiles from exfoliated epithelial cells. *Am. J. Physiol. Gastrointest. Liver Physiol.* **298**, G582-589 (2010).
15. Schlaberg, R. *et al.* Fecal Host Transcriptomics for Non-invasive Human Mucosal Immune Profiling: Proof of Concept in *Clostridium difficile* Infection. *Pathog. Immun.* **3**, 164–180 (2018).
16. Whitfield-Cargile, C. M. *et al.* The non-invasive exfoliated transcriptome (exfoliome) reflects the tissue-level transcriptome in a mouse model of NSAID enteropathy. *Sci. Rep.* **7**, 14687 (2017).
17. Ungar, B. *et al.* Host transcriptome signatures in human faecal-washes predict histological remission in patients with IBD. *Gut* gutjnl-2021-325516 (2022) doi:10.1136/gutjnl-2021-325516.
18. Bagnoli, J. W. *et al.* Sensitive and powerful single-cell RNA sequencing using mcSCR-seq. *Nat. Commun.* **9**, 2937 (2018).
19. Haber, A. L. *et al.* A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017).
20. Kashyap, P. C. *et al.* Complex Interactions Among Diet, Gastrointestinal Transit, and Gut Microbiota in Humanized Mice. *Gastroenterology* **144**, 967–977 (2013).

21. Padmanabhan, P., Grosse, J., Asad, A. B. M. A., Radda, G. K. & Golay, X. Gastrointestinal transit measurements in mice with 99mTc-DTPA-labeled activated charcoal using NanoSPECT-CT. *EJNMMI Res.* **3**, 60 (2013).
22. Woting, A. & Blaut, M. Small Intestinal Permeability and Gut-Transit Time Determined with Low and High Molecular Weight Fluorescein Isothiocyanate-Dextrans in C3H Mice. *Nutrients* **10**, 685 (2018).
23. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).
24. Moor, A. E. *et al.* Spatial Reconstruction of Single Enterocytes Uncovers Broad Zonation along the Intestinal Villus Axis. *Cell* (2018) doi:10.1016/j.cell.2018.08.063.
25. Buchauer, L. & Itzkovitz, S. cellanneal: A User-Friendly Deconvolution Software for Omics Data. (2021) doi:10.48550/ARXIV.2110.08209.
26. Keren-Shaul, H. *et al.* MARS-seq2.0: an experimental and analytical pipeline for indexed sorting combined with single-cell RNA sequencing. *Nat. Protoc.* **14**, 1841–1862 (2019).
27. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
28. Bahar Halpern, K. *et al.* Lgr5⁺ telocytes are a signaling source at the intestinal villus tip. *Nat. Commun.* **11**, 1936 (2020).
29. Czarnewski, P. *et al.* Conserved transcriptomic profile between mouse and human colitis allows unsupervised patient stratification. *Nat. Commun.* **10**, 2892 (2019).
30. Khaki-Khatibi, F. *et al.* Calprotectin in inflammatory bowel disease. *Clin. Chim. Acta* **510**, 556–565 (2020).
31. Reikvam, D. H. *et al.* Depletion of Murine Intestinal Microbiota: Effects on Gut Mucosa and Epithelial Gene Expression. *PLoS ONE* **6**, e17996 (2011).

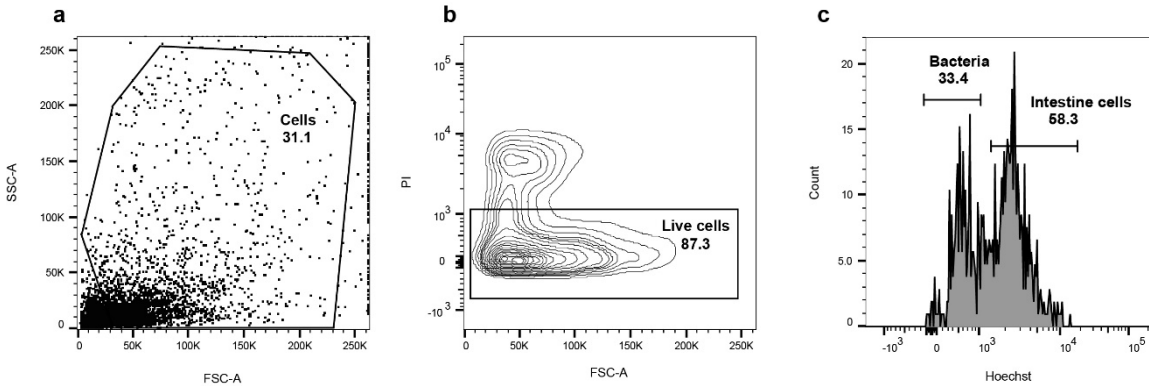
32. Angiogenins: a new class of microbicidal proteins involved in innate immunity - PubMed.
<https://pubmed.ncbi.nlm.nih.gov/12548285/>.
33. Propheter, D. C., Chara, A. L., Harris, T. A., Ruhn, K. A. & Hooper, L. V. Resistin-like molecule β is a bactericidal protein that promotes spatial segregation of the microbiota and the colonic epithelium. *Proc. Natl. Acad. Sci.* **114**, 11027–11033 (2017).
34. Cordes, F., Foell, D., Ding, J. N., Varga, G. & Bettenworth, D. Differential regulation of JAK/STAT-signaling in patients with ulcerative colitis and Crohn's disease. *World J. Gastroenterol.* **26**, 4055–4075 (2020).
35. Salas, A. *et al.* JAK–STAT pathway targeting for the treatment of inflammatory bowel disease. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 323–337 (2020).
36. Brazil, J. C. & Parkos, C. A. Pathobiology of neutrophil–epithelial interactions. *Immunol. Rev.* **273**, 94–111 (2016).
37. Vorperian, S. K., Moufarrej, M. N., Tabula Sapiens Consortium & Quake, S. R. Cell types of origin of the cell-free transcriptome. *Nat. Biotechnol.* **40**, 855–861 (2022).
38. Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S. & Medzhitov, R. Recognition of Commensal Microflora by Toll-Like Receptors Is Required for Intestinal Homeostasis. *Cell* **118**, 229–241 (2004).
39. Ichinohe, T. *et al.* Microbiota regulates immune defense against respiratory tract influenza A virus infection. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 5354–5359 (2011).
40. Fagarasan, S. *et al.* Critical Roles of Activation-Induced Cytidine Deaminase in the Homeostasis of Gut Flora. *Science* **298**, 1424–1427 (2002).
41. Parekh, S., Ziegenhain, C., Vieth, B., Enard, W. & Hellmann, I. zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. *GigaScience* **7**, (2018).
42. Biton, M. *et al.* T Helper Cell Cytokines Modulate Intestinal Stem Cell Renewal and Differentiation. *Cell* **175**, 1307-1320.e22 (2018).
43. Širvinskas, D. *et al.* Single-cell atlas of the aging mouse colon. *iScience* **25**, 104202 (2022).

44. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
45. Devroye, L. Chapter 4 Nonuniform Random Variate Generation. in *Handbooks in Operations Research and Management Science* (eds. Henderson, S. G. & Nelson, B. L.) vol. 13 83–121 (Elsevier, 2006).
46. Lyubimova, A. *et al.* Single-molecule mRNA detection and counting in mammalian tissue. *Nat. Protoc.* **8**, 1743–1758 (2013).

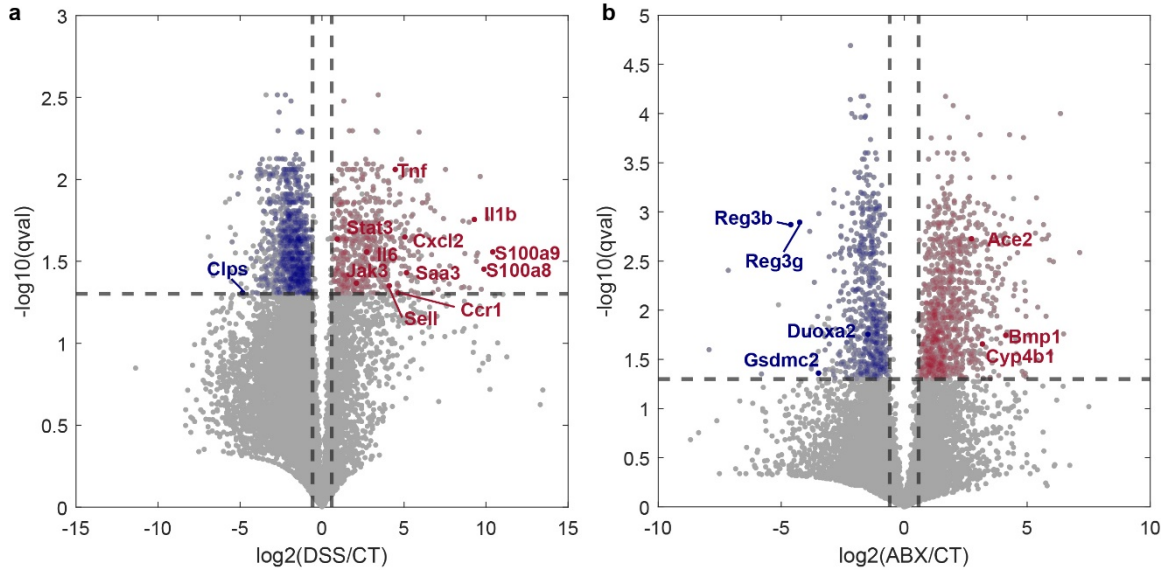
Extended Data



Extended Data Fig. 1 - Volcano plot showing differentially expressed genes between fecal washes and tissues in small (left) or large (right) intestine, each dot is a gene, x-axis is the \log_2 - ratio of expression between fecal wash and tissue, y axis is $-\log_{10}$ (q value), where p value is computed using Wilcoxon rank-sum tests and q- values computed using Benjamini- Hochberg FDR correction. Genes with corresponding q-values below 0.05 and fold change greater than 1.5 are marked in red or blue for fecal wash or tissue elevated genes respectively. Names of representative up-regulated genes are shown in red or blue for fecal wash or tissue respectively.



Extended Data Fig. 2 - Sorting strategy for isolation of mouse intestinal cells from fecal washes. **a.** FSC-A and SSC-A are used to select cells based on size. **b.** PI staining is used to gate out dead cells. **c.** Hoechst fluorescence is used to select eukaryotic cells while avoiding low DNA amount of bacteria cells. Numbers in A-C represent the percent of gated cells.



Extended Data Fig. 3 - Volcano plot showing differentially expressed genes between large intestine fecal wash in control (CT) and DSS (a) or ABX and CT (b), each dot is a gene, x-axis is the \log_2 - ratio of expression between fecal wash and tissue, y axis is $-\log_{10}(q\text{ value})$, where p value is computed using Wilcoxon rank-sum tests and q- values computed using Benjamini- Hochberg FDR correction. Genes with corresponding q-values below 0.05 and fold change greater than 1.5 are marked in red or blue for treatment or control elevated genes respectively. Names of representative up-regulated genes are shown in red or blue for treatment or control respectively.