Title: HIV-1 infection is associated with depletion of germinal center B cells and a decrease in IgA$^+$ plasma cells in the GI tract.

Authors: Francesca Cossarini[1,2], Joan Shang[2,3], Azra Krek[3], Zainab Al-taie[3], Ruixue Hou[3], Pablo Canales-Herrerias[1,4], Minami Tokuyama[1,4], Michael Tankelevich[1,4], Adam Tillowiz[1,4], Divya Jha[1,4], Alexandra E. Livanos[1,4], Louise Leyre[1,4], Mathieu Uzzan[1,4,5] , Gustavo Martinez-Delgado[1,4], Matthew Tylor[1,4], Keshav Sharma[1,4], Arno R Bourgonje[1,4], Michael Cruz[3], Giorgio Ioannou[2,6], Travis Dawson[2,6], Darwin D'Souza[2,6], Seunghee Kim-Schulze[2,6], Ahmed Akm[7], Judith A. Aberg[1], Benjamin K. Chen[1,2], Sacha Gnjatic[2,6], Alexandros D. Polydorides[7], Andrea Cerutti[2,8,9], Carmen Argmann[3], Ivan Vujkovic-Cvijin[10], Mayte Suarez-Farinas[3], Francesca Petralia[3], Jeremiah J. Faith[2,3], Saurabh Mehandru[2,3*]

[1]*Division of Infectious Diseases, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY*
[2]*Precision Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, NY*
[3]*Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY*
[4]*Division of Gastroenterology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY*
[5]Hopital Henri Mondor, APHP, Gastroenterology department, Créteil, France
[6]*Human Immune Monitoring Center, Icahn School of Medicine at Mount Sinai, New York, NY*
[7]*Department of Pathology, Molecular and Cell-Based Medicine, Icahn School of Medicine at Mount Sinai, New York, NY*
[8]*Translational Clinical Research Program, Hospital del Mar Medical Research Institute (IMIM), Barcelona, Spain.*
[9]*Catalan Institute for Research and Advanced Studies (ICREA), Barcelona, Spain*
[10] F Widjaja IBD Institute, Department of Medicine, Gastroenterology, Cedars-Sinai Medical Center, Los Angeles, CA.

[*]***Address correspondence to saurabh.mehandru@mssm.edu***

Author contributions:

FC: study concept and design; study participants enrollment, acquisition of data; analysis and interpretation of data; statistical analysis; drafting of the manuscript. JS: acquisition of data; analysis and interpretation of data; statistical analysis; drafting of the manuscript. AK: analysis and interpretation of data; statistical analysis; drafting of the manuscript. ZAT: analysis and interpretation of data; statistical analysis; drafting of the manuscript. RH: analysis and interpretation of data; statistical analysis; manuscript revision. PCH: acquisition of data analysis and interpretation of data; statistical analysis; manuscript revision MTo: acquisition of data analysis and interpretation of data; statistical analysis; drafting of the manuscript. MTa: study participants enrollment, acquisition of data, technical, or material support. AT: acquisition of data; analysis and interpretation of data; manuscript revision. DJ: acquisition of data; analysis and interpretation of data; manuscript revision. AEL: acquisition of data; analysis and interpretation of data; manuscript revision. LL: acquisition of data; analysis and interpretation of data; manuscript revision. MU: acquisition of data; interpretation of data; manuscript revision. GMD: acquisition of data, technical, or material support. MTy: acquisition of data, technical, or material support. KS: acquisition of data, interpretation of data, technical, or material support. ARB: interpretation of data, manuscript revision. MC: study participants enrollment, technical, or material support. GI: acquisition of data, interpretation of data, technical, or material support, manuscript revision. TD: acquisition of data, technical, or material support. DDS: acquisition of data, technical, or material support. SKS: interpretation of data, technical, or material support supervision. AK: acquisition of data, technical, or material support supervision. JAA: interpretation of data; manuscript revision. BKC: interpretation of data; manuscript revision. SG: interpretation of data; technical, or material support supervision, manuscript revision. ADP: acquisition of data, interpretation of data interpretation of data, manuscript revision. AC: interpretation of data; manuscript revision. CA: analysis and interpretation of data; statistical analysis; drafting of the manuscript. IVC: analysis and interpretation of data; statistical analysis; drafting of the manuscript. MSF: analysis and interpretation of data; statistical analysis; drafting of the manuscript. FP: analysis and interpretation of data; statistical analysis; drafting of the manuscript. JF: study concept and design, analysis and interpretation of data; statistical analysis; drafting of the manuscript. SM: study supervision, study concept and design, study participants enrollment, acquisition of data; analysis and interpretation of data; statistical analysis; drafting of the manuscript.

Abstract

Gastrointestinal (GI) B cells and plasma cells (PCs), critical to mucosal homeostasis, play an important role in the host response to HIV-1 infection. Here, high resolution mapping of human B cells and PCs from colon and ileum during both viremic and suppressed HIV-1 infection identified a significant reduction in germinal center (GC) B cells and Follicular Dendritic Cells (FDCs) during HIV-1 viremia. Further, IgA^+ PCs, the major cellular output of intestinal GCs were significantly reduced during viremic HIV-1 infection. PC-associated transcriptional perturbations, including type I interferon signaling persisted in antiretroviral therapy (ART) treated individuals, suggesting ongoing disruption of the intestinal immune milieu during ART. GI humoral immune perturbations associated with changes in intestinal microbiome composition and systemic inflammation. Herein, we highlight a key immune defect in the GI mucosa due to HIV-1 viremia, with major implications.

One Sentence Summary

Major perturbations in intestinal GC dynamics in viremic HIV-1 infection relate to reduced IgA^+ plasma cells, systemic inflammation and microbiota changes.

INTRODUCTION

Neutralizing antiviral antibodies (NAbs) play a critical role in host-protection against invading pathogens such as HIV-1(*1*). While some of the viral countermeasures against host NAbs, such as prevention of immune recognition and generation of escape mutants are well studied (reviewed in(*2*)), there is a relative lack of understanding of the humoral immune dynamics within the tissues with high rates of HIV-1 infection, where the NAb response is likely generated. Gut-associated lymphoid tissue (GALT) is one such site where viral replication is concentrated during acute and chronic HIV-1(*3-5*) resulting in a profound and preferential $CD4^+$ T cell depletion within the gastrointestinal (GI) tract (*6-9*). Further, GALT-associated HIV-1 replication is associated with seeding of HIV-1 reservoirs(*10-12*), disruption of the mucosal barrier integrity and systemic inflammation(*13-16*). Therefore, a detailed examination of the humoral immune system of the GALT is important to understanding HIV-1 pathogenesis and to inform HIV-1 cure efforts.

The cellular output of GALT is dominated by the generation of secretory IgA (SIgA) that regulates barrier integrity (*17-22*) and mediates mucosal homeostasis (*23*). SIgA exists in both low affinity and high affinity states(*24, 25*). While the low affinity, polyreactive IgA coats the commensal microbiota and enables symbiosis(*26-28*), high affinity IgA provides the first line of defense against pathogens. The generation of both polyreactive IgA to commensals and high affinity IgA to pathogens requires B cell-T cell interaction within the highly organized microenvironment of germinal centers (GC) (*23, 29, 30*). Within the GCs, specialized stromal cells called follicular dendritic cells (FDCs) are crucial to the maturation of the B cell response through affinity-based selection of B cells that drives the production of highly specific antibodies(*31*). B cells then egress the GC as memory B cells or as plasma cells (PCs). When primed in Peyer's patches or mesenteric lymph nodes, PCs migrate to effector sites such as the intestinal lamina propria or to the bone marrow for long term survival(*31-34*).

Intimate bi-directional cross-talk exists between SIgA and intestinal bacteria. Intestinal dysbiosis is recognized in HIV-1 infection(*35-37*), where an increase in pro-inflammatory bacteria (*35, 37-39*) and a decrease in bacterial taxa with anti-inflammatory properties (*39-42*) is reported. However, the impact of humoral perturbations on microbiota alterations is not well-elucidated in persons with HIV-1 infection (PWH). Additionally, while a number of gut-barrier homeostatic

mechanisms have been shown to be disrupted by HIV-1 infection(*21, 43-45*), B cells and PCs, key regulators of gut barrier remain understudied in PWH.

The goals of the present study were to perform a high-resolution mapping of the phenotypic and molecular landscape of the humoral immune system of the GI tract during HIV-1 infection and to examine the downstream effects of HIV-1-mediated humoral immune perturbations on the local and systemic host inflammatory environment as well as on intestinal dysbiosis.

RESULTS

Viremic HIV-1 infection is associated with a major depletion of GC B cells in the GI tract.

To dissect the dynamics of intestinal PCs and B cells in PWH, scRNA-seq analysis was performed in 9 PWH. This included 4 people with viremic HIV (PWVH) with elevated HIV viral load (VL) off antiretroviral treatment (ART) (VL; plasma HIV RNA >20 copies/ml) and 5 people with suppressed HIV (PWSH) on ART (plasma HIV RNA <20 copes/ml for at least 12 months). Additionally, 9 HIV negative controls (NC) were included **(Table 1, Table S1)**. The colon and ileum were analyzed separately to sample the immune effector and inductive sites respectively (*46*). Unsupervised cell clustering of 62365 cells with an average of 9016 Unique Molecular Identifiers (UMIs) per cell from the colon generated 23 unique clusters **(Fig. 1, A and B, Fig. S1A and Table S2),** while 43291 cells with an average of 6292 UMIs per cell from the ileum mapped to 26 unique clusters **(Fig. 1, C and D, Fig. S1B and Table S2).** Based on canonical gene expression, we identified clusters representing epithelial cells (*KRT8, SLC25A5*, *CA2*), endothelial cells (*PLVAP, PECAM*) stromal cells (*PDGFRB, CXCL14, CALD1,PMP22*) and immune cell lineages that included PCs (*JCHAIN, MZB1, XBP1*), B cells (*CD79A, BANK1*), $CD4^+$ T cells (*CD3D, CD4, TRAC*), $CD8^+$ T cells (*CD3D, CD8, TRAC*), myeloid cells (*CD14, CD68, MNDA*), innate lymphoid cells (*IL7R, KIT, IL23R*) and mast cells (*TPS, KIT)* both in the colon and ileum **(Fig. 1, E and F and Tables S3-6).** We then quantified cell type frequencies across groups **(Fig. 1, G and H, Fig. S1, C and D and Table S7).**

Next, we selected B cells (methods) and re-clustered them using an unsupervised clustering algorithm(*47*). In the colon, analysis of 3382 B cells yielded 5 clusters, while in the ileum 5095 B cells yielded 6 clusters **(Fig. 1, I and K and Table S8).** Based on canonical gene expression, we

annotated naïve B cells (*IGHM, IGHD, FCER2, no CD27*; colon cluster #1, ileum cluster #2), memory B cells (*TNFRSF13B, CD27* no *IGHM, IGHD*; colon clusters #0 and 2, ileum clusters #0, 1, 3 and 5) and GC B cells (*AICDA, BCL6, CD38, MME;* colon clusters #3, ileum clusters #4) **(Fig. 1, J and L and Tables S9-12).** Frequencies of total B cells trended lower while naïve B cells trended higher in the colon in PWVH compared to NC **(Fig. 1M and Table S13).** While overall frequencies of memory B cells were comparable between the three groups, GC B cells were significantly reduced in the colon and trended lower in the ileum in PWVH compared to PWSH and NC, both **(Fig. 1, M and N and Table S13).**

To validate scRNA-seq data in a larger cohort, we next examined alterations in B cell subsets using multiparameter flow cytometry (FC) in 18 PWVH, 49 PWSH and 87 NC **(Table 1, Table S1, Table S14).** We observed that total B cells (live $CD45^{+}CD19^{+}CD10^{-}CD38^{-}$ cells, **Fig. S2A**) trended lower in the colon and ileum, in PWVH compared to NC (**Fig. 2, A and B**) consistent with the scRNA-seq data. The frequency of naïve B cells (live $CD45^{+}CD19^{+}CD10^{-}CD38^{-}IgD^{+}$ cells, **Fig. S2A**) trended lower in the ileum in PWVH compared to NC (**Fig. 2, A and B).** Among memory B cells, there were no significant differences in frequencies between the three groups (**Fig. 2, A and B**). Consistent with the scRNA-seq data, we identified a significant reduction of GC B cells ($CD45^{+}CD3^{-}CD19^{+}CD38^{+}IgD^{-}$ **Fig. S2A)** in both colon and ileum, in PWVH when compared to NC using flow cytometry (**Fig. 2, A and B)**.

To further confirm changes in GC B cells, we performed immunohistochemistry (IHC) on intestinal biopsy tissues from the colon and the ileum where a lymphoid follicle was identified **(Table S1 and Table S15)**. The size of GCs was reduced in PWVH compared to NC, despite overall comparable lymphoid follicle size **(Fig. 2, C and D).** Additionally, in line with the scRNA-seq and flow cytometry findings, a reduction in $CD4^{-}BCL6^{+}$ cells within the GC (putative GC B cells, GC manually annotated by GI pathologist, methods) was noted in PWVH when compared to NC **(Fig. 2, C and D, Table S15).**

To better understand the loss of GC B cells in PWVH, we examined for follicular dendritic cells (FDCs), given their importance in maintaining the GC response(*31*), using CD23(*48*). While robust CD23 expression was detected in NC in putative light zones of the GC (**Fig. S3**), a major loss of CD23 expression was noted in both the colon and the ileum PWVH **(Fig. 2, C and D, Table S15).** Further, CD23 expression correlated with BCL6 expression (**Fig. 2E and S4**).

Altogether, based on scRNA-seq, flow cytometry and IHC, we identify a significant loss of GC B cells in the GI tract of PWVH that appears to be driven by disruption of the FDC network.

Total PCs and IgA$^+$ PCs are decreased in the colon during viremic HIV-1 infection and demonstrate an altered transcriptional profile both during viremic and suppressed HIV infection.

We next examined whether PC changes accompany the significant loss of GC B cells in PWVH. Using multiparametric FC we defined intestinal PCs **(Fig. S2A)** and observed a reduction in the frequency of colonic PCs (live CD45$^+$CD19$^{+/-}$CD27$^+$ CD38^{high}cells) in PWVH, when compared to PWSH and NC **(Fig. 3, A and B and Table S16).** Among the PC isotypes, the frequency of IgA$^+$ PCs was significantly reduced in PWVH compared to NC **(Fig. 3, C and D and Table S16).** Concomitantly, the frequencies of IgG$^+$ and IgM$^+$ PCs were significantly higher in PWVH when compared to the other two groups **(Fig. 3, C and D and Table S16).** PC-related changes in the ileum were less striking than that observed in the colon **(Fig. S5A and Table S16).** Among the PC isotypes in the ileum, IgM$^+$ PCs were significantly increased in PWSH compared to NC **(Fig. S5A and Table S16).**

Next, we quantified the absolute numbers of IgA$^+$ and IgA$^-$ PCs in formalin fixed paraffin embedded (FFPE) colonic biopsy sections and noted that PWVH had significantly increased numbers of IgA$^-$ PCs and a higher IgA$^-$:IgA$^+$ PC ratio when compared to NC respectively **(Fig. 3, E and F and Table S17).** Using the criteria defined by Landsverk et al(*49*), we characterized changes in short-lived (SL; live CD45$^+$CD19$^+$D27$^+$CD38^{high} cells) and long-lived (LL; live CD45$^+$CD19$^-$D27$^+$CD38^{high} cells) PCs **(Fig. S2A and Table S16)** and noted a significant decrease in both SL and LL PCs in colon of PWVH **(Fig. S5B and Table S16),** while changes in SL and LL PCs in the ileum were not significant when compared to PWSH and NC **(Fig. S5C and Table S16).** Again, when comparing PC isotypes, we observed a decrease in SL IgA$^+$ PCs frequency and an increase in SL IgM$^+$ PC frequency in both the colon and ileum in PWVH compared to NC **(Fig. S5, D and E, Table S16).** In contrast, only colonic SL IgG$^+$ PCs were increased in PWVH compared to NC **(Fig. S5, D and E, Table S16).** Among the LL PCs isotype frequencies, colonic LL IgA$^+$ PCs were decreased, while colonic LL IgM$^+$ PCs were increased in PWVH compared to

NC **(Fig. S5F, Table S16).** Notably, the LL IgM$^+$ PC subsets remained increased in PWSH compared to NC controls in the ileum **(Fig. S5G, Table S16).**

In contrast to the intestinal mucosa, and similar to prior reports(*50*), an increased frequency of PCs was observed in the peripheral blood (PB) in PWVH compared to the other two groups **(Fig. 3, G and H, Table S16).** Similar to what was observed in the colonic LP, a reduced frequency of IgA$^+$ PCs was noted in the PB. The frequency of IgM$^+$ and IgG$^+$ peripheral PCs was overall not significantly different between the three groups. **(Fig. 3I and Table S16).**

Altogether, these data demonstrate that perturbed GC output leads to a significant reduction in mucosal antibody-secreting PCs during viremic HIV-1 infection. Further, there is a switch in the PC isotype expression from the homeostatic IgA$^+$ PCs to potentially pro-inflammatory non-IgA PC subsets.

To further dissect the dynamics of intestinal PCs in PWH, we utilized two orthogonal approaches, scRNA-seq on isolated mononuclear cells from the colon and ileum LP, and bulk RNA sequencing of flow cytometrically-sorted colonic PCs. Similar to B cells, we performed sub-clustering of the PC compartment based on scRNA-seq **(Fig. 1, I and K**; details in methods). In the colon, 12 PC clusters and in the ileum 8 PC clusters were identified **(Table S18)**. Based on immunoglobulin gene expression, PCs were grouped into IgA (colon clusters #0,1,2,3,4,5,7,8,9 and ileum clusters #1,2,5,6), IgM (colon cluster #11 and ileum clusters #0,3 [mixed IgM and IgA PCs clusters]) and IgG (colon clusters#6,10 and ileum clusters #4,7) **(Fig. 4A and Tables S19-22).** To further examine PCs, we investigated pathways differentially expressed between groups of study participants (i.e., PWVH, PWSH and NC) for each PC isotype subset (FDR <0.1 **Fig. 4 B, Tables S23-25**). In colonic PCs, a significant enrichment in interferon (IFN) response (*IRF7, ISG15, IFI6, STAT1, IRF9, IRF2, OAS2, IFNAR1, STAT2*), cell cycle regulation (*NUMA1, TUBGCP2, RAD21*), apoptosis and cell fate (*CFLAR, CASP8, FADD, TNFRSF10*), antigen presentation (*HLA-DR, TAPBP, AP1G1*) and antiviral response (*PSMA7, PSMC1, APOBEC3G*) pathways was noted, especially for IgA$^+$ PCs, in PWVH. Importantly, in PWSH, significant differences in IFN response, cell cycle, apoptosis and cell metabolism persisted, especially in IgA$^+$ and IgG$^+$ PCs. Furthermore, respiratory chain pathways (*SDHA, HADHB, IDH2, FH*) were specifically upregulated in PWSH, suggesting an altered metabolic state. Pathways dysregulation

was less striking in ileum PCs **(Table S26).** These data indicate a significant dysregulation of GI PCs even during effective ART.

To further validate these findings, we FACS-sorted colonic PCs and examined their transcriptional profile via bulk RNAseq from 12 study participants (PWVH, n=3; PWSH, n=5; NC, n=4, **Table S27**). Due to the small sample size, we were underpowered to detect differentially expressed genes (at adjusted $p<0.05$). However, gene set enrichment analysis (GSEA) analysis leveraging different databases (i.e. KEGG, Reactome and Bioplanet) indicated dysregulation of multiple pathways (**Table S28**).To reduce redundancy of pathway information (more than 600 pathways were significantly enriched in the PWSH as compared to NC) we performed Sumer analysis(*51*) that uses a weighted set cover algorithm to select the fewest number of gene sets that cover all genes associated with the enriched sets. Prioritizing gene sets based on enrichment p-value, the algorithm reduced the original number of gene sets by 52–77% (methods (*51*) and **Tables S29-34**). In alignment with the scRNA-seq dataset, PCs from PWVH were enriched in cell cycle (*CHAF1A*, *SETD1A, CREBBP*), IFN signaling (*IFI6, IRF7, STAT1, ISG15, OAS2, OAS3)* and innate and adaptive immune pathways (*HLA-DRA, HLA-DQB1, CD226, KIF11*) (**Fig. 4, C and D; Tables S29-34**), indicating a pro-inflammatory response in PWVH. On the other hand, PCs from PWSH showed upregulation in protein degradation (*PSMB, PSMD)* and mitochondrial metabolism pathways (*ATP6, COX5A, SDHD*) (**Fig. 4, C and D; Tables S29-34**), again demonstrating an altered metabolic state of PCs in PWSH during effective ART.

Taken together, these data demonstrate significant perturbations in the number and transcriptional profile intestinal PCs that accompany the disrupted GC dynamics during HIV-1 viremia. Additionally, the transcriptional profile of intestinal PCs remains abnormal during ART, with persistent interferon signaling and an altered metabolic profile, further indicating that the mucosal milieu remains dysfunctional despite suppression of HIV-1 viremia.

HIV-1 induced intestinal T cell disruption associates with PC and B cell abnormalities.

Next, we studied the association of HIV-induced T cell changes with alterations in intestinal B cells and PCs. First, using scRNA-seq, we identified T cell subsets by re-clustering the T cell compartment (methods). Among 18204 T cells in colon and 17676 T cells in ileum, 9 colonic clusters and 10 ileal clusters were identified respectively **(Fig. 5, A and B, Table S35).** We annotated clusters using canonical markers and identified T cells mapping to $CD8^+$ CTLs

(*CD8, GZMM GZMA, GZMB, IFNG,* colon clusters #0, 2, 7 and ileum clusters #0,7), naïve and central memory T cells (*CCR7, CD27*, colon clusters #3 and ileum clusters #4) as well as effector memory T cells (*CD27, ICOS*, colon clusters #1, 4, 8 and ileum clusters #1, 2, 3, 6 and 9*)* (**Figure 5 C, D, Table S36-39**). Additionally, we identified regulatory T cells (Tregs; *FOXP3, MAF*) both in the colon (cluster #5) and in the ileum (cluster #8) as well as a T follicular helper cell (TFH)-like cluster (*CD40L, BCL6*) in the colon (cluster #6) and in the ileum (cluster #5). As expected, in PWH, the fraction of $CD4^+$ T cells was decreased and the fraction of $CD8^+$ T cells was increased compared to NC **(Fig. 5, E and F, Table S40).** Within the T cells subsets, we observed a decrease in naïve and central memory T cells and an increase in effector memory T cells in PWVH in both colon and ileum. In the colon, a non-significant decrease in the fraction of Tregs were observed in PWVH compared to NC **(Fig. 5, E and F, Table S40).**

Using FC, we confirmed a significant depletion of $CD4^+$ T cells in the intestinal mucosa in PWVH VL compared to NC, which did not normalize in PWSH **(Fig. S6, A and B, Table S16).** We also detected increased frequencies of activated ($CD38^+$) and cycling ($Ki67^+$) $CD4^+$ and $CD8^+$ T cells in PWVH compared to NC **(Fig. S2C; Fig. S6, C-F, Table S16).** Next, using FC-derived data, we assessed the relationship between intestinal T cells and intestinal PCs, and noted an inverse correlation between the frequency of colonic PCs and colonic IgA^+ PCs and activated and cycling T cells **(Fig. 5G, Table S16).** Conversely, a direct correlation was found between the frequency of colonic IgM^+ (and to a less extent IgG^+) PCs and activated and cycling T cell subsets **(Fig. 5G, Table S16),** suggesting an association between the local inflammatory milieu and a decreased PC frequency and the shift of PC isotype to non-IgA^+ PCs. In the ileum, a direct correlation between total B cells, naïve B cells, SM B cells and GC B cell was observed with $CD4^+$ T cells. An inverse correlation was detected between $CD8^+$ T cells and total B cells, naïve B cells, SM B cells and GC B cells **(Fig. 5H, Table S16).**

As TFH cells offer critical support for B cell maturation within the GC, we next examined for the contribution of TFH cells in the perturbed GC dynamics. In the scRNA-seq dataset, a higher proportion of colonic $CD4^+$ T cells mapping to the TFH-like cluster was noted in PWSH compared to NC. A similar, although non-significant trend was also noted for PWVH in the colon. as well as for PWH in in the ileum (**Fig. 5, E and F**). Since functional abnormalities within TFH cells are reported in PWH(*52, 53*), we examined the transcriptional profile of TFH cells using scRNA- seq. Pathway analysis on the TFH-like cell clusters revealed an increased activation of type-I IFN, IFN-

response, cell-division and apoptosis pathways in PWVH both in the ileum and in the colon **(Fig. 5I and Fig. S6H, Tables S41-42).** Similar, but somewhat less significant changes were seen in the TFH-like cell clusters of PWSH, suggesting that despite no detectable in the relative frequency of TFH-like cells, the transcriptional profile of this population is altered during HIV-1 infection.

Changes in intestinal PCs are associated with systemic inflammation.

To determine whether the observed perturbations in intestinal PCs impacted the systemic immune milieu, we examined soluble and cell-associated inflammatory parameters in the plasma samples of the study participants. As reported previously(*54-56*), we confirmed higher frequency of peripheral blood activated T cells ($CD38^+CD4^+$ and $CD38^+CD8^+$ T cells) in PWVH and, to a lesser extent in PWSH, compared to NC **(Fig. 6, A and B, Table S16).** PWVH also had an increased frequency of cycling ($Ki67^+CD4^+$ and $Ki67^+CD8^+$) T cells compared to HIV-NC **(Fig. 6 C and D, Table S16).** Colonic IgA^+ PCs inversely correlated with the frequency of activated and cycling $CD4^+$ and $CD8^+$ T cells in circulation. Conversely, colonic IgM^+ PCs were directly correlated with both activated and cycling $CD4^+$ and $CD8^+$ T cells in circulation **(Fig. 6E, Table S16).**

We then examined a panel of soluble circulating inflammatory parameters using multiplex proteomic assay in participants (n=86) where plasma samples were available. After correcting for multiple comparison, 34 analytes were found to be significantly increased during HIV-1 infection. PWVH showed higher levels of analytes relating to T cell activation (IL-17C, SLAMF1, IL10, IFN-γ, CD6, IL-15RA, IL6, EN-RAGE, IL-13, CCL20, CD8A, PD-L1, IL-20, CD244, TNFB, CCL11, FLT3L), myeloid cell activation (SLAMF1, TNF, CXCL9, IL6, EN-RAGE, CXCL10, TNFRSF9, CXCL11, CCL20, PD-L1, CCL19, IL-20, CD244, IL-12B), B cell activation (SLAMF1, CD6, CD5, PD-L1, IL-20) and NK cell activation (IFN-γ, IL-20, IL-18, CD244, TRAIL, FLT3L). Of these analytes, 17 remained increased in PWSH, including markers of T cell activation (IL-17C, SLMAF1, CD6, IL-15RA, IL6, CCL20, CD8A), myeloid cell activation (SLAMF1, TNF, CXCL9, CCL20, IL-20), B cell activation (SLAMF1) and NK cell activation (TRAIL) **(Fig. 6F, Table S43).** NC on the other end, had increased levels of analytes involved in stem cell activation (SCF), T cell and dendritic cell proliferation (TNFSF14) and cell apoptosis (CASP-8). We found an inverse association of colonic IgA^+ PCs and a direct association of IgG^+ and IgM^+ with most of the analytes that were increased during HIV infection **(Fig. 6G).**

Taken together, these results demonstrate perturbed T cell dynamics in the GI tract of PWH, with T cell activation associated with a decrease in colonic PCs and a shift towards non-homeostatic PC isotypes and an inverse correlation between IgA^+ PCs and markers of systemic inflammation during HIV-1 infection.

Discrete intestinal microbiome changes associate with distinct colonic PCs isotypes.

Given the intimate, bidirectional communication of the intestinal microbiota with the mucosal immune system(*25, 57-60*), we examined the impact of reduced IgA^+ PCs on the microbiota by sorting IgA-bound and unbound bacteria from stool suspensions (methods) in a subset of study participants (7 PWVH, 18 PWSH and 16 NC). The frequency of IgA-bound bacteria trended lower in PWVH compared to both PWSH as well as NC **(Fig. 7A, Table S44).** Similar findings were observed in a separate validation cohort (Consortium for the Evaluation and Performance of HIV Incidence Assays (CEPHIA) of the University of California, San Francisco, methods) where IgA binding of stool bacteria was analyzed for 5 PWVH, 13 PWSH and 12 NC **(Fig. S7A, Table S44).** To delineate bacterial species targeted by IgA-mediated host immune response, we performed IgA-seq analysis of sorted IgA-bound vs unbound bacteria in each of the three groups and calculated the log ratio of bound to unbound bacteria at genus level (*61*). The IgA log Palm score for *Actinomyces* (R=0.44, p=0.049) and *Eubacterium* (R=0.52, p=0.015) correlated with the log abundance of IgA^+ PCs within each individual and the IgA log Palm score for *Eubacterium* inversely correlated with the log abundance of IgG^+ PCs (**Fig. 7C** and **Table S45)** suggesting an additional homeostatic role for this genus which has previously been shown in the context of healthy gut microbiome (*62-64*).

To define more broadly the impact of the altered host immune landscape on the overall composition of the intestinal microbiota, we examined the stool metagenome of PWVH (n=7), PWSH (n=18) and NC (n=16). Among the 26 PWH where stool microbiota was analyzed, 15 participants identified as men who have sex with men (MSM) (**Table 1**). Since sexual practices may impact on the intestinal microbiome (*37, 65-67*), we specifically recruited HIV-negative study participants who identified as MSM (n=8, 50%). Beta diversity clustering by Principal Component Analysis (PCoA) revealed similar composition between PWH and HIV-NC. **(Fig. 7D).** Notably, microbiota from MSM participants clustered together and distinct from the microbiota from non-

MSM participants regardless of the HIV-infection status and the microbiome distribution of HIV-infected non-MSM participants, spanned across all other subject groups **(Fig. 7E).** Consistent with previous reports(*65-67*), a higher relative abundance of *Prevotella* and lower relative abundance of *Bacteroides* was noted in MSM participants, regardless of their HIV status **(Fig. 7F).**

When examined across all three groups (PWSH, PWVH and NC), the major bacterial phyla were represented by *Firmicutes, Bacteroides and Actinobacteria,* with marginal differences in their relative abundance across the three groups **(Fig. 7G).** No major differences across the three groups were noted in the relative abundance of different bacteria genera **(Fig. S7B).** Nonetheless, Linear discriminant analysis Effect Size (Lefse), revealed enrichment of *Enterococci* and *Mitsuokella* in PWH and enrichment of *Roseburia* and *Alistipes* in NC **(Fig. 7H and Fig S7C).** Interestingly, in the validation cohort, *Roseburia* was found to be less IgA-coated in PWH compared to NC as well **(Fig. 7I).**

Finally, we explored the relationship between mucosal PC frequency and the intestinal microbiota at the genus level. Frequency of colonic PCs were directly related to several microbial genera (**Fig. S7D and Table S46**) including *Roseburia* **(Fig. 7J),** which was found to be enriched in the NC in the Lefse analysis and more IgA-coated in our validation cohort. Similarly, IgA^+ PCs were directly associated with the relative abundance of *Bilophila, Parabacteroides, Odoribacter and Streptococcus* (**Fig. S7D** and **Table S46**), and inversely associated with the relative abundance of *Enterobacter* and *Mitsukella* (**Fig. 7J**), both significantly enriched in PWVH. Interestingly, *Odoribacter* was the only genus that was directly related to IgA^+ PCs and inversely related to IgG^+ PCs, suggesting a dichotomic relation of this genus with the mucosal PC isotypes (**Fig. S7D** and **Table S46**).

Taken together these results demonstrate reduced bacterial targeting by IgAs during viremic HIV infection and alterations in the relative abundance of specific bacteria, despite the lack of major changes in the overall microbiome composition in PWH.

DISCUSSION

The present report provides a unique insight into changes in the humoral immune compartment of the GI tract a site of enhanced HIV-1 replication (*3, 6, 10, 11*). Through high-resolution examination of intestinal PCs and B cells in PWH, we describe a major loss of GC B

cells and a reduction in the GC size in the GALT during viremic HIV infection. Additionally, we describe a loss of intestinal PCs, specifically IgA^+ PCs with a concomitant increase in IgG^+ and IgM^+ PCs during viremic HIV infection.

NAb production is a highly coordinated process requiring close, antigen-directed participation of TFH cells, GC B cells and FDCs. Triggered by antigen, and upon entering the GC response, B cells are instructed by light zone predominant TFH cells to proliferate in a MYC-dependent fashion, culminating in progressive induction of high affinity antibodies by PCs or leading to the development of B cell memory (*68-71*). GC B cell frequency relates to the magnitude of NAb titers (*72*). Comparison of colonic antibody repertoire in PWH with age- and sex-matched HIV-1-uninfected controls reveals a significantly altered antibody repertoire and significantly decreased rates of immunoglobulin somatic hypermutation (SHM) (*73*). Further, a majority of the identified broadly neutralizing antibodies are highly mutated (*74*). All of these data emphasize a pivotal role of GC B cells in the induction of NAb response. Our report therefore, highlights a fundamental injury to the high affinity NAb-generating apparatus in the GI tract in the face of ongoing virus replication. We posit that the loss of intestinal GC response likely contributes to a failure of virologic control due to impaired induction of anti-viral Nabs, resulting in the establishment of viral sanctuaries in the intestines. Our findings presented herein therefore add another important dimension to better understanding mucosal pathogenesis of HIV-1 with broad implications.

HIV-1 associated GC dysfunction is likely multifactorial. Prior data have demonstrated Nef-mediated perturbations of B cell function (*75*) including a disruption of cognate interactions at the T cell- B cell immune synapses (*76*). In addition to Nef, other viral proteins like Vpr and Vpu are implicated in reduced B cell class switching (*77*) and antigen presentation respectively (*78*). Further, the decreased GC B cell frequencies in the GALT herein could be related to TFH cell dysfunction, infiltration of GC by cytotoxic immune cells and injury to the stromal FDC network of the GC. Chronic antigen persistence within GCs, as is seen during HIV-1 infection is associated with an IL-6-enriched, cytokine driven expansion of TFH cells (*79*). Yet, TFH cells remain functionally altered in PWH and likely provide poor support to the GC reaction (*52, 53, 80*). Additionally, GC infiltration by $CD8^+$ T cells and NK cells was reported by Levesque et al (*81*), although their study did not examine intestinal GC B cells. Finally, loss or fibrosis of the stromal network of the GALT likely contributes to altered GC dynamics in HIV-1 (*82*). Reductions

in stromal fibroblastic reticular cells (FRCs), predominant in lymphoid T cell zones relate to reductions in naïve T cells (as a consequence of reductions in FRC-produced IL7 levels, (reviewed in(*82*)). Additionally, in the present report we identify a significant reduction of GALT FDCs, a stromal cell type that predominates in the GCs. Studies performed on peripheral lymph nodes early in the course of the HIV-1 pandemic identified degenerative changes in the FDC network, particularly in patients with advanced HIV-1 disease (*83, 84*). Examination of FDCs in the GALT is under-appreciated, although it may underpin intestinal B cell dysfunction as detailed in the present study. Direct infection of FDCs by HIV-1 remains unproven (*85*), while targeted killing by cytotoxic $CD8^+$ T cells and NK cells, given high HIV-1 loads within FDCs, is a favored hypothesis (*86*). Additionally, HIV-1-associated T cell and B cell immunopathology could contribute to FDC disruption, given the recently identified FDC-associated genes related to T cell regulation by FDCs (*48*). This likely sets forth a maladaptive forward feedback loop that culminates in a highly impaired GC B cell reaction at a time when it is needed the most for anti-HIV-1 immune responses. While our study does not examine peripheral lymph node and GALT-resident FDCs contemporaneously, we hypothesize an early disruption of GALT FDCs given high levels of HIV-1 antigenemia and immune cell targeting of the GALT in early stages of HIV-1 infection. Further studies pertaining to GALT FDCs will be important to better understand and potentially rescue intestinal immune responses during HIV-1 viremia.

Under homeostatic conditions, the predominant antibody to emerge from intestinal GCs is IgA (*87, 88*), making it the most abundant immunoglobulin in the intestines (*89, 90*). Inflammatory stimuli, in particular T_{H1} cytokines such as IFN-γ redirect immunoglobulin class switching towards IgG (*91*). While this is well described in the context of chronic inflammation such as in inflammatory bowel diseases (*92-95*), and in response to intestinal pathogenic infections (*96*), here we describe an increase in intestinal IgG and IgM response during viremic HIV infection. We observe that suppression of plasma HIV VL results in 'normalization' of intestinal IgA responses, consistent with a prior report by Zaunders *et al*. (*97*) but in contrast to *Planchais* and colleagues' findings (*98*). Notably, in the latter study, participants who had started treatment during acute HIV infection, showed normalization of IgA^+ B cells and PCs, suggesting that timing of initiating of ART may influence immune reconstitution (*99*).

We have also identified a number of potent anti-viral defense and pro-inflammatory programs elaborated by intestinal PCs in PWVH. These included IFN-responses, Ag presentation

and intracellular and cytokine signaling, consistent with the role of PCs in host defense. However, while numerically similar to NC, the transcriptome of intestinal PCs to remain highly perturbed in PWSH even after 5 years of fully suppressive ART. In this regard, we observed an increase in mRNA transcripts encoding for oxidative phosphorylation and mitochondrial metabolism. Changes in cellular metabolism such as increase in both glycolysis and mitochondrial metabolism have been described in T cells and monocyte/macrophages during chronic infections, including HIV-1 (*100-102*). Additionally, an increase in oxidative phosphorylation pathways has been described in an *in vitro* model of HIV infection (*103*) and might represent a compensatory mechanism to maintain adequate intracellular levels of energy. Our data demonstrating that the intestinal PC transcriptome remains significantly abnormal in PWSH, emphasize that mucosal dynamics remain perturbed even after years of apparently 'well controlled' HIV-1 infection.

We hypothesized that alterations in the intestinal PCs would result in significant shifts in the microbiota. In our data, consistent with previous reports (*65-67*), sexual practices appeared to a significant driver of differences in the microbiota composition. Additionally, our analyses pointed to important differences in the microbiota composition of PWH compared to NC. For example, the microbiome in PWVH, lacked bacterial genera and species known to be producers of butyrate (*Lachnospiraceae and Ruminococcaceae*), a short chain fatty acid (SCFA) important for mucosal integrity and health (*104-106*) while it was enriched for Mitsoukella (*36*), a bacterium associated with increased levels of pro-inflammatory metabolites (*107-109*). Furthermore, in our data, Mitsoukella was inversely associated with colonic PC frequency, suggesting bi-directional, potentially antagonistic relationship between intestinal PCs and Mitsoukella. Conversely, other butyrate-producing bacteria of the Odoribacter genus were positively associated with the frequency of the homeostatic IgA^+ PCs and inversely with the frequency of potentially pro-inflammatory IgG^+ PCs, suggesting a contribution by Odoribacter to mucosal homeostasis in patients with HIV-1 infection, as in other inflammatory diseases (*110-112*).

Roseburia a genus that is positively associated with general health (*113, 114*), was depleted in subjects with HIV-1 viremia in our cohort, consistent with prior reports (*39, 40, 115, 116*). Furthermore, in our data, *Roseburia* abundance was directly related to the frequency of colonic PCs. In validation of our findings, samples from the CEPHIA cohort (*117*) also demonstrated significantly lower IgA binding of *Roseburia* in PWH as compared to NC. Additionally, we observed a trend towards lower IgA-bound bacteria in PWVH and a higher abundance of IgA-

bound *Actinomyces* and *Eubacterium*. The abundance of IgA-bound *Eubacterium* was also inversely correlated with the abundance of IgG^+ PCs, highlighting the homeostatic properties of this genus as shown in the context of intestinal health (*62-64*). These data in aggregate demonstrate a bidirectional relationship between mucosal plasma cells and homeostatic microbiota in health that is disrupted in PWH, in confirmation of our hypothesis.

Finally, we examined the associations between mucosal PCs and measures of systemic inflammation during HIV-1 infection. Consistent with prior reports (*56*), we confirmed heightened systemic inflammation and its persistence despite effective ART. The frequency of mucosal IgA^+ PCs was inversely associated with the level of the majority of pro-inflammatory parameters analyzed. The opposite was true for IgG^+ and IgM^+ PCs, demonstrating an association between the mucosal humoral compartment and control of systemic inflammation in HIV-1 infection during ART.

In summary, the present report details previously underappreciated disruptions in intestinal B cell and PC compartments during HIV-1 infection. Significant reduction in intestinal GC B cells during viremic infection may underlie the key changes in PC frequency that were observed. Further, alterations in the frequency, isotype expression and transcriptional profile of PCs and their association with the local and systemic inflammatory environment emphasize potential contributions of the intestinal humoral response to chronic morbidity associated with HIV even during effective ART.

MATERIALS AND METHODS

Study design and ethical considerations

Participants were recruited from the Infectious disease and HIV primary care clinics as well as from the Gastroenterology clinics at the Icahn School of Medicine at Mount Sinai. Informed consent was obtained from all study participants. The study protocol was approved by the Mount Sinai institutional review board (HS#16-00512/ GCO# 16-0583). We enrolled PWVH who were either: ART-naive or within ten days of starting antiretroviral treatment; or who self-discontinued ART and had HIV RNA >1000 copies/mL at the time of enrollment. We also enrolled PWSH who had well suppressed HIV-1 RNA (<20copies/mL) for >3months while receiving ART. Isolated viral "blips" (one time HIV RNA >20 copies/mL but <1000 copies/mL) were allowed.

Additionally, we enrolled HIV-negative controls who underwent ileocolonoscopy for routine colon cancer screening or other benign conditions.

Ileocolonoscopy, Biopsy Processing and Lamina Propria Isolation

Study participants underwent ileocolonoscopic procedure which was done either for clinical indications (e.g., colorectal cancer screening) or solely for study purposes (Table 1). Mucosal biopsies were obtained using standard large cup biopsy forceps (n=30, limiting the effect of variability in sampling) and were transported in Roswell Park Memorial Institute medium (RPMI, GIBCO car #11875085) on ice and processed within 2 hours of collection.

Intestinal mononuclear cells were obtained as previously described (*118*). First, two 20 minute rounds of dissociation to remove the epithelial layer (Dissociation media: 10 ml HBSS (free of calcium and magnesium) EDTA (0.5M, pH 8, Invitrogen) and HEPES (1M, Lonza)) was followed by transfer to 10 ml of digestion solution (containing RPMI, FBS (Corning 35-010-CV), 0.005 g of Type IV collagenase (Sigma-Aldrich, cat # C5138-1G) and DNAase I (Sigma Aldrich Fine Chemicals Biosciences cat # 10104159001)) and incubated for 40min at 37°C using gentle agitation (180 rpm). The digested tissue was then further mechanically broken down by aspirating 10-times with a 10ml syringe by sequentially filtering through a 100-µm and a 40-µm cell strainer and washed with RPMI twice to obtain a single cell suspension of the intestinal lamina propria.

Multiparameter Flow Cytometry

Intestinal lamina propria cell suspension was obtained as described above and then pelleted by centrifugation for 10min at 1800rpm and 100 µl of resuspension was plated for each cell panel and stained with cell type-specific antibody panels. The cell suspension was then fixed using either 2% formaldehyde (for extracellular markers staining only) or fixation/permeabilization buffer (Invitrogen # 00-5523-00) for intracellular staining. The list of used antibodies is shown in Table S14. Samples were then acquired for flow cytometry analysis d using LSR Fortessa (BD), and data were analyzed using FlowJo v10 (Tree Star). Dead cells and doublets were excluded from all analyses. Statistical analysis was performed using Prism GraphPad software by Dotmatics. Comparison across groups was done with one-way ANOVA for non-parametric variables (Kruskal-Wallis test) with Dunn's correction for multiple comparisons. Median and interquartile range was used to plot summary data.

Immunofluorescent microscopy

FFPE-embedded intestinal biopsy tissue was cut in 5µm sections. Tissue was dewaxed in xylene and rehydrated in graded alcohol and phosphate-buffered saline (PBS). Heat-induced epitope retrieval was performed on sections submerged in target retrieval solution (Dako, S1699) in a pressure cooker. Slides were cooled to room temperature and washed in PBS. Non-specific binding was blocked with 10% goat serum for 1 hour at room temperature. Tissue was then incubated in primary antibodies diluted in 10% goat serum overnight at 4°C. Slides were washed in PBS thrice and then incubated in secondary antibody and 4′,6-diamidino-2-phenylindole (1 µg/mL) for 1 hour at room temperature. Antibodies are summarized in Table S14. Sections were washed thrice in PBS and then mounted with Fluoromount-G (Electron microscopy sciences, #1798425). Controls included, omitting primary antibody (no primary control), or substituting primary antibodies with non-reactive antibodies of the same isotype (isotype control). Sections were imaged using a Nikon Eclipse Ni microscope and digital SLR camera (Nikon, DS-Qi2).

Cell quantification: Image analysis was done in ImageJ/Fiji®. Immunofluorescence was performed on biopsies from NC (n=7), PWSH (n=7) and PWVH (n=10), IgA^+ PCs (IgA^+ $CD138^+$) and IgA^- PCs ($CD138^+$ IgA^-) were quantified per area of lamina propria (median area 0.62 mm^2, range= 0.21 mm^2 to 1.48 mm^2). CD138 is highly expressed in PCs and in several epithelial subtypes (*119*). CD138+ cells were only counted in the lamina propria, excluding positive staining in the epithelium, matching the flow cytometry data. Statistical analysis was performed using Prism GraphPad software by Dotmatics. Frequency of cells per unit area of PWVH, PWSH and NC were compared using one-way ANOVA for non-parametric variables (Kruskal-Wallis test) with Dunn's correction for multiple comparisons. Median and interquartile range was used to plot summary data.

Immunohistochemistry staining

Immunohistochemistry (ICH) staining was performed on FFPE sections in a subset of NC (n=6), PWSH (n=5) and PWVH (n=3) where lymphoid aggregates and Germinal Centers (GCs) could be identified. Slides were baked in an oven at 70°C for 60 minutes and then loaded onto an automated stainer (Leica Bond III for CD4 staining and Dako Omnis (Dako / Agilent) for BCL6 staining. Heat Induced Epitope Retrieval (HIER) was achieved with Epitope Retrieval Solution 2 (ER2) for

30 minutes at 100°C (CD4 staining) Envision Flex TRS High Solution for 30 minutes at 97°C (BCL6 staining). Ready to Use Antibody reagent was added as marker and incubated for 15 min (CD4: Leica Bio systems, clone 4B12; BCL6: Dako / Agilent, Clone PG-B6p). Detection was achieved using Leica DAB Detection Kit (Bond Polymer Refine Detection) for CD4 and Envision FLEX DAB+ (Diaminobenzidine) (Dako Omnis) for BCL6. Slides were dehydrated and mounted. Positive and internal negative controls were employed as appropriate. Lymphoid aggregates and GCs were demarcated by an expert gastrointestinal pathologist on hematoxylin and eosin (H&E)-stained sections and $BCL6^{+}$ cells were counted manually within the GCs on IHC slides. CD23 staining within the GC was measured via the *thresholder* function in QuPath®. After color deconvolution, a threshold of 0.25 was set for positive staining. The area of positive staining was calculated as % of GC area. Statistical analysis was performed using Prism GraphPad software by Dotmatics. Comparison across groups was performed by one-way ANOVA for non-parametric variables (Kruskal-Wallis test) with Dunn's correction for multiple comparisons. Median and interquartile range was used to plot summary data.

Lamina propria cell sorting and mRNA extraction

To further characterize the transcriptional signature of intestinal mucosa PCs, lamina propria Single cell suspension obtained as described above, was stained with a dedicated Ab cocktail for B cell lineage and PCs (Live $CD45^{+}CD38^{high}CD27^{+}CD19^{+}$ cells) were sorted using BD FACS Aria II system. Sorted cells were then pelleted and RNA was isolated using RNeasy Plus Micro extraction Kit Qiagen # 74034). Briefly, the lysate was transferred to a gDNA eliminator spin column and centrifuged. Flow through was collected while the column was discarded. 1 volume of 70% ethanol was added to the flow-through and mixed well by pipetting. The sample was transferred to a RNeasy MinElute spin column and centrifuged. The flow-through was discarded and 700 µl Buffer RW1 was added to the RNeasy MinElute spin column, centrifuged. 500 µl Buffer RPE was added to the RNeasy MinElute spin column and centrifuged. Flow through was discarded and 500 µl of 80% ethanol added to the RNeasy MinElute spin column and centrifuged. The collection tube was discarded, and a dry spin was run for 5min to dry the membrane. RNeasy MinElute spin column was placed in a new 1.5 ml collection tube. RNase-free water was added, and the collection tube was centrifuged for 1 min at full speed to elute the RNA.

RNA libraries were prepared using either SMART-Seq v4 Ultra Low Input Nextera XT or SureSelect XT RNA Direct v6 (**Table S27**)

Bulk RNA sequencing analysis of flow-sorted PCs

*Pre-processing***:** Reads from FASTQ format files were mapped against Homo sapiens reference genome hg38 using STAR aligner. FastQC and RseQC were used to check quality of raw read and aligned reads. The mapping results were analyzed by featureCounts of the Subread package to quantify transcript reads. Features with fewer than 4 normalized counts across all samples were filtered out. The counts were normalized by using voom (*120*) which estimates the mean variance relationship and generates a precision weight for individual normalized observations. Normalized log counts per million and associated precision weights were then entered into the limma analysis pipeline. (*121-123*)which takes into account the paired structure of samples, as performed for microarrays and RT-PCR.

Statistical analysis: RNA-seq data were adjusted for RIN score using a linear model(*122-124*). Expression profiles were modeled using linear mixed-effects models including age and group variable (NC, PWSH, PWVH) as fixed effects, and a random intercept for each subject. Gene Set Enrichment Analysis (GSEA) was performed using clusterProfiler r package to identify the enriched terms (p.adjusted<0.05) in Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome and Bioplanet, for the comparison of the three groups (NC, PWSH, PWVH). GSEA resulted in numerous enrichments indicating dysregulation of a large number of pathways: 659 for suppressed vs control; 320 for PWVH vs NC; 640 for PWVH vs PWSH **(Table S28)**. To summarize and consolidate GSEA results, pathways with padj<0.05 were selected from each database into the sumer r package(*51*), which was used to select the most representative pathways. This process reduced the number of pathway terms organized to various clusters as follows: for PWVH vs NC 5 clusters of 87 pathways, for PWSH vs NC 8 clusters of 143 Pathways and for PWVH vs PWSH 4 clusters of 67 pathways **(Tables S29-34).** The top 5 pathways (by adj P value) per cluster were subselected for further analysis. Gene set variance analysis (GSVA) (z-score method) was performed on the HIV gene expression data (RIN adjusted) to generate a pathway score with the core enriched genes for each pathway selected by the sumer package. Bubble plot of Net Expression Score (NES) of selected pathway scores across groups was constructed. The

expression (RIN adjusted) levels of selected Leading Edge Genes (LEGs) across each sample of each group were displayed in a summary heatmap. The color of the heatmap in each row is based on the Z score values calculated by centering and scaling the data by standard deviation using the formula (X ~m)/SD, where X is individual values, and m is the mean of the row.

Lamina Propria processing for single cell RNA sequencing:

Lamina propria mononuclear cell suspension was obtained from freshly processed intestinal biopsies as detailed above and dead cells were removed by magnetic beads using Dead Cell Removal by magnetic binding (Miltenyi Biotec# 130-090-101) as per manufacturer instruction. Single cell suspensions were counted using a Nexcelom Cellometer Auto2000 and each loaded on one lane of the 10x Genomics NextGem 3'v3.1 assay as per the manufacturer's protocol with a targeted cell recovery of 8,000 cells. Gene expression libraries were prepared as per 10x Genomics demonstrated protocol(*125*). Libraries were quantified via Agilent 2100 hsDNA Bioanalyzer and KAPA library quantification kit (Roche Cat# 0796014001) and sequenced at a targeted depth of 25,000 reads per cell. Libraries were pooled in equimolar concentration and sequenced on the Illumina NovaSeq 100 cycle kit with run parameters set to 28x8x0x60 (R1xi7xi5xR2).

Single-cell RNA data processing and analysis

Filtering, Normalization, Batch correction: Cells were filtered based on the total number of UMI counts and mitochondria genes fractions. Only cells with total UMI counts greater than 1,000 and mitochondria gene fraction less than 10% were considered for the analysis. The analysis of single-cell RNA data including data normalization and batch effect correction was performed via the Seurat package(*47, 126*). First, each cell was normalized independently using function *NormalizeData* which divides feature counts for each cell by the total counts for that cell. Data was then natural log transformed. Next, we found variable features using function FindVariableFeatures(*47, 126*).This function first fits a line to the relationship of log(variance) and log(mean) using local polynomial regression; then, standardizes the feature values using the observed mean and expected variance (given by the fitted line). Feature variance was then calculated based on the standardized values. Finally, we selected features which were repeatedly variable across datasets for integration via function *SelectIntegrationFeatures(47, 126).* We then identified anchors using the *FindIntegrationAnchors* function, which takes a list of Seurat objects

as input, and use these anchors to integrate different datasets together via the *IntegrateData* function(*47, 126*). Dimensionality reduction to identify anchors was performed using reciprocal principal component analysis. For visualization, the dimensionality of each dataset was further reduced using Uniform Manifold Approximation and Projection (UMAP) implemented with Seurat functions *RunUMAP*. The PCs used to calculate the embedding were as the same as those used for clustering.

Clustering: Clustering was performed via function *Find.Cluster* available in the Seurat package, with resolution parameter r=0.5. This function identifies clusters of cells using a shared nearest neighbor algorithm. Once that unsupervised clustering was performed, we identified cluster-specific markers using function *FindMarkers* available in the *Seurat* package(*47, 126*). Markers were ranked based on area under the receiver operating characteristic curve (AUC). Based on the cluster-specific markers and leveraging existing signatures from the literature(*94, 127-131*), clusters were annotated into 9 cell-types for the ileum and the colon.

Cell-type signatures:

We considered the lineage cell-type categories to obtain gene-signature from the single-cell RNA data. This was performed via function *FindMarkers* available in the Seurat package. Only markers with an AUC greater than 0.7 were considered as cell-type specific markers. Mitochondrial and ribosomal protein genes were removed in the final list of genes since they were found expressed in multiple cell-types(*47*).

Re-clustering of B cells, PCs and T cells:

B-cells, PCs and T cells were re-clustered as smaller individual subsets using the same nearest neighborhood clustering algorithm(*47*). The resolution parameters used in re-clustering were 0.5 for PCs and T cells and 0.4 for B cells. After identifying markers for each of the subclusters we could assign them to specific subtypes of each cell type: naïve, B cells, memory B cells, GC B cells; and IgA, IgM and IgG isotypes for PCs.

Quantification of cell-type fractions:

Cell types from primary clusters were quantified as a fraction of total cells in each sample. Cell types from subclusters were quantified as a fraction of the corresponding primary cell type in each sample unless otherwise specified. Comparisons across groups were performed using Mann-Whitney U test.

Pathway enrichment analysis:

For each cell-type compartment (i.e., B cells and PCs), the log fold-change was computed between the average gene expression in cells derived from PWVH, PWSH and NC. Enrichment scores and adjusted p values (FDR) were derived via gene set enrichment analysis(*132*) for PWVH vs NC and PWSH vs NC separately. Specifically, genes were ordered in decreasing order of fold-change and a Kolmogorov-Smirnov statistics was utilized to find pathways significantly enriched at the top of the list. A permutation-based strategy was utilized to derive p-values and FDR for different pathways(*132*). Molecular pathway gene sets used in this analysis were from the Reactome database(*133*).

Stool metagenomics

Stool samples were collected from study participants before or at least two weeks after the colonoscopy procedure to avoid changes due to the colonic preparation. Fresh samples were frozen at -80C. Frozen samples were then mechanically pulverized. Fecal DNA for metagenomic sequencing was extracted from approximately 50mg of pulverized feces using physical lysis by bead beating in phenol:chloroform:isoamyl alcohol protocol(*134*). DNA was quantified using the Broad Range or High Sensitivity Quant-IT dsDNA Assay kit (Thermo Fisher,Q32853 and Q33130) with a BioTek Synergy HTX Multi-Mode Reader. Using NEBNext Ultra II DNA Library Prep Kit, metagenomic libraries were prepared. DNA that was fragmented by Diagenode Bioruptor Pico solictor were subjected to end repair, Illumina adaptor ligation, and purification with Beckman Coulter SPRI beads. NEBNext Ultra Q5 Master Mix was utilized to amplify ligated product using custom i5 and i7 index primer and Beckman Coulter AMPure XP beads quantified the final product before final sequencing using Samples were then sequenced with an Illumina HiSeq (paired-end 150 bp).

Computational analysis of metagenomic profiles

To process and taxonomically profile whole-metagenome shotgun samples, MetaPhlAn2 was used(*135*). To ensure high quality reads were sequenced, any metagenomic sample with reads of less than 100,000 was filtered out. Uncharacterized phylae and taxa that did not meet the minimum prevalence threshold of 3% of all samples were also filtered out. Linear discriminant analysis (LDA) effect size (LEfSe) was implemented to estimate bacterial candidates associated with study participants groups with a logarithmic LDA cut off ≥2.0. (*136*). To conduct and visualize LEfse analysis, we utilized the R package *microbial.* Beta diversity (using Bray-Curtis distances) was calculated and visualized using with a principal coordinate analysis (PCoA) plot. Distance and diversity index calculations were performed using *phyloseq(137)*. Spearman correlation was performed on bacterial genus relative abundances and immune cell abundances. Given the exploratory nature of this study, reported p-values are unadjusted.

IgA bound and unbound stool bacteria sorting

Fifty mg of frozen stool samples were resuspended in PBS (10x stool weight) and then pelleted via gentile spin. The supernatant was then filtered through 40nm cell strainer. 100µL of suspension was blocked for non-specific staining using anti-mouse serum and stained for IgA (PE Mylteny Biotec, Cat#130-113-476), washed three times and stained for DNA (SyborGreen Applied Biosciences). IgA-bound (SyborGreen+, IgA^+) and IgA-unbound (SyborGreen+, IgA^-) stool bacteria were then sorted using the BD FACS Aria II system. Sorted cells were then pelleted, and DNA was extracted using QIAamp UCP Pathogen Mini Kit (Qiagen Cat# 50214).

16S rRNA Sequencing and IgA-Seq

For DNA extraction, we utilized phenol:chloroform method on 50-200 mg of fecal specimen as previously described(*138*). We adjusted DNA templates to a concentration of 2 ng/mL. Following this, the V4 variable region of the 16S rRNA gene underwent PCR amplification using indexed primers, adopting the methodology outlined in(*139*). The resultant unique 16S rRNA V4 amplicons were merged and subsequently purified using AMPure XP beads (Beckman Coulter, A63880) at a bead-to-PCR reaction ratio of 1:1. Gel electrophoresis was employed to confirm the accurate amplicon size and to ensure the non-presence of primer dimers. The consolidated samples were then sequenced on an Illumina MiSeq platform, producing paired-end reads of 250bp each.To process and denoise paired-end sequence data, the DADA2 algorithm available within the QIIME2

pipeline was used. The Greengenes sequences database was used as a reference for pair-end reads merge and alignment.

To ensure that analysis included only high-quality samples, we first removed samples (and their associated IgA^+ or IgA^- pair) with less than 1000 reads. Ambiguously annotated phylum-level taxonomic features (e.g. "NA") and low prevalence ASVs were removed. Additionally, ASVs less than 0.01% relative abundance in both IgA bound and unbound dataset were filtered out. A pseudocount of minimum(relative abundance of genus)/2 was computed to entries with exact 0 relative abundance entries in ASV tables. Log Palm score was calculated (*61*) to quantify IgA binding. Spearman correlation was calculated on immune cell abundances and log Palm score of bacterial genus.

In addition, we also included a validation IgA-coating and IgA-seq cohort (Consortium for the Evaluation and Performance of HIV Incidence Assays (CEPHIA) of the University of California, San Francisco, methods) in which samples were processed as previously described (*117*). To obtain genus-level IgA scores for all taxa, the DADA2 pipeline was applied and taxonomy was assigned on resulting amplicon sequence variants (ASV) using the SILVA database (v138.1, nr99). ASVs belonging to the same genus were summed, and a pseudocount of 1 was applied. The IgA score was computed as a natural log of the ratio of each genus in the IgA-positive fraction over the IgA-negative fraction.

Multiplexed Proteomic Assay (Olink)

Circulating inflammatory biomarkers were measured using a multiplexed proteomic inflammation panel (Olink®). This platform uses an antibody-mediated proximity extension-based assay to simultaneously quantify 92 inflammation-related proteins. The level of each biomarker was compared between PWVH and NC as well as between PWSH and NC using Student's *t* test. The Benjamini-Hochberg procedure for multiple comparisons was used to adjust the resulting p value. Correlation with LP cell type frequencies and circulating inflammation-related proteins was obtained by Spearman's r correlation coefficient.

References

1. D. Corti, A. Lanzavecchia, Broadly neutralizing antiviral antibodies. *Annu Rev Immunol* **31**, 705-742 (2013).
2. D. R. Burton, L. Hangartner, Broadly Neutralizing Antibodies to HIV and Their Role in Vaccine Design. *Annu Rev Immunol* **34**, 635-659 (2016).
3. M. Guadalupe *et al.*, Severe CD4+ T-cell depletion in gut lymphoid tissue during primary human immunodeficiency virus type 1 infection and substantial delay in restoration following highly active antiretroviral therapy. *J Virol* **77**, 11708-11717 (2003).
4. Q. Li *et al.*, Peak SIV replication in resting memory CD4+ T cells depletes gut lamina propria CD4+ T cells. *Nature* **434**, 1148-1152 (2005).
5. R. S. Veazey *et al.*, Gastrointestinal tract as a major site of CD4+ T cell depletion and viral replication in SIV infection. *Science* **280**, 427-431 (1998).
6. J. M. Brenchley *et al.*, T-cell subsets that harbor human immunodeficiency virus (HIV) in vivo: implications for HIV pathogenesis. *J Virol* **78**, 1160-1168 (2004).
7. J. J. Mattapallil *et al.*, Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection. *Nature* **434**, 1093-1097 (2005).
8. S. Mehandru *et al.*, Primary HIV-1 infection is associated with preferential depletion of CD4+ T lymphocytes from effector sites in the gastrointestinal tract. *J Exp Med* **200**, 761-770 (2004).
9. S. Mehandru *et al.*, Lack of mucosal immune reconstitution during prolonged treatment of acute and early HIV-1 infection. *PLoS Med* **3**, e484 (2006).
10. A. Chaillon *et al.*, HIV persists throughout deep tissues with repopulation from multiple anatomical sources. *J Clin Invest* **130**, 1699-1712 (2020).
11. J. D. Estes *et al.*, Defining total-body AIDS-virus burden with implications for curative strategies. *Nat Med* **23**, 1271-1276 (2017).
12. S. A. Yukl *et al.*, Differences in HIV burden and immune activation within the gut of HIV-positive patients receiving suppressive antiretroviral therapy. *J Infect Dis* **202**, 1553-1561 (2010).
13. J. M. Brenchley *et al.*, Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat Med* **12**, 1365-1371 (2006).
14. A. K. Steele *et al.*, Contribution of intestinal barrier damage, microbial translocation and HIV-1 infection status to an inflammaging signature. *PLoS One* **9**, e97171 (2014).
15. A. S. Zevin, L. McKinnon, A. Burgener, N. R. Klatt, Microbial translocation and microbiome dysbiosis in HIV-associated immune activation. *Curr Opin HIV AIDS* **11**, 182-190 (2016).
16. G. Ancona *et al.*, Long-Term Suppressive cART Is Not Sufficient to Restore Intestinal Permeability and Gut Microbiota Compositional Changes. *Front Immunol* **12**, 639291 (2021).
17. J. D. Estes *et al.*, Damaged intestinal epithelial integrity linked to microbial translocation in pathogenic simian immunodeficiency virus infections. *PLoS Pathog* **6**, e1001052 (2010).
18. A. J. Macpherson, E. Slack, The functional interactions of commensal bacteria with intestinal secretory IgA. *Curr Opin Gastroenterol* **23**, 673-678 (2007).
19. M. A. McGuckin, S. K. Linden, P. Sutton, T. H. Florin, Mucin dynamics and enteric pathogens. *Nat Rev Microbiol* **9**, 265-278 (2011).
20. B. Corthesy, Multi-faceted functions of secretory IgA at mucosal surfaces. *Front Immunol* **4**, 185 (2013).
21. M. Somsouk *et al.*, Gut epithelial barrier and systemic inflammation during chronic HIV infection. *AIDS* **29**, 43-51 (2015).
22. B. Pietrzak, K. Tomela, A. Olejnik-Schmidt, A. Mackiewicz, M. Schmidt, Secretory IgA in Intestinal Mucosal Secretions as an Adaptive Barrier against Microbial Cells. *Int J Mol Sci* **21**, (2020).
23. J. Kabbert *et al.*, High microbiota reactivity of adult human intestinal IgA requires somatic mutations. *J Exp Med* **217**, (2020).
24. A. J. Macpherson, K. D. McCoy, F. E. Johansen, P. Brandtzaeg, The immune geography of IgA induction and function. *Mucosal Immunol* **1**, 11-22 (2008).
25. O. Pabst, E. Slack, IgA and the intestinal microbiota: the importance of being specific. *Mucosal Immunol* **13**, 12-21 (2020).
26. E. Slack *et al.*, Innate and adaptive immunity cooperate flexibly to maintain host-microbiota mutualism. *Science* **325**, 617-620 (2009).
27. E. Slack, M. L. Balmer, J. H. Fritz, S. Hapfelmeier, Functional flexibility of intestinal IgA - broadening the fine line. *Front Immunol* **3**, 100 (2012).
28. J. J. Bunker *et al.*, Natural polyreactive IgA antibodies coat the intestinal microbiota. *Science* **358**, (2017).
29. R. Elgueta *et al.*, Molecular mechanism and function of CD40/CD40L engagement in the immune system. *Immunol Rev* **229**, 152-172 (2009).

30. A. Reboldi *et al.*, IgA production requires B cell interaction with subepithelial dendritic cells in Peyer's patches. *Science* **352**, aaf4822 (2016).
31. G. D. Victora, M. C. Nussenzweig, Germinal Centers. *Annu Rev Immunol* **40**, 413-442 (2022).
32. J. P. Pereira, L. M. Kelly, J. G. Cyster, Finding the right niche: B-cell migration in the early phases of T-dependent antibody responses. *Int Immunol* **22**, 413-419 (2010).
33. T. A. Schwickert *et al.*, A dynamic T cell-limited checkpoint regulates affinity-dependent B cell entry into the germinal center. *J Exp Med* **208**, 1243-1252 (2011).
34. J. G. Cyster, C. D. C. Allen, B Cell Responses: Cell Interaction Dynamics and Decisions. *Cell* **177**, 524-540 (2019).
35. S. M. Dillon *et al.*, An altered intestinal mucosal microbiome in HIV-1 infection is associated with mucosal and systemic immune activation and endotoxemia. *Mucosal Immunol* **7**, 983-994 (2014).
36. C. A. Lozupone *et al.*, Alterations in the gut microbiota associated with HIV-1 infection. *Cell Host Microbe* **14**, 329-339 (2013).
37. I. Vujkovic-Cvijin *et al.*, Dysbiosis of the gut microbiota is associated with HIV disease progression and tryptophan catabolism. *Sci Transl Med* **5**, 193ra191 (2013).
38. D. M. Dinh *et al.*, Intestinal microbiota, microbial translocation, and systemic inflammation in chronic HIV infection. *J Infect Dis* **211**, 19-27 (2015).
39. J. F. Vazquez-Castellanos *et al.*, Altered metabolism of gut microbiota contributes to chronic immune activation in HIV-infected individuals. *Mucosal Immunol* **8**, 760-772 (2015).
40. E. A. Mutlu *et al.*, A compositional look at the human gastrointestinal microbiome and immune activation parameters in HIV infected subjects. *PLoS Pathog* **10**, e1003829 (2014).
41. C. L. Monaco *et al.*, Altered Virome and Bacterial Microbiome in Human Immunodeficiency Virus-Associated Acquired Immunodeficiency Syndrome. *Cell Host Microbe* **19**, 311-322 (2016).
42. W. Lu *et al.*, Association Between Gut Microbiota and CD4 Recovery in HIV-1 Infected Patients. *Front Microbiol* **9**, 1451 (2018).
43. J. M. Brenchley *et al.*, Differential Th17 CD4 T-cell depletion in pathogenic and nonpathogenic lentiviral infections. *Blood* **112**, 2826-2835 (2008).
44. V. Cecchinato *et al.*, Altered balance between Th17 and Th1 cells at mucosal sites predicts AIDS progression in simian immunodeficiency virus-infected macaques. *Mucosal Immunol* **1**, 279-288 (2008).
45. S. M. Tugizov *et al.*, HIV-associated disruption of mucosal epithelium facilitates paracellular penetration by human papillomavirus. *Virology* **446**, 378-388 (2013).
46. P. Brandtzaeg, H. Kiyono, R. Pabst, M. W. Russell, Terminology: nomenclature of mucosa-associated lymphoid tissue. *Mucosal Immunol* **1**, 31-37 (2008).
47. T. Stuart *et al.*, Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821 (2019).
48. B. A. Heesters *et al.*, Characterization of human FDCs reveals regulation of T cells and antigen presentation to B cells. *J Exp Med* **218**, (2021).
49. O. J. Landsverk *et al.*, Antibody-secreting plasma cells persist for decades in human intestine. *J Exp Med* **214**, 309-317 (2017).
50. C. M. Buckner *et al.*, Characterization of plasmablasts in the blood of HIV-infected viremic individuals: evidence for nonspecific immune activation. *J Virol* **87**, 5800-5811 (2013).
51. S. R. Savage, Z. Shi, Y. Liao, B. Zhang, Graph Algorithms for Condensing and Consolidating Gene Set Analysis Results. *Mol Cell Proteomics* **18**, S141-S152 (2019).
52. M. Perreau *et al.*, Follicular helper T cells serve as the major CD4 T cell compartment for HIV-1 infection, replication, and production. *J Exp Med* **210**, 143-156 (2013).
53. E. Moysi *et al.*, Altered immune cell follicular dynamics in HIV infection following influenza vaccination. *J Clin Invest* **128**, 3171-3185 (2018).
54. Z. Liu *et al.*, CD8+ T-lymphocyte activation in HIV-1 disease reflects an aspect of pathogenesis distinct from viral burden and immunodeficiency. *J Acquir Immune Defic Syndr Hum Retrovirol* **18**, 332-340 (1998).
55. J. V. Giorgi *et al.*, Shorter survival in advanced human immunodeficiency virus type 1 infection is more closely associated with T lymphocyte activation than with plasma virus burden or virus chemokine coreceptor usage. *J Infect Dis* **179**, 859-870 (1999).
56. I. Sereti, M. Altfeld, Immune activation and HIV: an enduring relationship. *Curr Opin HIV AIDS* **11**, 129-130 (2016).
57. L. A. van der Waaij, P. C. Limburg, G. Mesander, D. van der Waaij, In vivo IgA coating of anaerobic bacteria in human faeces. *Gut* **38**, 348-354 (1996).

58. K. Suzuki *et al.*, Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut. *Proc Natl Acad Sci U S A* **101**, 1981-1986 (2004).
59. C. Gutzeit, G. Magri, A. Cerutti, Intestinal IgA production and its role in host-microbe interaction. *Immunol Rev* **260**, 76-85 (2014).
60. A. Nakajima *et al.*, IgA regulates the composition and metabolic function of gut microbiota by promoting symbiosis between bacteria. *J Exp Med* **215**, 2019-2034 (2018).
61. N. W. Palm *et al.*, Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* **158**, 1000-1010 (2014).
62. S. H. Duncan *et al.*, Reduced dietary intake of carbohydrates by obese subjects results in decreased concentrations of butyrate and butyrate-producing bacteria in feces. *Appl Environ Microbiol* **73**, 1073-1078 (2007).
63. T. S. Ghosh *et al.*, Mediterranean diet intervention alters the gut microbiome in older people reducing frailty and improving health status: the NU-AGE 1-year dietary intervention across five European countries. *Gut* **69**, 1218-1228 (2020).
64. G. Jakobsdottir, J. Xu, G. Molin, S. Ahrne, M. Nyman, High-fat diet reduces the formation of butyrate, but increases succinate, inflammation, liver fat and cholesterol in rats, while dietary fibre counteracts these effects. *PLoS One* **8**, e80476 (2013).
65. M. Noguera-Julian *et al.*, Gut Microbiota Linked to Sexual Preference and HIV Infection. *EBioMedicine* **5**, 135-146 (2016).
66. C. F. Kelley *et al.*, The rectal mucosa and condomless receptive anal intercourse in HIV-negative MSM: implications for HIV transmission and prevention. *Mucosal Immunol* **10**, 996-1007 (2017).
67. A. J. S. Armstrong *et al.*, An exploration of Prevotella-rich microbiomes in HIV and men who have sex with men. *Microbiome* **6**, 198 (2018).
68. A. D. Gitlin, Z. Shulman, M. C. Nussenzweig, Clonal selection in the germinal centre by regulated proliferation and hypermutation. *Nature* **509**, 637-640 (2014).
69. A. D. Gitlin *et al.*, Independent Roles of Switching and Hypermutation in the Development and Persistence of B Lymphocyte Memory. *Immunity* **44**, 769-781 (2016).
70. J. M. Tas *et al.*, Visualizing antibody affinity maturation in germinal centers. *Science* **351**, 1048-1054 (2016).
71. J. Ersching *et al.*, Germinal Center Selection and Affinity Maturation Require Dynamic Regulation of mTORC1 Kinase. *Immunity* **46**, 1045-1058 e1046 (2017).
72. M. Pauthner *et al.*, Elicitation of Robust Tier 2 Neutralizing Antibody Responses in Nonhuman Primates by HIV Envelope Trimer Immunization Using Optimized Approaches. *Immunity* **46**, 1073-1088 e1076 (2017).
73. S. T. Jones *et al.*, Altered Immunoglobulin Repertoire and Decreased IgA Somatic Hypermutation in the Gut during Chronic HIV-1 Infection. *J Virol* **96**, e0097622 (2022).
74. B. F. Haynes *et al.*, Strategies for HIV-1 vaccines that induce broadly neutralizing antibodies. *Nat Rev Immunol* **23**, 142-158 (2023).
75. W. Xu *et al.*, HIV-1 evades virus-specific IgG2 and IgA responses by targeting systemic and intestinal B cells via long-range intercellular conduits. *Nat Immunol* **10**, 1008-1017 (2009).
76. S. Kaw *et al.*, HIV-1 infection of CD4 T cells impairs antigen-specific B cell function. *EMBO J* **39**, e105594 (2020).
77. P. Eldin *et al.*, Impact of HIV-1 Vpr manipulation of the DNA repair enzyme UNG2 on B lymphocyte class switch recombination. *J Transl Med* **18**, 310 (2020).
78. L. Manganaro *et al.*, HIV Vpu Interferes with NF-kappaB Activity but Not with Interferon Regulatory Factor 3. *J Virol* **89**, 9781-9790 (2015).
79. M. Lindqvist *et al.*, Expansion of HIV-specific T follicular helper cells in chronic HIV infection. *J Clin Invest* **122**, 3271-3280 (2012).
80. R. A. Cubas *et al.*, Inadequate T follicular cell help impairs B cell immunity during HIV infection. *Nat Med* **19**, 494-499 (2013).
81. M. C. Levesque *et al.*, Polyclonal B cell differentiation and loss of gastrointestinal tract germinal centers in the earliest stages of HIV-1 infection. *PLoS Med* **6**, e1000107 (2009).
82. J. D. Estes, Pathobiology of HIV/SIV-associated changes in secondary lymphoid tissues. *Immunol Rev* **254**, 65-77 (2013).
83. M. A. Piris *et al.*, Persistent and generalized lymphadenopathy: a lesion of follicular dendritic cells? An immunohistologic and ultrastructural study. *Am J Clin Pathol* **87**, 716-724 (1987).
84. K. Tenner-Racz, Human immunodeficiency virus associated changes in germinal centers of lymph nodes and relevance to impaired B-cell function. *Lymphology* **21**, 36-43 (1988).

85. J. Schmitz *et al.*, Follicular dendritic cells retain HIV-1 particles on their plasma membrane, but are not productively infected in asymptomatic patients with follicular hyperplasia. *J Immunol* **153**, 1352-1359 (1994).
86. O. Devergne *et al.*, Activation of cytotoxic cells in hyperplastic lymph nodes from HIV-infected patients. *AIDS* **5**, 1071-1079 (1991).
87. S. W. Craig, J. J. Cebra, Peyer's patches: an enriched source of precursors for IgA-producing immunocytes in the rabbit. *J Exp Med* **134**, 188-200 (1971).
88. P. Brandtzaeg, Presence of J chain in human immunocytes containing various immunoglobulin classes. *Nature* **252**, 418-420 (1974).
89. A. Cerutti, K. Chen, A. Chorny, Immunoglobulin responses at the mucosal interface. *Annu Rev Immunol* **29**, 273-293 (2011).
90. N. Y. Lycke, M. Bemark, The regulation of gut mucosal IgA B-cell responses: recent developments. *Mucosal Immunol* **10**, 1361-1374 (2017).
91. C. M. Snapper, W. E. Paul, Interferon-gamma and B cell stimulatory factor-1 reciprocally regulate Ig isotype production. *Science* **236**, 944-947 (1987).
92. K. Baklien, P. Brandtzaeg, Comparative mapping of the local distribution of immunoglobulin-containing cells in ulcerative colitis and Crohn's disease of the colon. *Clin Exp Immunol* **22**, 197-209 (1975).
93. T. Castro-Dopico *et al.*, Anti-commensal IgG Drives Intestinal Inflammation and Type 17 Immunity in Ulcerative Colitis. *Immunity* **50**, 1099-1114 e1010 (2019).
94. M. Uzzan *et al.*, Ulcerative colitis is characterized by a plasmablast-skewed humoral response associated with disease activity. *Nat Med* **28**, 766-779 (2022).
95. J. F. Scheid *et al.*, Remodeling of colon plasma cell repertoire within ulcerative colitis patients. *J Exp Med* **220**, (2023).
96. N. Kamada *et al.*, Humoral Immunity in the Gut Selectively Targets Phenotypically Virulent Attaching-and-Effacing Bacteria for Intraluminal Elimination. *Cell Host Microbe* **17**, 617-627 (2015).
97. J. Zaunders *et al.*, CD4(+) T Follicular Helper and IgA(+) B Cell Numbers in Gut Biopsies from HIV-Infected Subjects on Antiretroviral Therapy Are Similar to HIV-Uninfected Individuals. *Front Immunol* **7**, 438 (2016).
98. C. Planchais *et al.*, Early Antiretroviral Therapy Preserves Functional Follicular Helper T and HIV-Specific B Cells in the Gut Mucosa of HIV-1-Infected Individuals. *J Immunol* **200**, 3519-3529 (2018).
99. A. Kok *et al.*, Early initiation of combined antiretroviral therapy preserves immune function in the gut of HIV-infected patients. *Mucosal Immunol* **8**, 127-140 (2015).
100. S. Loisel-Meyer *et al.*, Glut1-mediated glucose transport regulates HIV infection. *Proc Natl Acad Sci U S A* **109**, 2549-2554 (2012).
101. C. S. Palmer *et al.*, Increased glucose metabolic activity is associated with CD4+ T-cell activation and depletion during chronic HIV infection. *AIDS* **28**, 297-309 (2014).
102. P. K. Datta *et al.*, Glutamate metabolism in HIV-1 infected macrophages: Role of HIV-1 Vpr. *Cell Cycle* **15**, 2288-2298 (2016).
103. I. Clerc *et al.*, Entry of glucose- and glutamine-derived carbons into the citric acid cycle supports early steps of HIV-1 infection in CD4 T cells. *Nat Metab* **1**, 717-730 (2019).
104. D. J. Morrison, T. Preston, Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut Microbes* **7**, 189-200 (2016).
105. V. Singh *et al.*, Butyrate producers, "The Sentinel of Gut": Their intestinal significance with and beyond butyrate, and prospective use as microbial therapeutics. *Front Microbiol* **13**, 1103836 (2022).
106. R. R. Fagundes *et al.*, Beyond butyrate: microbial fiber metabolism supporting colonic epithelial homeostasis. *Trends Microbiol*, (2023).
107. M. M. Seldin *et al.*, Trimethylamine N-Oxide Promotes Vascular Inflammation Through Signaling of Mitogen-Activated Protein Kinase and Nuclear Factor-kappaB. *J Am Heart Assoc* **5**, (2016).
108. B. C. Fu *et al.*, Associations of plasma trimethylamine N-oxide, choline, carnitine, and betaine with inflammatory and cardiometabolic risk biomarkers and the fecal microbiome in the Multiethnic Cohort Adiposity Phenotype Study. *Am J Clin Nutr* **111**, 1226-1234 (2020).
109. X. Liu *et al.*, Trimethylamine-N-oxide-stimulated hepatocyte-derived exosomes promote inflammation and endothelial dysfunction through nuclear factor-kappa B signaling. *Ann Transl Med* **9**, 1670 (2021).
110. S. F. Lima *et al.*, Transferable Immunoglobulin A-Coated Odoribacter splanchnicus in Responders to Fecal Microbiota Transplantation for Ulcerative Colitis Limits Colonic Inflammation. *Gastroenterology* **162**, 166-178 (2022).

111. M. Osawa *et al.*, Reduced abundance of butyric acid-producing bacteria in the ileal mucosa-associated microbiota of ulcerative colitis patients. *J Clin Biochem Nutr* **73**, 77-83 (2023).
112. T. Wang *et al.*, Autoimmune diseases exhibit shared alterations in the gut microbiota. *Rheumatology (Oxford)*, (2023).
113. F. Asnicar *et al.*, Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. *Nat Med* **27**, 321-332 (2021).
114. R. Gacesa *et al.*, Environmental factors shaping the gut microbiome in a Dutch population. *Nature* **604**, 732-739 (2022).
115. I. H. McHardy *et al.*, HIV Infection is associated with compositional and functional shifts in the rectal mucosal microbiota. *Microbiome* **1**, 26 (2013).
116. I. Vujkovic-Cvijin *et al.*, HIV-associated gut dysbiosis is independent of sexual practice and correlates with noncommunicable diseases. *Nat Commun* **11**, 2448 (2020).
117. I. Vujkovic-Cvijin *et al.*, The systemic anti-microbiota IgG repertoire can identify gut bacteria that translocate across gut barrier surfaces. *Sci Transl Med* **14**, eabl3927 (2022).
118. M. Uzzan *et al.*, Anti-α4β7 therapy targets lymphoid aggregates in the gastrointestinal tract of HIV-1-infected individuals. *Sci Transl Med* **10**, (2018).
119. M. J. McCarron, P. W. Park, D. R. Fooksman, CD138 mediates selection of mature plasma cells by regulating their survival. *Blood* **129**, 2749-2759 (2017).
120. C. W. Law, Y. Chen, W. Shi, G. K. Smyth, voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29 (2014).
121. C. W. Law *et al.*, RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Res* **5**, (2016).
122. M. E. Ritchie *et al.*, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
123. B. Phipson, S. Lee, I. J. Majewski, W. S. Alexander, G. K. Smyth, Robust Hyperparameter Estimation Protects against Hypervariable Genes and Improves Power to Detect Differential Expression. *Ann Appl Stat* **10**, 946-963 (2016).
124. G. K. Smyth, Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**, Article3 (2004).
125. 10xgenomics.
126. Y. Hao *et al.*, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529 (2021).
127. J. S. Weinstein *et al.*, Global transcriptome analysis and enhancer landscape of human primary T follicular helper and T effector lymphocytes. *Blood* **124**, 3719-3729 (2014).
128. P. A. Szabo *et al.*, Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat Commun* **10**, 4706 (2019).
129. A. B. Holmes *et al.*, Single-cell analysis of germinal-center B cells informs on lymphoma cell of origin and outcome. *J Exp Med* **217**, (2020).
130. S. W. Kazer *et al.*, Integrated single-cell analysis of multicellular immune dynamics during hyperacute HIV-1 infection. *Nat Med* **26**, 511-518 (2020).
131. H. W. King *et al.*, Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. *Sci Immunol* **6**, (2021).
132. A. Subramanian *et al.*, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).
133. D. Croft *et al.*, Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res* **39**, D691-697 (2011).
134. A. Reyes, M. Wu, N. P. McNulty, F. L. Rohwer, J. I. Gordon, Gnotobiotic mouse model of phage-bacterial host dynamics in the human gut. *Proc Natl Acad Sci U S A* **110**, 20236-20241 (2013).
135. D. T. Truong *et al.*, MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods* **12**, 902-903 (2015).
136. N. Segata *et al.*, Metagenomic biomarker discovery and explanation. *Genome Biol* **12**, R60 (2011).
137. P. J. McMurdie, S. Holmes, phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217 (2013).
138. E. J. Contijoch *et al.*, Gut microbiota density influences host physiology and is shaped by host and microbial factors. *Elife* **8**, (2019).
139. J. J. Faith *et al.*, The long-term stability of the human gut microbiota. *Science* **341**, 1237439 (2013).

Acknowledgements
This work was supported by R01 DK123749 (SM, JF and AC) and R01 DK112296 (SM). We also acknowledge the support of the Human Immune Monitoring Core, the Flow cytometry Core and the computational and data resources and staff expertise provided by Scientific Computing at the Icahn School of Medicine at Mount Sinai. We thank the Consortium for the Evaluation and Performance of HIV Incidence Assays (CEPHIA) and its funding sources, including the Bill and Melinda Gates Foundation (OPP1017716, OPP1062806, and OPP1115799); the NIH (P01 AI071713, R01 HD074511, P30 AI027763, R24 AI067039, U01 AI043638, P01 AI074621, and R24 AI106039); the HIV Prevention Trials Network (HPTN) sponsored by the NIAID, National Institutes of Child Health and Human Development (NICH/HD), National Institute on Drug Abuse, National Institute of Mental Health, and Office of AIDS Research of the NIH, DHHS (UM1 AI068613 and R01 AI095068); the California HIV-1 Research Program (RN07-SD-702); the Brazilian Program for STD and AIDS, Ministry of Health (914/BRA/3014-UNESCO); and the São Paulo City Health Department (2004–0.168.922–7).

Table 1: Participants demographics

	PWVH (n=18)	PWSH (n=49)	HIV NC (n=87)	p value
Male sex	16 (89%)	33 (67%)	44(51%)	0.004
Median (IQR) Age	44 (35 – 54)	56 (52 – 61)	51 (35 - 62)	0.003
Ethnicity				
Caucasian	0	10 (20%)	36 (41%)	
African-American	12 (67%)	12 (24%)	13 (15%)	
Hispanic/Latino	6 (28%)	25 (51%)	21 (24%)	<.0001
Other	0	2 (4%)	15 (17%)	
Risk factor for HIV acquisition				
MSM	9 (50%)	18 (37%)		
Heterosexual contact	6 (33%	16 (33%)	-	0.662
IVDU	2 (11%)	5 (10%)		
Other	0	1 (2%)		
NA	1 (6%)	9 (18%)		
Years since HIV diagnosis	5.2 (0.04 – 25)	19.7 (11.6 – 24.5)	-	0.021
CD4 count at study baseline	307 (114 – 547)	617 (457 – 841)	-	<.0001
Nadir CD4 count	298 (55 – 379)	193 (94 -229)	-	0.191
HIV RNA (copies/ml)	4,705 (33,950 – 206,750)	-	-	-
Years of virological suppression	-	5.0 (3.8 – 10.1)	-	-
Type of ARV				
INsTI		30 (61%)		
PI	-	6 (12%)	-	-
NNRTI		6 (12%)		
Other/NA		7 (14%)		

Values are expressed as frequency or median (interquartile range) as appropriate.

PWVH: People with viremic HIV-1 infection. PWSH: People with suppressed HIV-1 infection. HIV NC: HIV negative control. MSM:Men who have Sex with Men; IVDU: Intravenous Drug Use; N/A: not available; ARV: antiretroviral therapy; INsTI: INtegrase Strand Inhibitors; PI: Protease Inhibitor; NNRTI: Non-Nucleoside Reverse Transcriptase Inhibitor. Comparisons. Comparisons between groups performed with One-way ANOVA with Dunn's correction for multiple comparisons.

Figure 1. scRNA sequencing-based Identification of major immune cell types in the intestinal lamina propria during HIV infection. (A) Uniform Manifold Approximation and Projection (UMAP) of single cell RNA sequencing (scRNA-seq) data from colonic lamina propria with representation of sample distribution across the groups (HIV negative controls (NC), n=9, grey; People with suppressed HIV infection (PWSH), n=5, light red; People with viremic HIV infection (PWVH), n=4, deep red). **(B)** UMAP representing cell-type annotated clusters derived from the colonic lamina propria. **(C)** UMAP of ileal lamina propria-derived samples with representation of sample distribution across the groups (NC n=5, grey; PWSH, n=5, light red; PWVH, n=4, deep red). **(D)** UMAP representing cell-type annotated clusters derived from the ileal lamina propria. **(E, F)** Heatmap showing the average z-score-normalized log expression of canonical cell type markers across different colonic **(E)** and ileum-derived **(F)** cell clusters. **(G, H)** Bar plots showing cell-type distribution within each individual for colonic-derived **(G)** and ileum-derived **(H)** samples. **(I)** UMAPs showing re-clustered, colon-derived B cell and PC clusters using scRNA-seq. **(J)** Heatmap showing the average normalized log expression z-score of canonical B cell markers across different colonic-derived B cell sub-clusters. **(K)** UMAPs showing re-clustered, ileum-derived B cell and PC clusters using scRNA-seq. **(L)** Heatmap showing the average z-score normalized log expression of canonical B cell markers across different ileum-derived B cell sub-clusters. **(M, N)** scRNA-seq data demonstrating the frequency of B cell sub-clusters derived from the colon **(M)** and ileum **(N)** across the three groups (HIV- controls (NC), n=9 for colon and n=5 for ileum; people with suppressed HIV (PWSH), n=5 for colon and ileum, both; people with viremic HIV (PWVH) n=4 for colon and ileum, both). Bar represents median values. P-values from Wilcoxon signed-rank sum test are reported.

Figure 2. Depletion of intestinal Germinal Center B cells during viremic HIV-1 infection.
(A, B) Flow cytometric data quantifying colonic **(A)** and ileal **(B)** total B cells and B cell subsets. Comparisons between groups performed with one-way ANOVA with Dunn's correction for multiple comparisons, p-values as indicated. **(C)** Representative Immunohistochemistry (IHC) images of intestinal lymphoid follicles for each of the three groups showing staining for CD4 (top panels), and BCL6 (middle panels) and CD23 (bottom panels). Scale bar represents 100μm. **(D)** Summary data for IHC (NC n=6; PWSH, n=5; PWVH, n=4), quantifying lymphoid follicle size, GC size, the ratio between GC and lymphoid follicle size, the number of $BCL6^+$ cells in each GC and the % area positive for CD23 staining within each GC. **(E)** Correlation analysis between the area with positive CD23 staining and the number of $BCL6^+$ cells within each GC. Comparisons between groups performed one-way ANOVA for non-parametric variables (Kruskal-Wallis test) with Dunn's correction for multiple comparisons. Median and interquartile range was used to plot summary data, p values <0.1 are shown as indicated (p-values > 0.1 are not shown). Correlation was obtained using Spearman's r correlation coefficient, p values as indicated.

Figure 3. Colonic Plasma cells are decreased during untreated HIV infection and show switch in isotype expression. (A) Representative flow cytometry plots of Plasma Cells (PCs) (live $CD45^{+}CD19^{+/-}CD27^{+}CD38^{high}$cells) derived from NC, n=51; PWSH, n=35; PWVH, n=16. **(B)** Summary data quantifying colonic PCs for each of the three groups. **(C)** Representative flow cytometry plots demonstrating PC isotype expression. **(D)** Summary data quantifying colonic IgA^{+} IgG^{+} and IgM^{+} PCs. **(E)** Representative Immunofluorescence (IF) images showing the expression of CD138 (red), IgA (green) and DAPI (blue). Top panel (10X magnification) with red insert shown in the bottom panels (20X magnification). Red cells indicate IgA^{-} PCs, green cells indicate IgA^{+} cells and yellow cells (merge) indicate IgA^{+} PCs. **(F)** Summary IF data showing total IgA^{-} PC per square mm area (left panel) and ratio of IgA^{-} : IgA^{+} PCs per square mm area (right panel). **(G)** Representative flow cytometry plots of peripheral blood (PB)-derived PCs (live $CD45^{+}CD19^{+/-}CD27^{+}IgD^{-}CD38^{high}$cells) for each of the three groups. **(H)** Summary data quantifying PB PCs. **(I)** Summary data quantifying PB-derived IgA^{+}, IgG^{+} and IgM^{+} PCs. Comparisons between groups performed one-way ANOVA for non-parametric variables (Kruskal-Wallis test) with Dunn's correction for multiple comparisons. Median and interquartile range was used to plot summary data, p values <0.1 are shown as indicated (p values > 0.1 are not shown).

Figure 4. Plasma Cell transcriptional profile is altered during HIV-1 infection. (A) Heatmaps showing the average normalized log expression z-score of Immunoglobulin genes across different colon (top panel) and ileum (bottom panel) PCs sub-clusters. **(B)** Heatmaps showing the mean pathway z-score for each colon (left panel)- and ileum (right panel)- derived PC isotype, across the three groups ((NC), n=9 for colon and n=5 for ileum; PWSH, n=5 for colon and ileum, both; PWVH n=4 for colon and ileum, both). **(C)** Bubble plot showing the Normalized Enrichment Score (NES) of selected pathways of interest derived from the top five pathways identified with Sumer analysis (methods) in flow cytometrically-sorted colonic PCs of each of the three groups (NC n=4; PWSH n=5; PWVH n=3)._The bubble size represents the NES, and the color represents the direction of the change (purple = increased, yellow = decreased). Color intensity indicates statistical significance levels, with lighter colors indicating $p < 0.05$, darker colors for $p < 0.01$, and the darkest colors for $p < 0.001$. Pathway names ending with "K" are from the KEGG database, "B" represents Bioplanet, and "R" stands for the Reactome database. **(D)** Heatmap showing the expression of selected Leading-Edge Genes (LEGs) from pathways in panel (E) for each participant across the three groups.

Figure 5. Viremic HIV-1 infection is associated with perturbations in T cell subsets including TFH-like cells. (A, B) Uniform Manifold Approximation and Projections (UMAPs) showing re-clustered, colon **(A)**- and ileum **(B)**- derived T cell clusters using scRNA-seq. **(C, D)** Heatmaps showing the average z-score normalized log expression of T cell related genes across different

colonic **(C)** and ileum-derived **(D)** T cell subclusters. **(E, F)** Boxplots showing the scRNA-seq derived frequencies of colon-derived **(E)** and ileum-derived **(F)** $CD4^+$ and $CD8^+$ T cells as a fraction of all T cells (top panels), across the three groups (HIV- controls (HIV NC), n=9 for colon and n=5 for ileum; people with suppressed HIV (PWSH), n=5 for colon and ileum, both; people with viremic HIV (PWVH) n=4 for colon and ileum, both). In panels E and F, bottom boxplots show the scRNA-seq derived frequencies of colon-derived **(E)** and ileum-derived **(F)** subtypes of CD4+ T cells, across the three groups. In panels E and F, comparisons were performed with Wilcoxon test, p values as indicated **(G)** Correlation matrix showing relationships between frequencies of colonic PCs and PC isotypes and frequency of colonic T cell subsets.. **(H)** Correlation matrix showing relationships between frequencies of ileum-derived B cells and frequency of ileal T cell subsets. Panels G and H represent derived from the flow cytometric dataset (NC, n=51; PWSH, n=35; PWVH, n=16)). **(I)** Heatmap representing the mean z-score enrichment for selected pathways within the ileum-derived TFH-like T cell subcluster.

Figure 6: Altered frequency and Isotype expression of colonic PCs contribute to persistent systemic inflammation and immune activation during HIV infection. **(A, B)** Frequency of peripheral blood-derived activated ($CD38^+$) $CD4^+$T cells **(A)** and $CD8^+$ T cells **(B)** obtained via flow—cytometry analysis, across the three groups of study participants (NC, n=33; PWSH, n=32 ; and PWVH, n=12) . **(C,D)** Frequency of peripheral blood-derived cycling ($Ki67^+$) $CD4^+$ T cells **(C)** and $CD8^+$ T cells **(D)** across the three groups of study participants. For panels A-D, comparisons between groups performed one-way ANOVA for non-parametric variables (Kruskal-Wallis test) with Dunn's correction for multiple comparisons. Median and interquartile range was used to plot summary data, p values <0.1 are shown as indicated (p values > 0.1 are not shown).**(E)** Correlation matrix showing the correlation between colonic PCs as well as PCs isotypes and peripheral blood-derived activated and cycling T cells subsets using flow cytometry data, correlation was obtained using Spearman's r correlation coefficient, p values as indicated **(F)** Bar plot showing systemic inflammatory markers in PWVH compared to NC (dark red) and in PWSH vs NC (light red). Dark gray bars represent systemic inflammatory markers found to be significantly higher in NC compared to PWVH and light gray bars indicate systemic inflammatory markers found to be significantly higher in NC to PWSH. Comparisons were performed by Student's t-test with Benjamani-Hochberg correction for multiple comparisons, p values as indicated. **(G)** Correlation matrix showing the relationship between flow-cytometry derived colonic PCs as well as PCs isotypes and systemic inflammatory markers noted to be elevated in People with HIV (PWH) compared to NC. Correlation was obtained using Spearman's r correlation coefficient, p values as indicated.

Figure 7: Distinct intestinal microbiota changes and targeting by secretory IgAs during HIV infection relate to altered colonic PCs Isotype expression. **(A)** Representative flow plots showing IgA-bound and unbound stool bacteria across the three groups (NC, PWSH and PWVH).

(B) Summary data representing the frequency of IgA-bound bacteria across the three groups where stool samples were available for stool bacteria sorting in the primary cohort (NC, n=16; PWSH, n=18; PWVH, n=7. Bar represents median and IQR values. Statistical comparisons performed with Mann-Whitney test, p values as indicated. **(C)** Correlation plots between Log Palm Score and Log Cell abundance of IgA^+ or IgG^+ colonic PCs for bacteria where a significant correlation was found. **(D)** Principal component analysis (PCoA) plot showing Beta diversity (using Bray-Curtis distances) between People with HIV (PWH) and NC. **(E)** PCoA plot showing Beta diversity (using Bray-Curtis distances) across PWH who are men who have sex with men (MSM) or not as well as NC who are MSM or not. **(F)** Box plots showing the relative abundance of Prevotella (top plot) and Bacteriodes (bottom plot) across PWH and NC who are MSM or not. **(G)** Barplot showing the distribution of bacterial phyla across the three groups. **(H)** Linear discriminant analysis (LDA) effect size (LEfSe) performed on all Operational Taxonomic Units (OTUs) comparing PWVH and NC. **(I)** Log IgA score for *Roseburia* in the validation cohort. Bar represents median and IQR values. Statistical comparisons performed with Mann-Whitney test, p values as indicated. **(J)** Correlation Plots between the relative abundance of selected bacterial genera and the frequency of colonic PCs or IgA PCs.

Figure S1: scRNA-seq analyses of colonic and ileal tissue reveals different distribution of cell types across groups. (A) Histogram showing cell count per colonic cell cluster. **(B)** Histogram showing cell count per ileal cell cluster. **(C)** scRNA-seq-derived data showing cell type distribution across the three groups of study participants in colonic samples (HIV- controls (NC) n=9; people with suppressed HIV infection (PWSH) n=5; people with viremic HIV infection (PWVH) n=4). **(D)** scRNA-seq-derived data showing cell type distribution across the three groups of study participants in ileum-derived samples (NC, n=5; PWSH, n=5; PWVH, n=4). In panels E and F comparisons were performed with Wilcoxon test, p-values as indicated.

Figure S2: Gating Strategy for flow cytometry data. (A) Gating strategy used for lamina propria B cells and PCs showing PCs as live $CD45^{+}CD19^{+/-}CD27^{+}CD38^{high}$cells; Short-lived (SL) PCs as live $CD45^{+}CD27^{+}CD38^{high}CD19^{+}$cells and Long-lived (LL) PCs as live $CD45^{+}CD27^{+}CD38^{high}CD19^{-}$cells. Total B cells are defined as live $CD45^{+}CD19^{+}CD10^{-}CD38^{-}$ and Germinal Center (GC) B cells are defined as $CD45^{+}CD3^{-}CD19^{+}CD38^{int}IgD^{-}$ cells. Naïve B cells are defined as live $CD45^{+}CD19^{+}CD10^{-}CD38^{-}IgM^{+}IgD^{+}$cells. Memory IgM expressing B Cells are defined as live $CD45^{+}CD19^{+}CD10^{-}CD38^{-}IgM^{+}$cells, Switched Memory (SM) B cells are defined as live $CD45^{+}CD19^{+}CD10^{-}CD38^{-}IgM^{-}IgD^{-}$cells. SM B cells are further classified into IgA^{+}(live $CD45^{+}CD19^{+}CD10^{-}CD38^{-}IgM^{-}IgD^{-}IgA^{+}$) and IgA^{-} (live $CD45^{+}CD19^{+}CD10^{-}CD38^{-}IgM^{-}IgD^{-}$ IgA) **(B)** Gating strategy used for peripheral blood (PB)-derived B cells and PCs showing PCs as live $CD45^{+}CD19^{+/-}CD27^{+}CD38^{high}IgD^{-}D27^{+}$ cells and total B cells as live $CD45^{+}CD19^{+}CD10^{-}$ $CD38^{-}$cells. Naïve B cells are defined as live $CD45^{+}CD19^{+}CD10^{-}CD38^{-}IgM^{+}IgD^{+}$ cells. Memory IgM expressing B cells are defined as live $CD45^{+}CD19^{+}CD10^{-}CD38^{-}IgM^{+}$cells, SM B Cells are defined as live $CD45^{+}CD19^{+}CD10^{-}CD38^{-}IgM^{-}IgD^{-}$cells. **(C)** Gating strategy used for lamina propria and PB T cells. $CD4^{+}$ T cells are defined as live $CD45^{+}CD3^{+}CD4^{+}$ cells. Similarly, $CD8^{+}$ T cells are defined as liveCD45^{+}CD3^{+}CD8+cells. Activated $CD4^{+}$ or $CD8^{+}$ T cells are defined as $CD38^{+}$ only, $HLADR^{+}$ only or $CD38^{+}HLADR^{+}$ cells. Cycling T cells are defined as $Ki67^{+}$ $CD4^{+}$ or $CD8^{+}$ T cells.

Figure S3: Expression of CD23 in the putative light zones of the GCs. Representative Immunohistochemistry (IHC) images of intestinal lymphoid aggregates from a NC showing the expression of BCL6 (top panel) and CD23 (bottom panel). LZ= light zone, DZ= dark zone, T= T cell zone. Scale bar represents 100μm.

Figure S4: Expression of BCL6 correlates with CD23 in PWVH. Representative IHC images from a PWVH (HIV066) showing 4 different lymphoid aggregates, comparing the expression of BCL6 and CD23. Yellow arrows indicate enhanced expression of BCL6 in a lymphoid follicle that demonstrated CD23 staining.

Figure S5: Flow cytometric analyses demonstrate reduced PCs and isotype changes across different subtypes of PCs. **(A)** Summary data quantifying ileal PCs and IgA^+, IgG^+ and IgM^+ PCs for each of the three groups of study participants (PWVH, PWSH and NC). **(B, C)** Summary data quantifying colonic **(B)** and ileal **(C)** Short-Lived (SL) and Long-Lived (LL) PCs for each of the three groups. **(D, E)** Summary data quantifying colonic **(D)** and ileal **(E)** IgA^+, IgG^+ and IgM^+ SL PCs for each of the three groups of study participants. **(F, G)** Summary data quantifying colonic **(F)** and ileal **(G)** IgA+, IgG^+ and IgM^+ LL PCs for each of the three groups of subjects. Comparisons between groups performed one-way ANOVA for non-parametric variables (Kruskal-Wallis test) with Dunn's correction for multiple comparisons. Median and interquartile range was used to plot summary data, p values <0.1 are shown as indicated (p values > 0.1 are not shown).

Figure S7: Alterations in frequency and transcriptional profile of intestinal T cell subsets in PWH. **(A, B)** Barplots comparing the frequency of colon- **(A)** and ileum-derived **(B)** $CD4^+$ and $CD8^+$ T cells across the three groups (PWVH, PWSH and NC) **(C, D)** Barplots comparing the frequency of colon- **(C)** and ileum-derived **(D)** activated ($CD38^+$) $CD4^+$ and $CD8^+$ T cells across the three groups of study participants. **(E, F)** Barplots comparing the frequency of colon- **(E)** and ileum-derived **(F)** cycling ($Ki67^+$) $CD4^+$ and $CD8^+$ T cells across the three groups. For panels A - F data is derived from the flow-cytometry dataset: (NC, n=51; PWSH, n=35; PWVH, n=16). Comparisons between groups performed one-way ANOVA for non-parametric variables (Kruskal-Wallis test) with Dunn's correction for multiple comparisons. Median and interquartile range was used to plot summary data, p-values are as indicated. **(G)** Boxplots comparing the frequency of ileum-associated $CD8^+$ T cell subsets. Data is derived from scRNA-seq analyses. Comparisons performed with Wilcoxon test, p-values as indicated **(H)** Heatmap representing the mean z-score enrichment for selected pathways within the colonic TFH-like T cell subcluster.

Figure S8: Intestinal microbiota changes and targeting by secretory IgAs during HIV infection relate to altered colonic PCs Isotype expression. **(A)** Summary data representing the frequency of IgA-bound bacteria across the three groups where stool samples were available for stool bacteria sorting in the validation cohort (Consortium for the Evaluation and Performance of HIV Incidence Assays (CEPHIA) of the University of California, San Francisco; 5 PWVH, 13 PWSH and 12 NC) **(B)** Barplot showing the distribution of bacteria genera across the three groups of study participants (NC, n=16; PWSH, n=18 and PWVH, n=7). **(C)** Linear discriminant analysis (LDA) effect size (LEfSe) performed on all Operational Taxonomic Units (OTUs) comparing PWSH and NC (top panel) and comparing PWSH and PWVH (bottom panel). **(D)** Correlation Plots between the OTU-based relative abundance of bacterial genera and the frequency of colonic PCs and IgA^+ PCs where a significant correlation was found using Spearman correlation. Unadjusted p values as indicated.

Figure 1

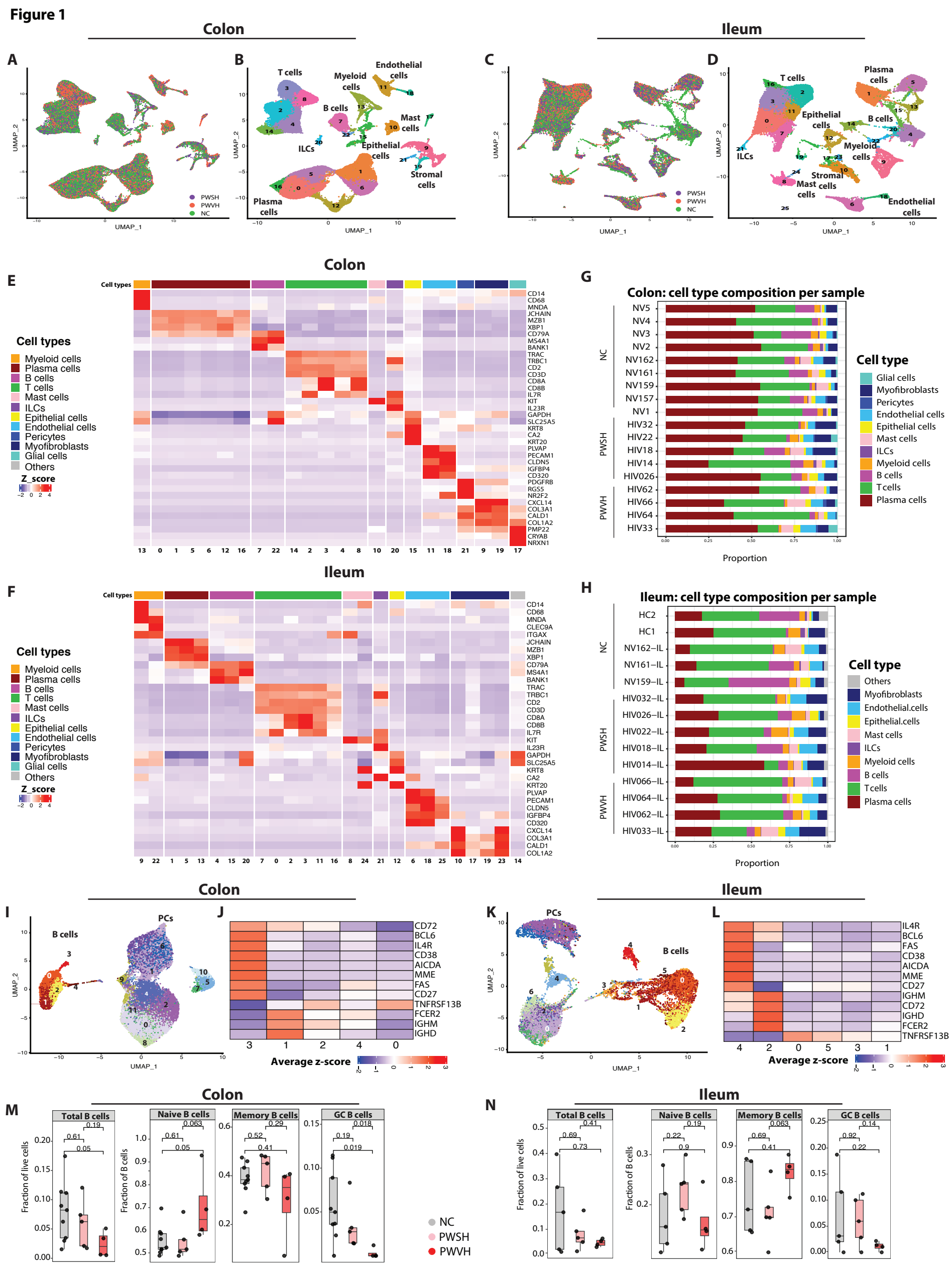

Colon
Ileum
A
B
C
D
T cells
Myeloid cells
Endothelial cells
B cells
Mast cells
ILCs
Epithelial cells
Stromal cells
Plasma cells
PWSH
PWVH
NC
UMAP_1
UMAP_2
E
F
Cell types
Myeloid cells
Plasma cells
B cells
T cells
Mast cells
ILCs
Epithelial cells
Endothelial cells
Pericytes
Myofibroblasts
Glial cells
Others
Z_score
G
H
Colon: cell type composition per sample
Ileum: cell type composition per sample
Cell type
Glial cells
Myofibroblasts
Pericytes
Endothelial cells
Epithelial cells
Mast cells
ILCs
Myeloid cells
B cells
T cells
Plasma cells
Proportion
I
J
K
L
PCs
B cells
Average z-score
M
N
Total B cells
Naive B cells
Memory B cells
GC B cells
Fraction of live cells
Fraction of B cells
NC
PWSH
PWVH

Figure 2

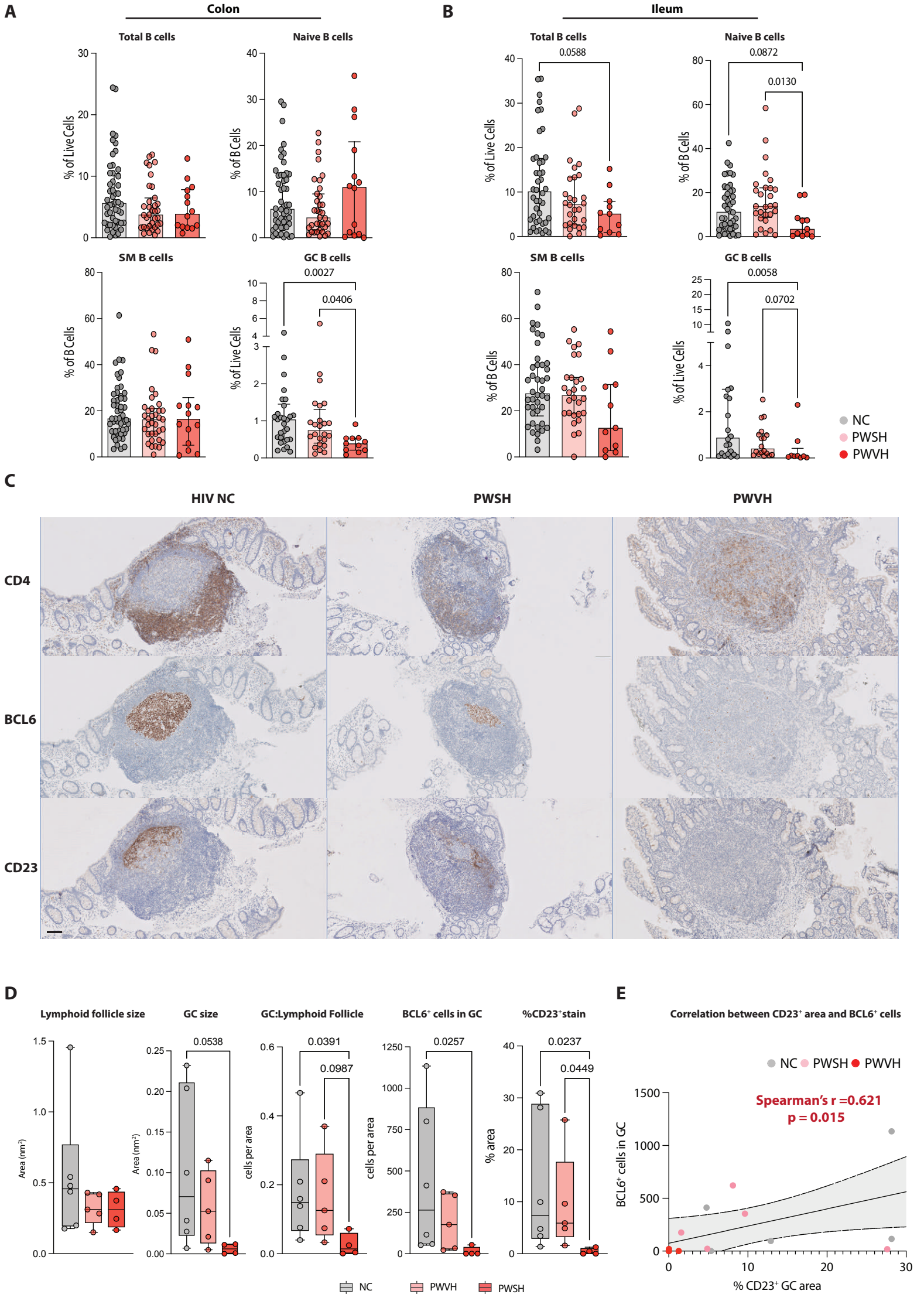

A
Colon
Total B cells
Naive B cells
SM B cells
GC B cells
% of Live Cells
% of B Cells
0.0027
0.0406
B
Ileum
0.0588
0.0872
0.0130
0.0058
0.0702
NC
PWSH
PWVH
C
HIV NC
PWSH
PWVH
CD4
BCL6
CD23
D
Lymphoid follicle size
GC size
GC:Lymphoid Follicle
BCL6⁺ cells in GC
%CD23⁺stain
Area (nm²)
cells per area
% area
0.0538
0.0391
0.0987
0.0257
0.0237
0.0449
NC
PWVH
PWSH
E
Correlation between CD23⁺ area and BCL6⁺ cells
NC PWSH PWVH
Spearman's r =0.621
p = 0.015
BCL6⁺ cells in GC
% CD23⁺ GC area

Figure 3

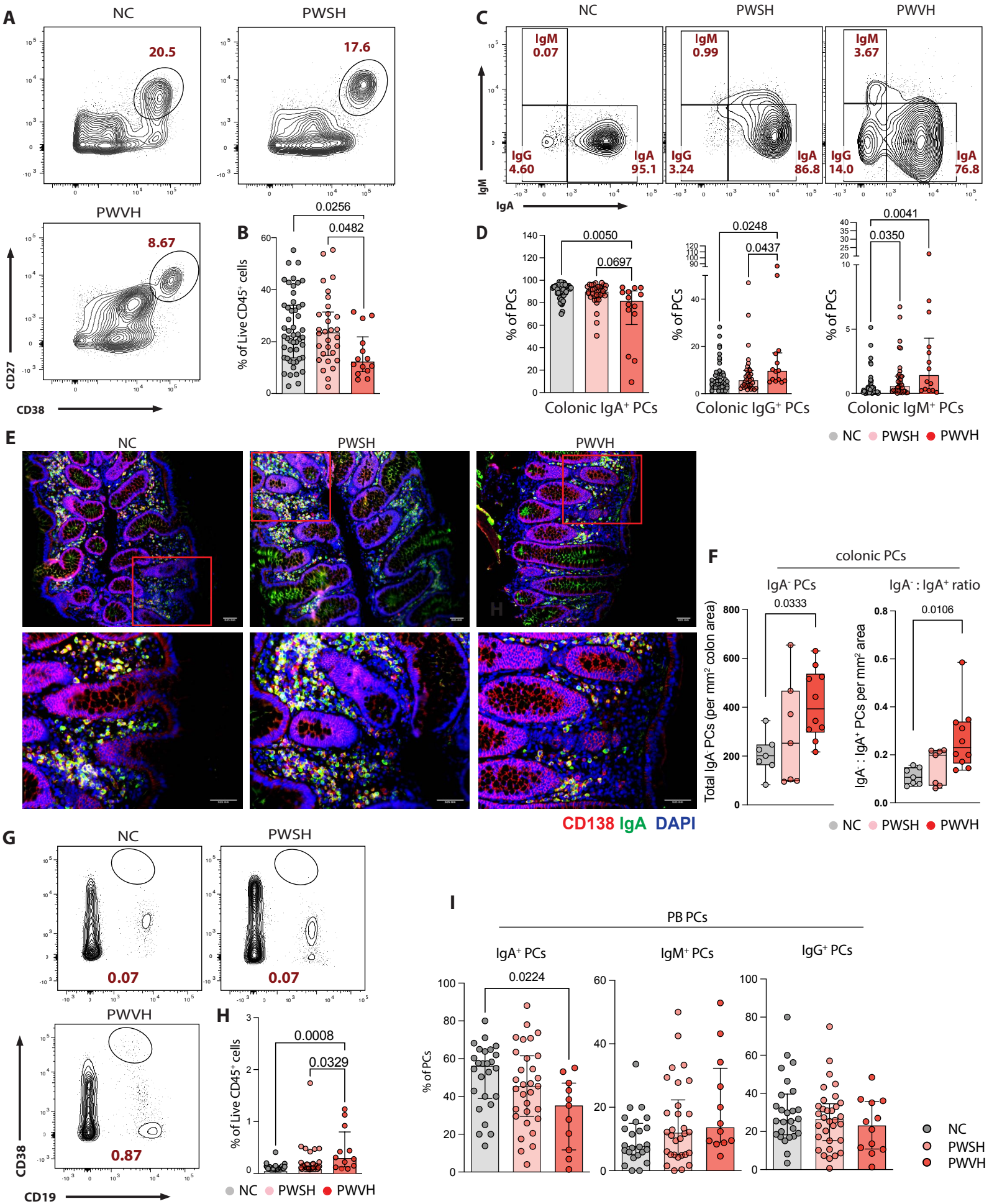

A
NC
PWSH
PWVH
20.5
17.6
8.67
CD27
CD38
B
% of Live CD45⁺ cells
0.0256
0.0482
C
NC
PWSH
PWVH
IgM
IgA
IgM 0.07
IgG 4.60
IgA 95.1
IgM 0.99
IgG 3.24
IgA 86.8
IgM 3.67
IgG 14.0
IgA 76.8
D
% of PCs
0.0050
0.0697
0.0248
0.0437
0.0041
0.0350
Colonic IgA⁺ PCs
Colonic IgG⁺ PCs
Colonic IgM⁺ PCs
NC PWSH PWVH
E
NC
PWSH
PWVH
CD138 IgA DAPI
F
colonic PCs
IgA⁻ PCs
IgA⁻ : IgA⁺ ratio
Total IgA⁻ PCs (per mm² colon area)
IgA⁻ : IgA⁺ PCs per mm² area
0.0333
0.0106
NC PWSH PWVH
G
NC
PWSH
PWVH
0.07
0.07
0.87
CD38
CD19
H
% of Live CD45⁺ cells
0.0008
0.0329
NC PWSH PWVH
I
PB PCs
IgA⁺ PCs
IgM⁺ PCs
IgG⁺ PCs
% of PCs
0.0224
NC
PWSH
PWVH

Figure 4

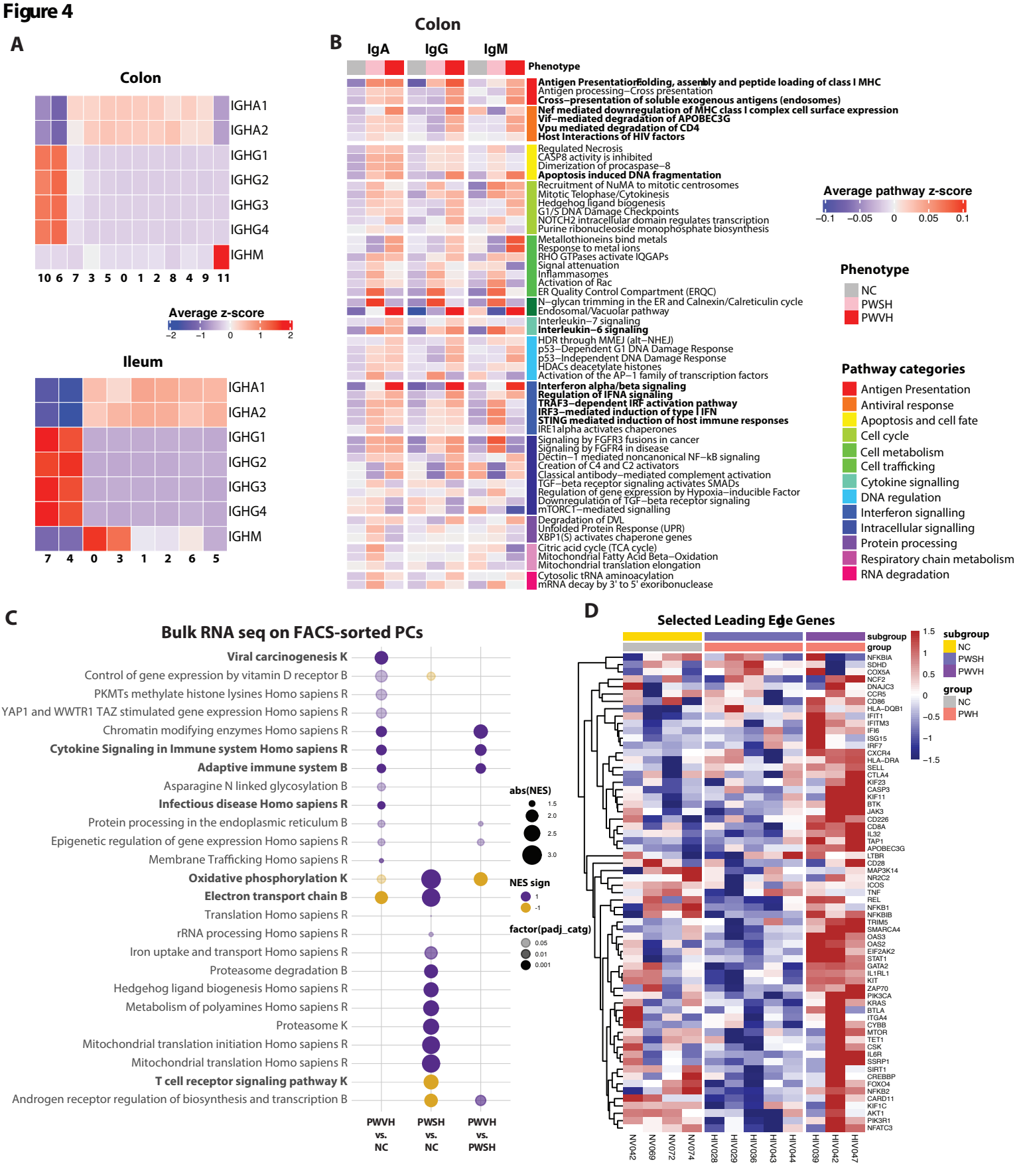

A
Colon
IGHA1
IGHA2
IGHG1
IGHG2
IGHG3
IGHG4
IGHM
10 6 7 3 5 0 1 2 8 4 9 11
Average z-score
−2 −1 0 1 2
Ileum
IGHA1
IGHA2
IGHG1
IGHG2
IGHG3
IGHG4
IGHM
7 4 0 3 1 2 6 5
B
Colon
IgA IgG IgM
Phenotype
Antigen Presentation: Folding, assembly and peptide loading of class I MHC
Antigen processing−Cross presentation
Cross−presentation of soluble exogenous antigens (endosomes)
Nef mediated downregulation of MHC class I complex cell surface expression
Vif−mediated degradation of APOBEC3G
Vpu mediated degradation of CD4
Host Interactions of HIV factors
Regulated Necrosis
CASP8 activity is inhibited
Dimerization of procaspase−8
Apoptosis induced DNA fragmentation
Recruitment of NuMA to mitotic centrosomes
Mitotic Telophase/Cytokinesis
Hedgehog ligand biogenesis
G1/S DNA Damage Checkpoints
NOTCH2 intracellular domain regulates transcription
Purine ribonucleoside monophosphate biosynthesis
Metallothioneins bind metals
Response to metal ions
RHO GTPases activate IQGAPs
Signal attenuation
Inflammasomes
Activation of Rac
ER Quality Control Compartment (ERQC)
N−glycan trimming in the ER and Calnexin/Calreticulin cycle
Endosomal/Vacuolar pathway
Interleukin−7 signaling
Interleukin−6 signaling
HDR through MMEJ (alt−NHEJ)
p53−Dependent G1 DNA Damage Response
p53−Independent DNA Damage Response
HDACs deacetylate histones
Activation of the AP−1 family of transcription factors
Interferon alpha/beta signaling
Regulation of IFNA signaling
TRAF3−dependent IRF activation pathway
IRF3−mediated induction of type I IFN
STING mediated induction of host immune responses
IRE1alpha activates chaperones
Signaling by FGFR3 fusions in cancer
Signaling by FGFR4 in disease
Dectin−1 mediated noncanonical NF−kB signaling
Creation of C4 and C2 activators
Classical antibody−mediated complement activation
TGF−beta receptor signaling activates SMADs
Regulation of gene expression by Hypoxia−inducible Factor
Downregulation of TGF−beta receptor signaling
mTORC1−mediated signalling
Degradation of DVL
Unfolded Protein Response (UPR)
XBP1(S) activates chaperone genes
Citric acid cycle (TCA cycle)
Mitochondrial Fatty Acid Beta−Oxidation
Mitochondrial translation elongation
Cytosolic tRNA aminoacylation
mRNA decay by 3' to 5' exoribonuclease
Average pathway z-score
−0.1 −0.05 0 0.05 0.1
Phenotype
NC
PWSH
PWVH
Pathway categories
Antigen Presentation
Antiviral response
Apoptosis and cell fate
Cell cycle
Cell metabolism
Cell trafficking
Cytokine signalling
DNA regulation
Interferon signalling
Intracellular signalling
Protein processing
Respiratory chain metabolism
RNA degradation
C
Bulk RNA seq on FACS-sorted PCs
Viral carcinogenesis K
Control of gene expression by vitamin D receptor B
PKMTs methylate histone lysines Homo sapiens R
YAP1 and WWTR1 TAZ stimulated gene expression Homo sapiens R
Chromatin modifying enzymes Homo sapiens R
Cytokine Signaling in Immune system Homo sapiens R
Adaptive immune system B
Asparagine N linked glycosylation B
Infectious disease Homo sapiens R
Protein processing in the endoplasmic reticulum B
Epigenetic regulation of gene expression Homo sapiens R
Membrane Trafficking Homo sapiens R
Oxidative phosphorylation K
Electron transport chain B
Translation Homo sapiens R
rRNA processing Homo sapiens R
Iron uptake and transport Homo sapiens R
Proteasome degradation B
Hedgehog ligand biogenesis Homo sapiens R
Metabolism of polyamines Homo sapiens R
Proteasome K
Mitochondrial translation initiation Homo sapiens R
Mitochondrial translation Homo sapiens R
T cell receptor signaling pathway K
Androgen receptor regulation of biosynthesis and transcription B
PWVH vs. NC
PWSH vs. NC
PWVH vs. PWSH
abs(NES)
1.5 2.0 2.5 3.0
NES sign
1 -1
factor(padj_catg)
0.05 0.01 0.001
D
Selected Leading Edge Genes
subgroup
group
NFKBIA
SDHD
COX5A
NCF2
DNAJC3
CCR5
CD86
HLA−DQB1
IFIT1
IFITM3
IFI6
ISG15
IRF7
CXCR4
HLA−DRA
SELL
CTLA4
KIF23
CASP3
KIF11
BTK
JAK3
CD226
CD8A
IL32
TAP1
APOBEC3G
LTBR
CD28
MAP3K14
NR2C2
ICOS
TNF
REL
NFKB1
NFKBIB
TRIM5
SMARCA4
OAS3
OAS2
EIF2AK2
STAT1
GATA2
IL1RL1
KIT
ZAP70
PIK3CA
KRAS
BTLA
ITGA4
CYBB
MTOR
TET1
CSK
IL6R
SSRP1
SIRT1
CREBBP
FOXO4
NFKB2
CARD11
KIF1C
AKT1
PIK3R1
NFATC3
NV042 NV069 NV072 NV074 HIV028 HIV029 HIV036 HIV043 HIV044 HIV039 HIV042 HIV047
1.5 1 0.5 0 −0.5 −1 −1.5
subgroup
NC
PWSH
PWVH
group
NC
PWH

Figure 5

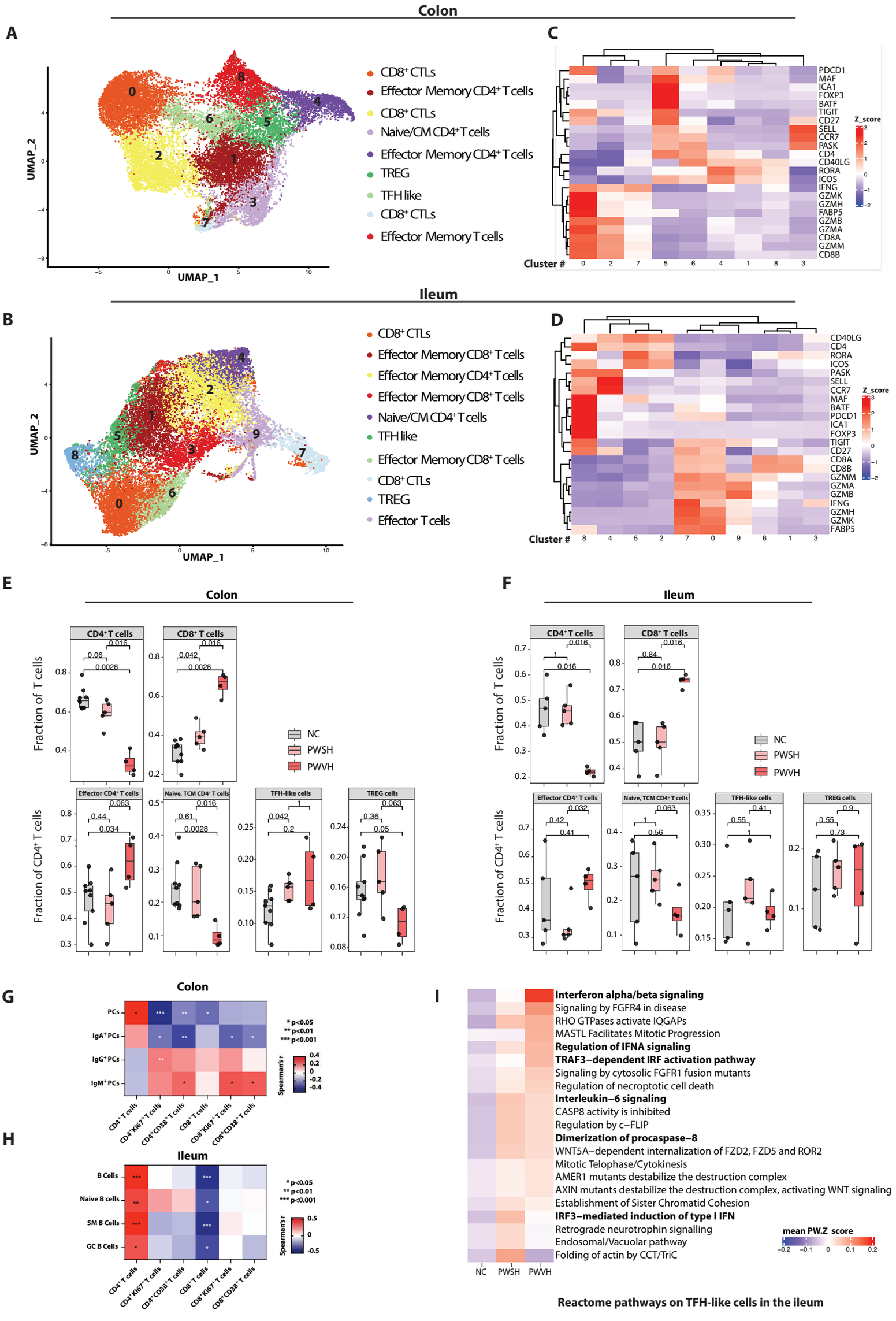

Colon
A
UMAP_2
UMAP_1
CD8+ CTLs
Effector Memory CD4+ T cells
CD8+ CTLs
Naive/CM CD4+ T cells
Effector Memory CD4+ T cells
TREG
TFH like
CD8+ CTLs
Effector Memory T cells
C
PDCD1
MAF
ICA1
FOXP3
BATF
TIGIT
CD27
SELL
CCR7
PASK
CD4
CD40LG
RORA
ICOS
IFNG
GZMK
GZMH
FABP5
GZMB
GZMA
CD8A
GZMM
CD8B
Z_score
Cluster #
Ileum
B
CD8+ CTLs
Effector Memory CD8+ T cells
Effector Memory CD4+ T cells
Effector Memory CD8+ T cells
Naive/CM CD4+ T cells
TFH like
Effector Memory CD8+ T cells
CD8+ CTLs
TREG
Effector T cells
D
CD40LG
CD4
RORA
ICOS
PASK
SELL
CCR7
MAF
BATF
PDCD1
ICA1
FOXP3
TIGIT
CD27
CD8A
CD8B
GZMM
GZMA
GZMB
IFNG
GZMH
GZMK
FABP5
Z_score
Cluster #
E
Colon
CD4+ T cells
CD8+ T cells
Fraction of T cells
NC
PWSH
PWVH
Effector CD4+ T cells
Naive, TCM CD4+ T cells
TFH-like cells
TREG cells
Fraction of CD4+ T cells
F
Ileum
CD4+ T cells
CD8+ T cells
Fraction of T cells
NC
PWSH
PWVH
Effector CD4+ T cells
Naive, TCM CD4+ T cells
TFH-like cells
TREG cells
Fraction of CD4+ T cells
G
Colon
PCs
IgA+ PCs
IgG+ PCs
IgM+ PCs
CD4+ T cells
CD4+Ki67+ T cells
CD4+CD38+ T cells
CD8+ T cells
CD8+Ki67+ T cells
CD8+CD38+ T cells
* p<0.05
** p<0.01
*** p<0.001
Spearman's r
H
Ileum
B Cells
Naive B cells
SM B Cells
GC B Cells
CD4+ T cells
CD4+Ki67+ T cells
CD4+CD38+ T cells
CD8+ T cells
CD8+Ki67+ T cells
CD8+CD38+ T cells
* p<0.05
** p<0.01
*** p<0.001
Spearman's r
I
Interferon alpha/beta signaling
Signaling by FGFR4 in disease
RHO GTPases activate IQGAPs
MASTL Facilitates Mitotic Progression
Regulation of IFNA signaling
TRAF3–dependent IRF activation pathway
Signaling by cytosolic FGFR1 fusion mutants
Regulation of necroptotic cell death
Interleukin–6 signaling
CASP8 activity is inhibited
Regulation by c–FLIP
Dimerization of procaspase–8
WNT5A–dependent internalization of FZD2, FZD5 and ROR2
Mitotic Telophase/Cytokinesis
AMER1 mutants destabilize the destruction complex
AXIN mutants destabilize the destruction complex, activating WNT signaling
Establishment of Sister Chromatid Cohesion
IRF3–mediated induction of type I IFN
Retrograde neurotrophin signalling
Endosomal/Vacuolar pathway
Folding of actin by CCT/TriC
NC
PWSH
PWVH
mean PW.Z score
Reactome pathways on TFH-like cells in the ileum

Figure 6

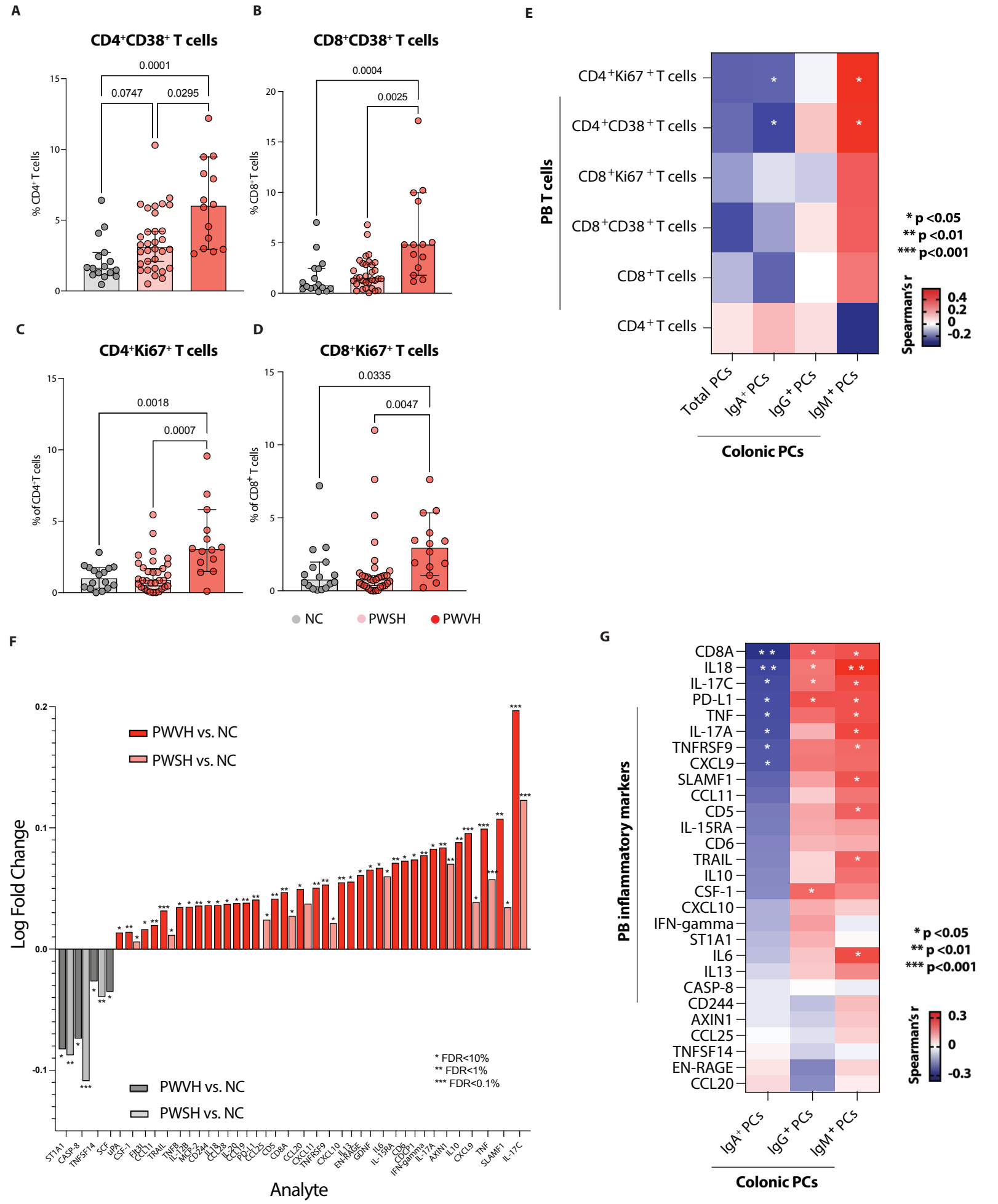

A
CD4+CD38+ T cells
0.0001
0.0747
0.0295
% CD4+ T cells
B
CD8+CD38+ T cells
0.0004
0.0025
% CD8+ T cells
C
CD4+Ki67+ T cells
0.0018
0.0007
% of CD4+T cells
D
CD8+Ki67+ T cells
0.0335
0.0047
% of CD8+ T cells
NC
PWSH
PWVH
E
PB T cells
CD4+Ki67+ T cells
CD4+CD38+ T cells
CD8+Ki67+ T cells
CD8+CD38+ T cells
CD8+ T cells
CD4+ T cells
Total PCs
IgA+ PCs
IgG+ PCs
IgM+ PCs
Colonic PCs
* p <0.05
** p <0.01
*** p<0.001
Spearman's r
0.4
0.2
0
-0.2
F
PWVH vs. NC
PWSH vs. NC
Log Fold Change
* FDR<10%
** FDR<1%
*** FDR<0.1%
Analyte
G
PB inflammatory markers
CD8A
IL18
IL-17C
PD-L1
TNF
IL-17A
TNFRSF9
CXCL9
SLAMF1
CCL11
CD5
IL-15RA
CD6
TRAIL
IL10
CSF-1
CXCL10
IFN-gamma
ST1A1
IL6
IL13
CASP-8
CD244
AXIN1
CCL25
TNFSF14
EN-RAGE
CCL20
IgA+ PCs
IgG+ PCs
IgM+ PCs
Colonic PCs
* p <0.05
** p <0.01
*** p<0.001
Spearman's r
0.3
0
-0.3

Figure 7

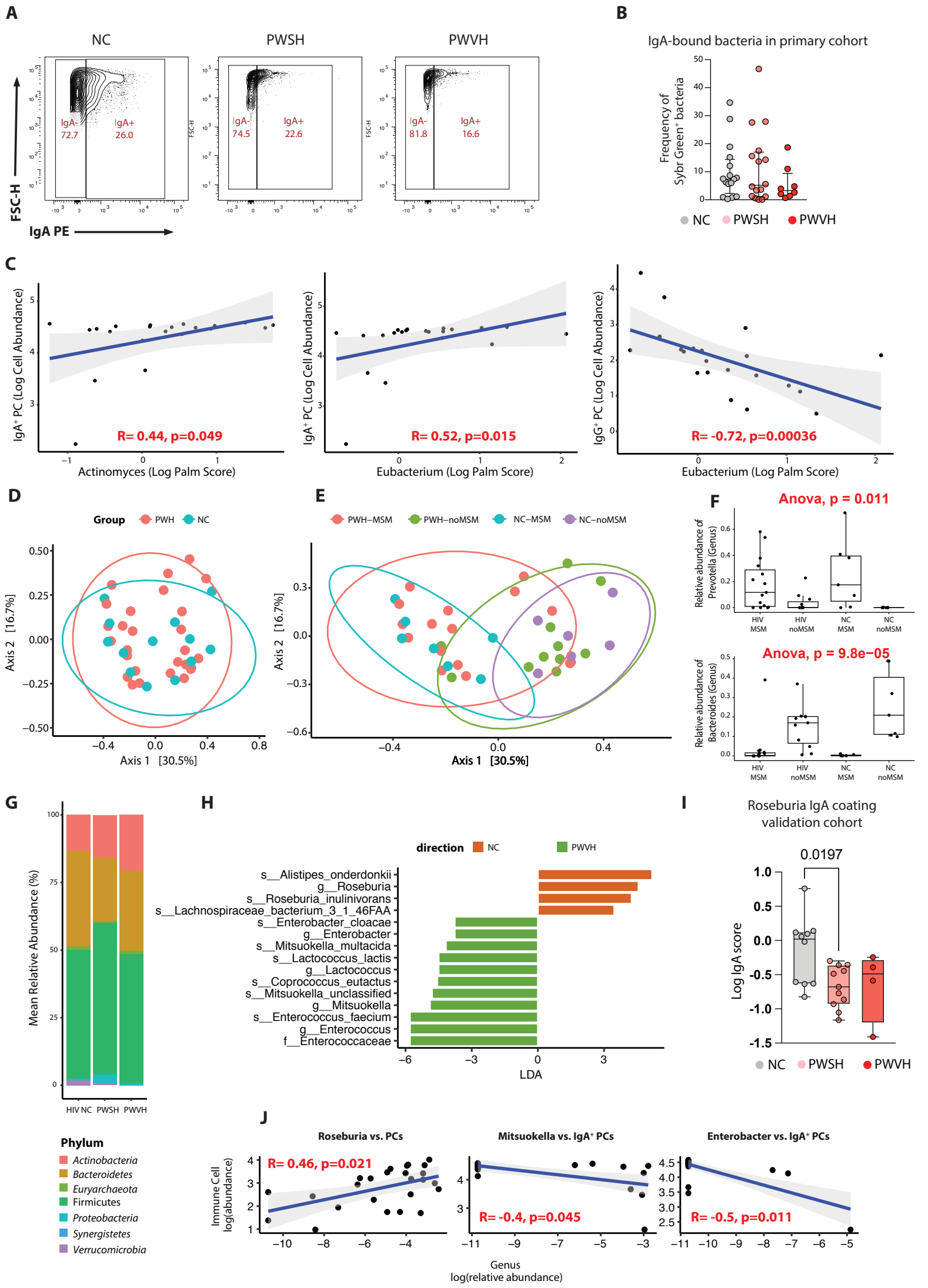

A
NC
PWSH
PWVH
IgA-
72.7
IgA+
26.0
IgA-
74.5
IgA+
22.6
IgA-
81.8
IgA+
16.6
FSC-H
IgA PE
B
IgA-bound bacteria in primary cohort
Frequency of Sybr Green⁺ bacteria
NC
PWSH
PWVH
C
IgA⁺ PC (Log Cell Abundance)
Actinomyces (Log Palm Score)
R= 0.44, p=0.049
IgA⁺ PC (Log Cell Abundance)
Eubacterium (Log Palm Score)
R= 0.52, p=0.015
IgG⁺ PC (Log Cell Abundance)
Eubacterium (Log Palm Score)
R= -0.72, p=0.00036
D
Group
PWH
NC
Axis 2 [16.7%]
Axis 1 [30.5%]
E
PWH−MSM
PWH−noMSM
NC−MSM
NC−noMSM
Axis 2 [16.7%]
Axis 1 [30.5%]
F
Anova, p = 0.011
Relative abundance of Prevotella (Genus)
HIV MSM
HIV noMSM
NC MSM
NC noMSM
Anova, p = 9.8e−05
Relative abundance of Bacteroides (Genus)
G
Mean Relative Abundance (%)
HIV NC
PWSH
PWVH
Phylum
Actinobacteria
Bacteroidetes
Euryarchaeota
Firmicutes
Proteobacteria
Synergistetes
Verrucomicrobia
H
direction
NC
PWVH
s__Alistipes_onderdonkii
g__Roseburia
s__Roseburia_inulinivorans
s__Lachnospiraceae_bacterium_3_1_46FAA
s__Enterobacter_cloacae
g__Enterobacter
s__Mitsuokella_multacida
s__Lactococcus_lactis
g__Lactococcus
s__Coprococcus_eutactus
s__Mitsuokella_unclassified
g__Mitsuokella
s__Enterococcus_faecium
g__Enterococcus
f__Enterococcaceae
LDA
I
Roseburia IgA coating validation cohort
0.0197
Log IgA score
NC
PWSH
PWVH
J
Roseburia vs. PCs
R= 0.46, p=0.021
Mitsuokella vs. IgA⁺ PCs
R= -0.4, p=0.045
Enterobacter vs. IgA⁺ PCs
R= -0.5, p=0.011
Immune Cell log(abundance)
Genus log(relative abundance)

Figure S1

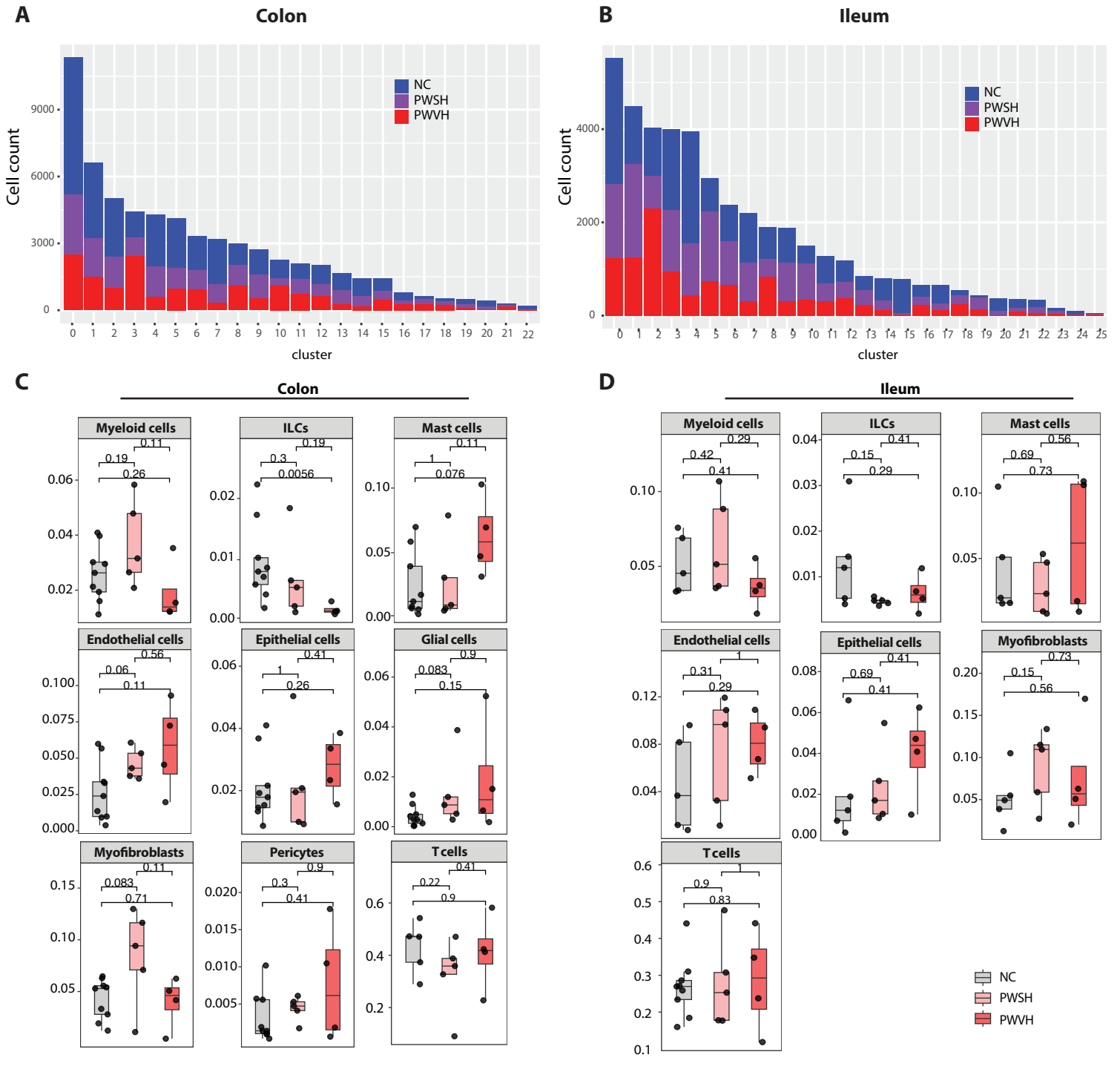
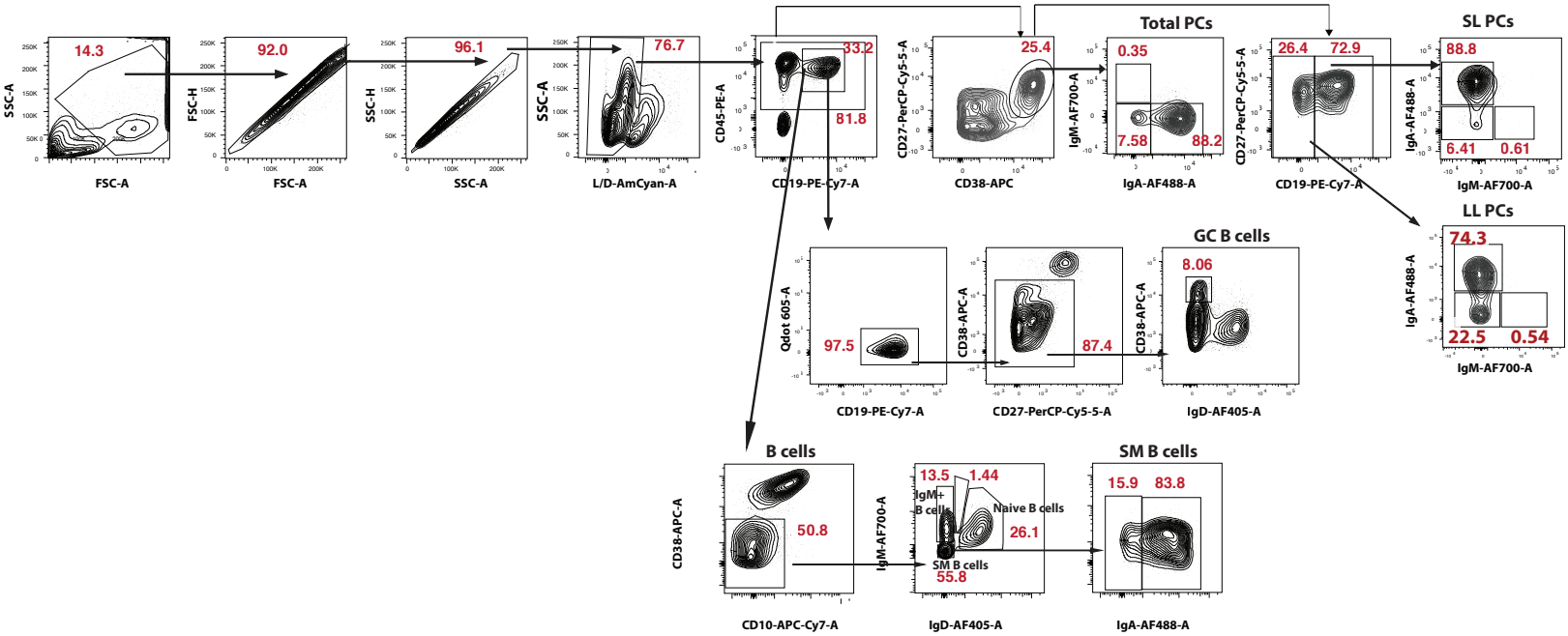
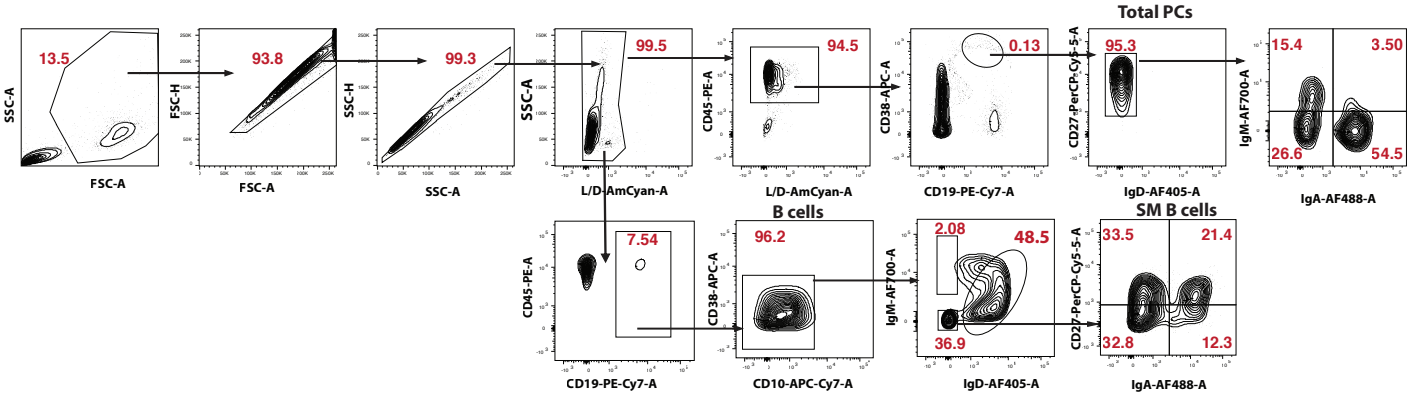

A
Colon
Cell count
cluster
NC
PWSH
PWVH
B
Ileum
Cell count
cluster
NC
PWSH
PWVH
C
Colon
Myeloid cells
ILCs
Mast cells
Endothelial cells
Epithelial cells
Glial cells
Myofibroblasts
Pericytes
T cells
D
Ileum
Myeloid cells
ILCs
Mast cells
Endothelial cells
Epithelial cells
Myofibroblasts
T cells
NC
PWSH
PWVH

Figure S2

A Gating strategies for intestinal B cells and PCs



B Gating strategies for peripheral blood B cells and PCs



C Gating strategies for activated and cycling T cells

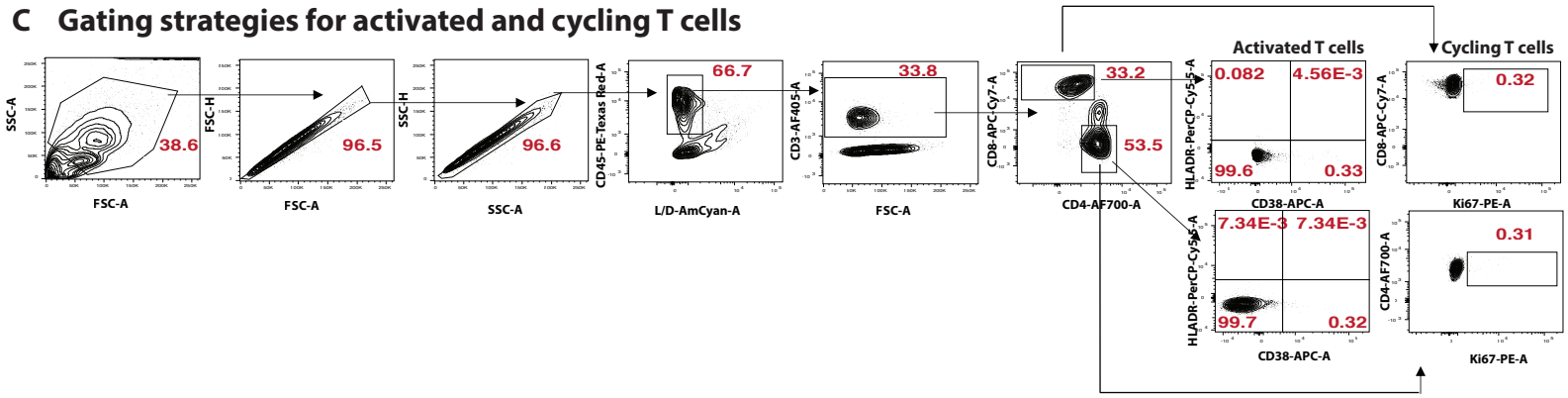
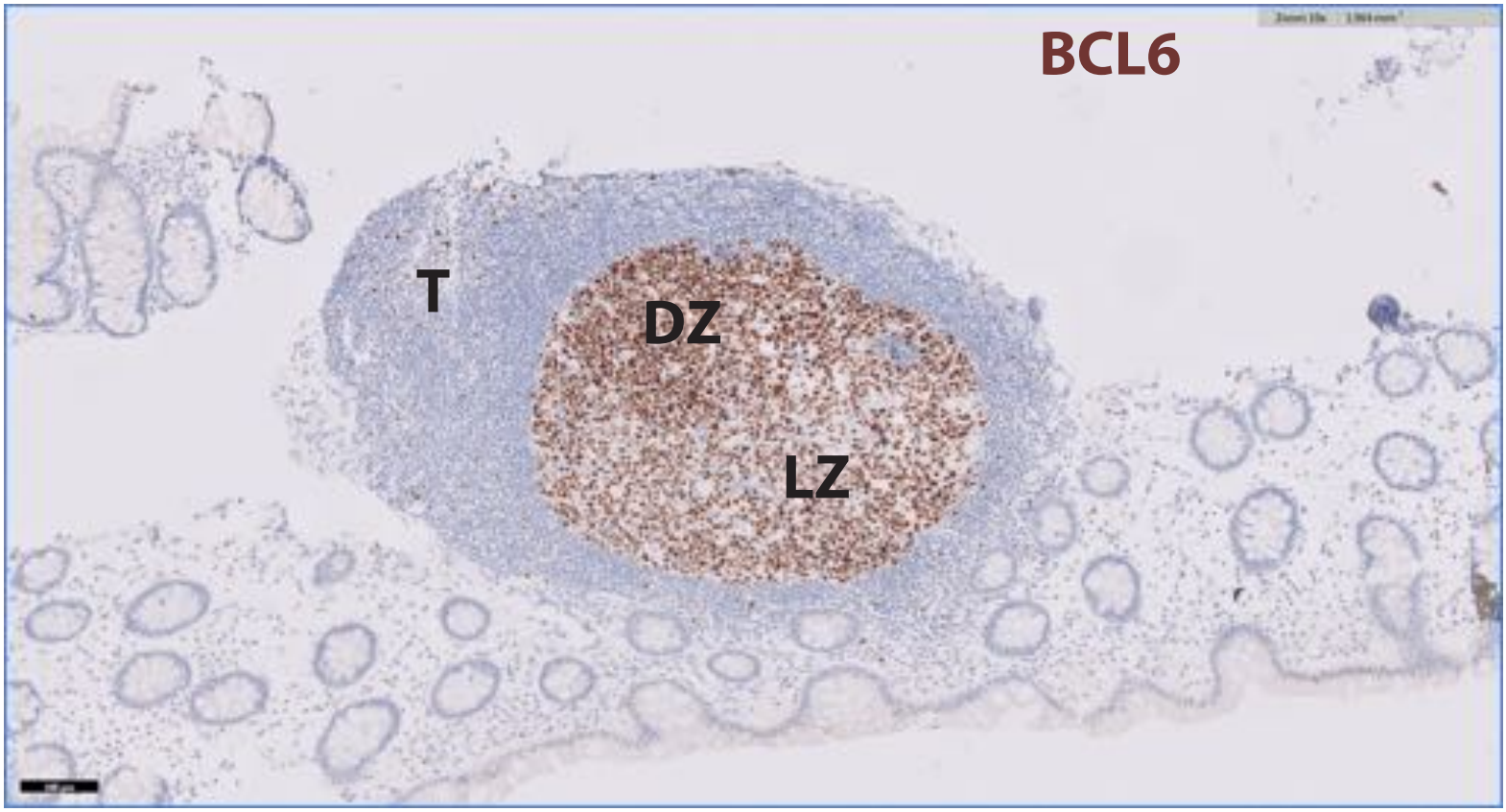


Figure S3



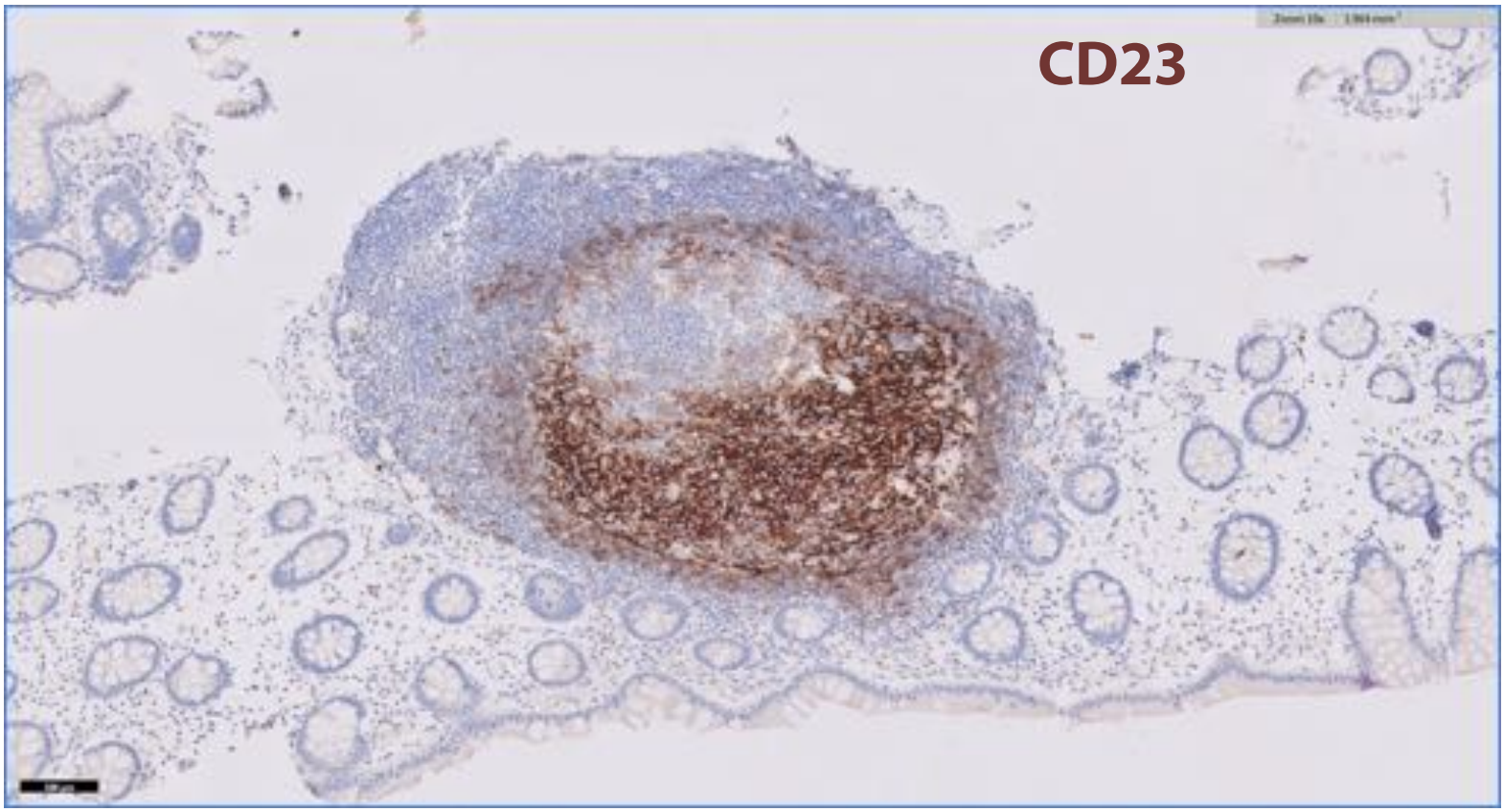
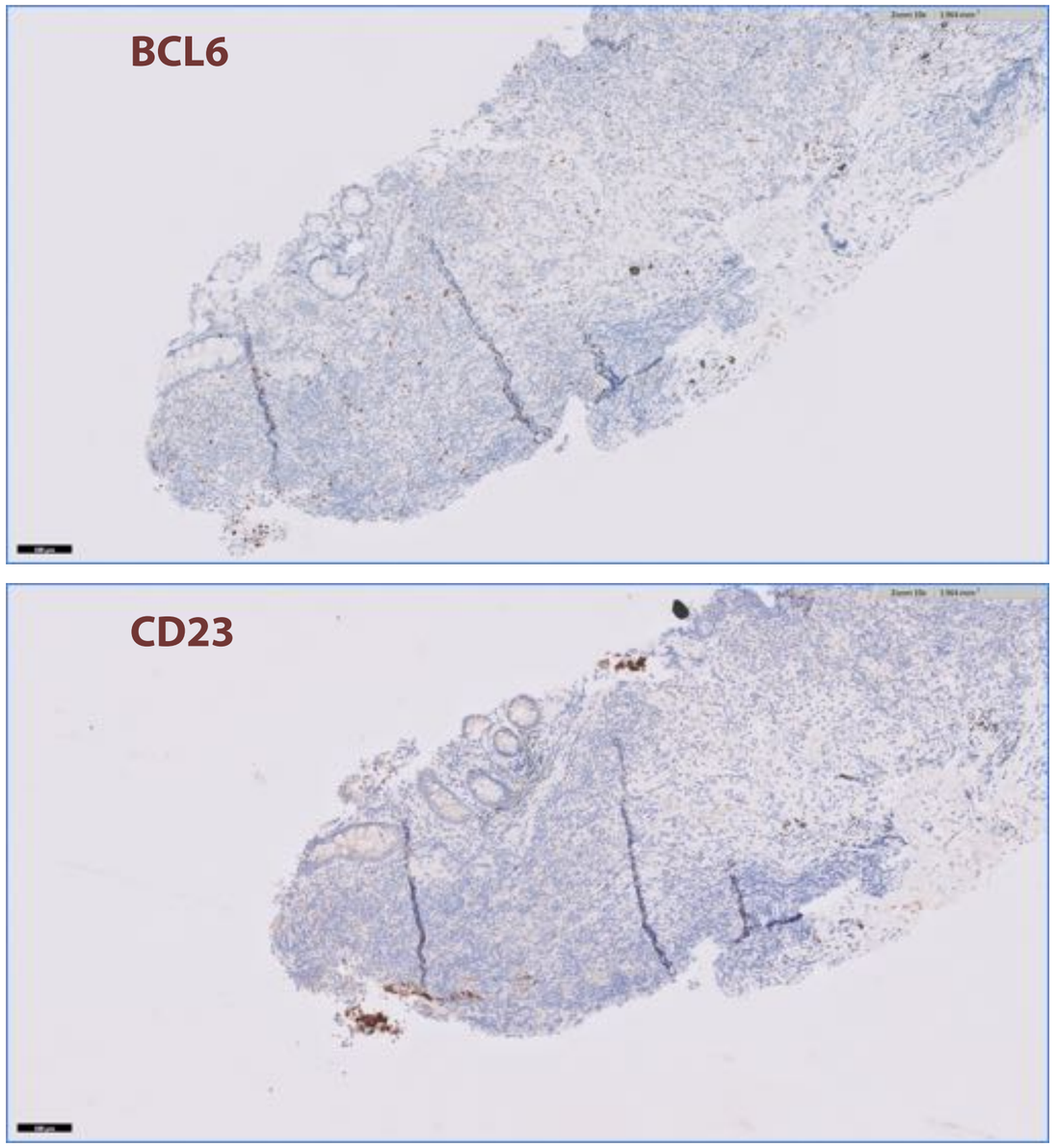
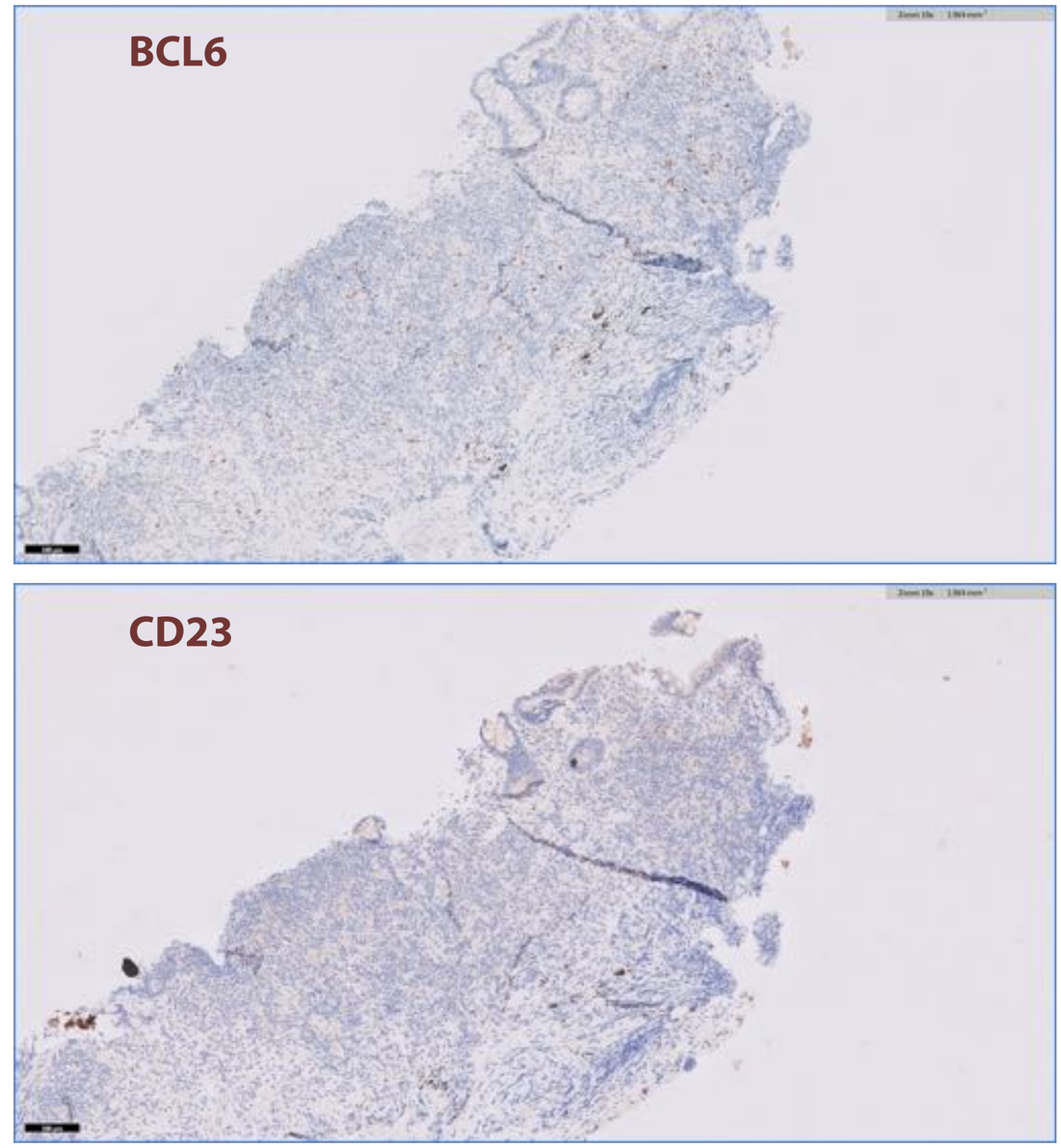


Figure S4

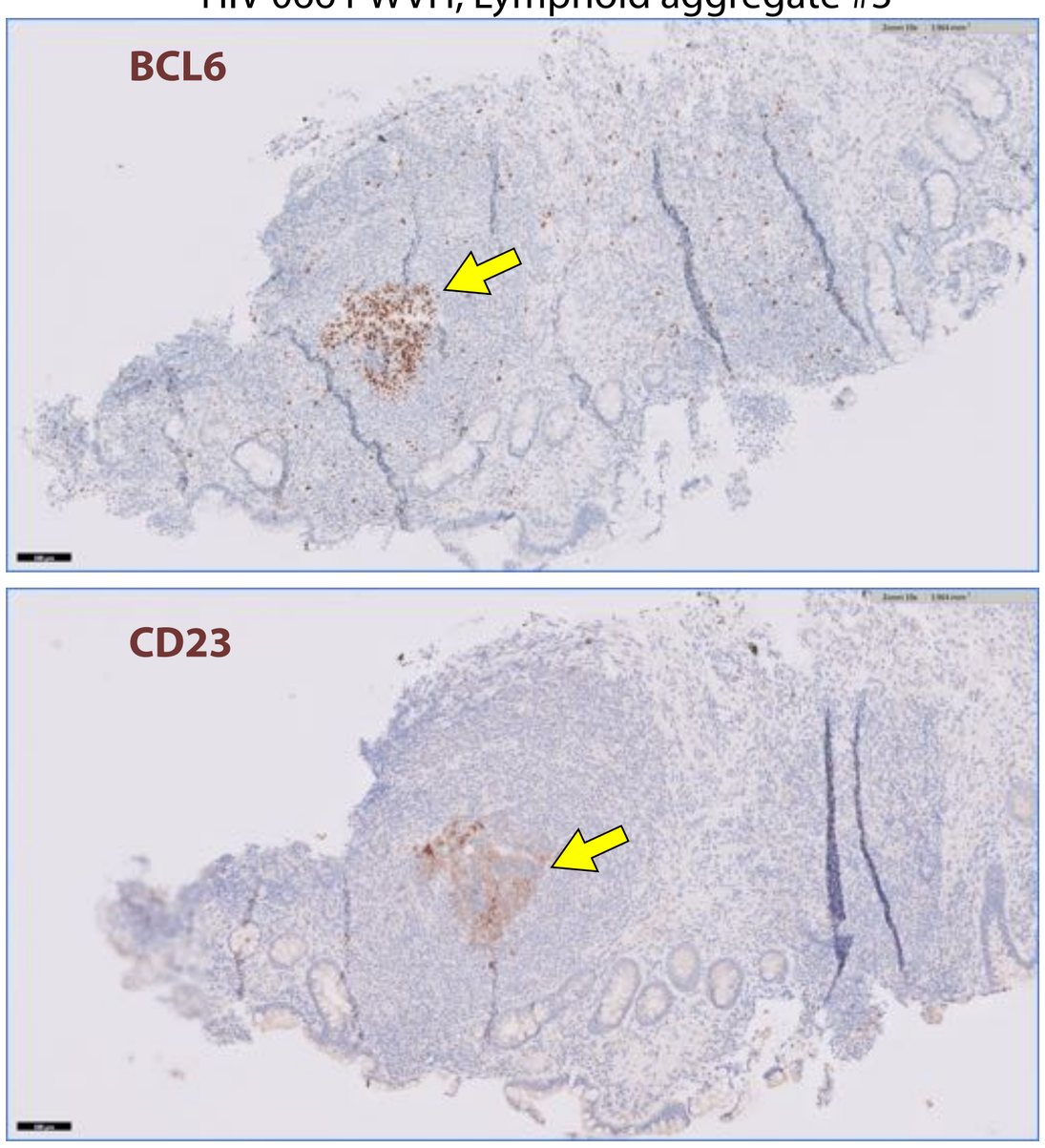
HIV 066 PWVH, Lymphoid aggregate #1



HIV 066 PWVH, Lymphoid aggregate #2



HIV 066 PWVH, Lymphoid aggregate #3



HIV 066 PWVH, Lymphoid aggregate #4

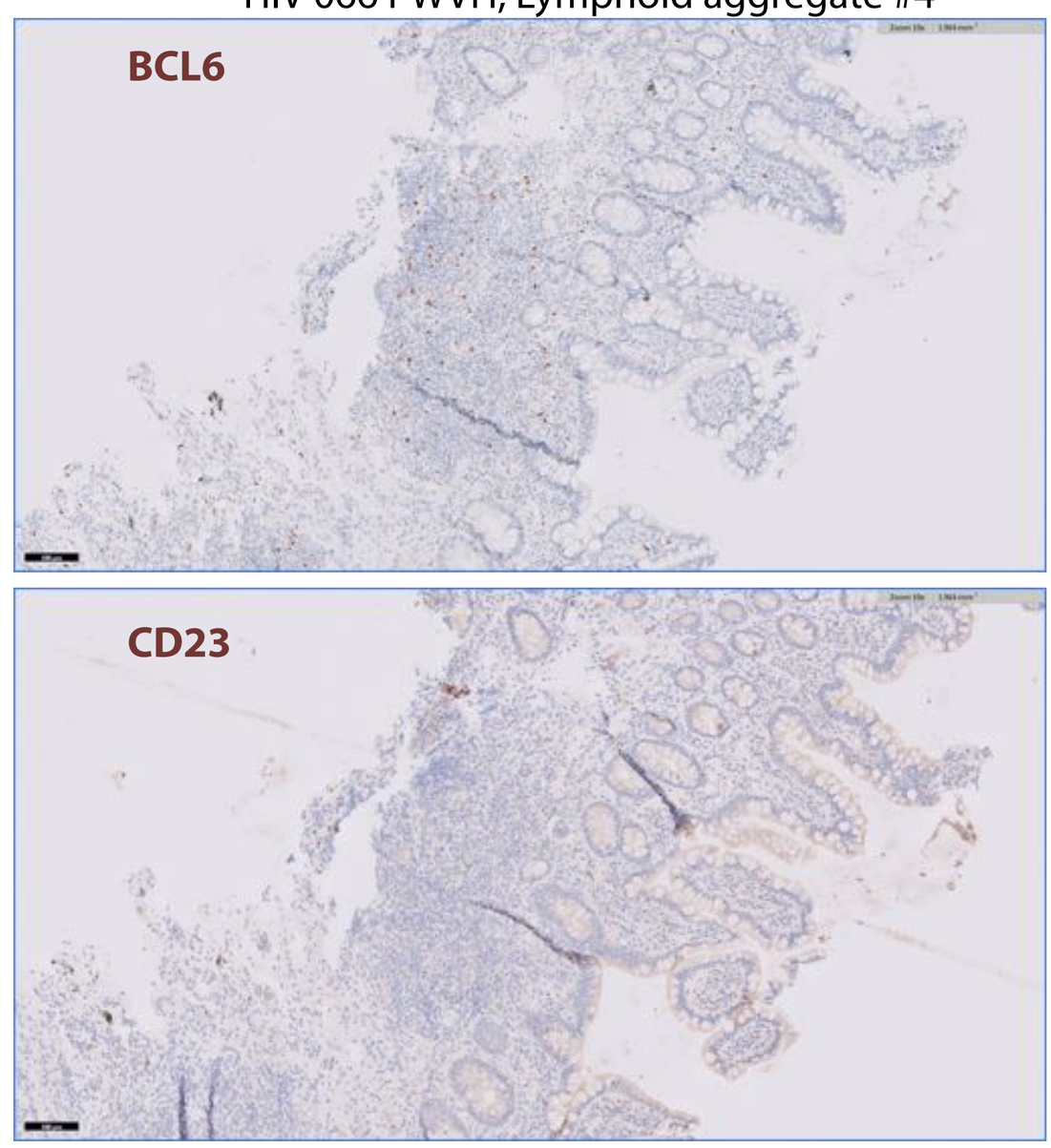


Figure S5

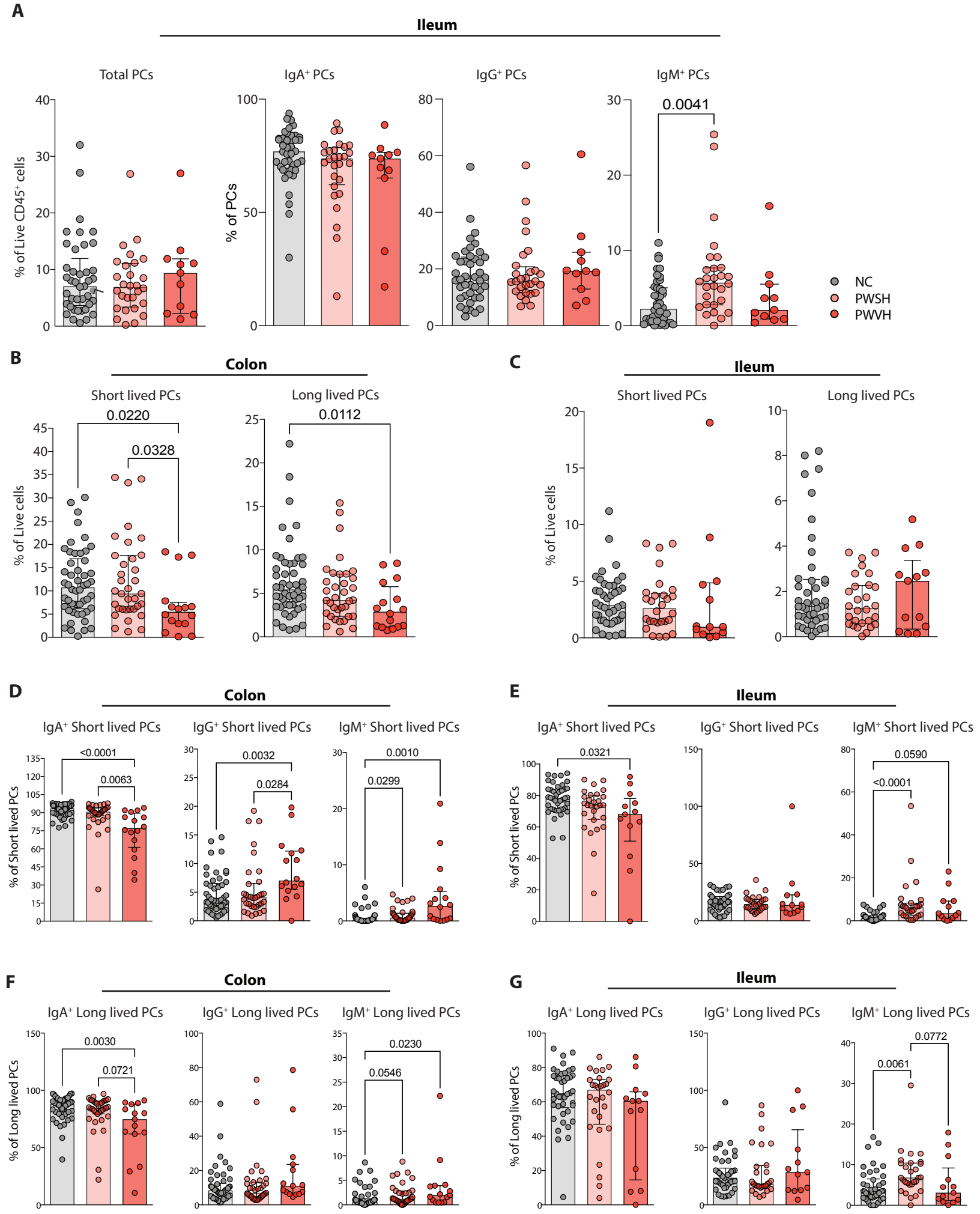

A
Ileum
Total PCs
IgA+ PCs
IgG+ PCs
IgM+ PCs
% of Live CD45+ cells
% of PCs
0.0041
NC
PWSH
PWVH
B
Colon
Short lived PCs
Long lived PCs
% of Live cells
0.0220
0.0328
0.0112
C
Ileum
Short lived PCs
Long lived PCs
% of Live cells
D
Colon
IgA+ Short lived PCs
IgG+ Short lived PCs
IgM+ Short lived PCs
% of Short lived PCs
<0.0001
0.0063
0.0032
0.0284
0.0010
0.0299
E
Ileum
IgA+ Short lived PCs
IgG+ Short lived PCs
IgM+ Short lived PCs
% of Short lived PCs
0.0321
0.0590
<0.0001
F
Colon
IgA+ Long lived PCs
IgG+ Long lived PCs
IgM+ Long lived PCs
% of Long lived PCs
0.0030
0.0721
0.0230
0.0546
G
Ileum
IgA+ Long lived PCs
IgG+ Long lived PCs
IgM+ Long lived PCs
% of Long lived PCs
0.0772
0.0061

Figure S6

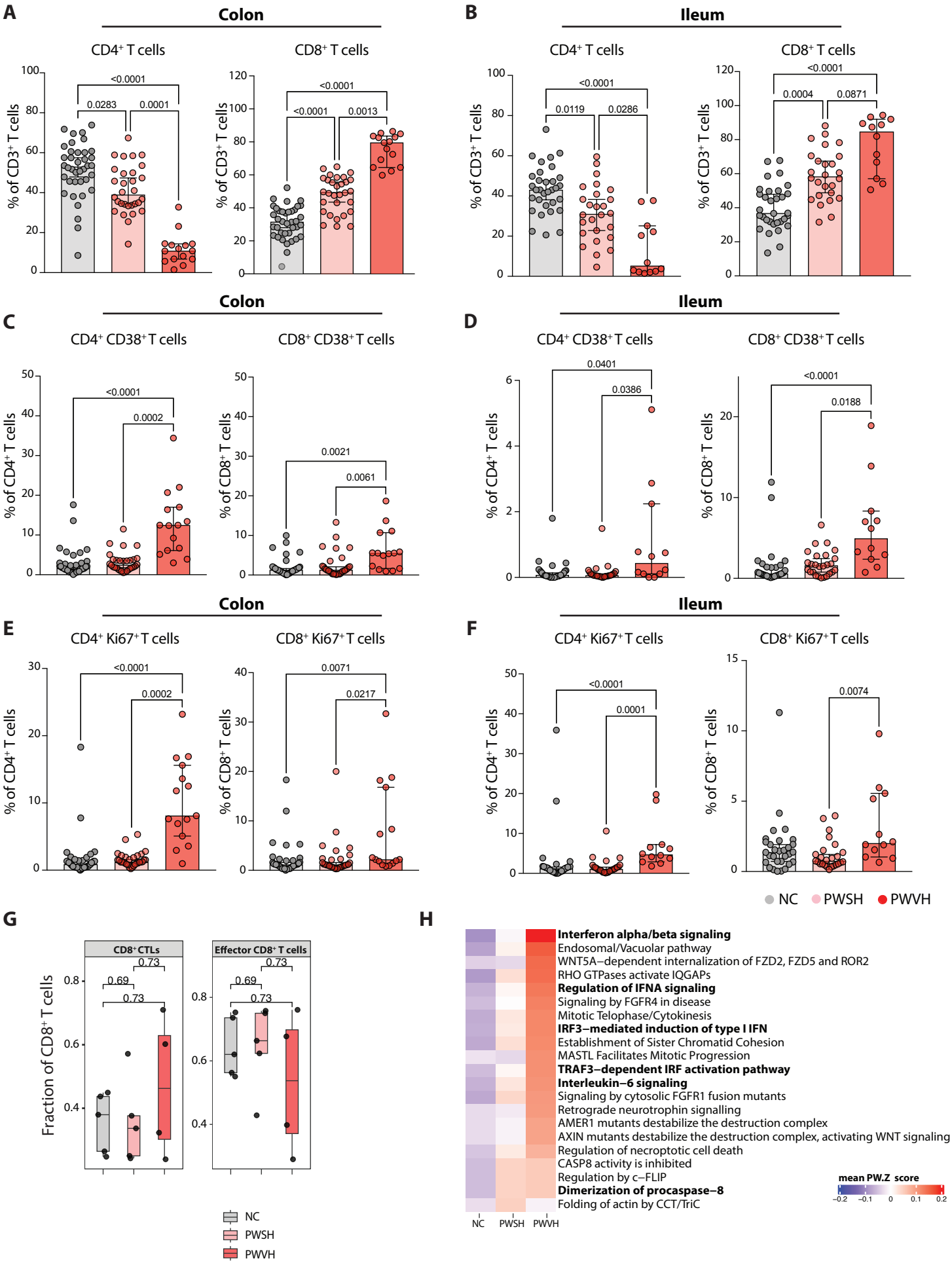
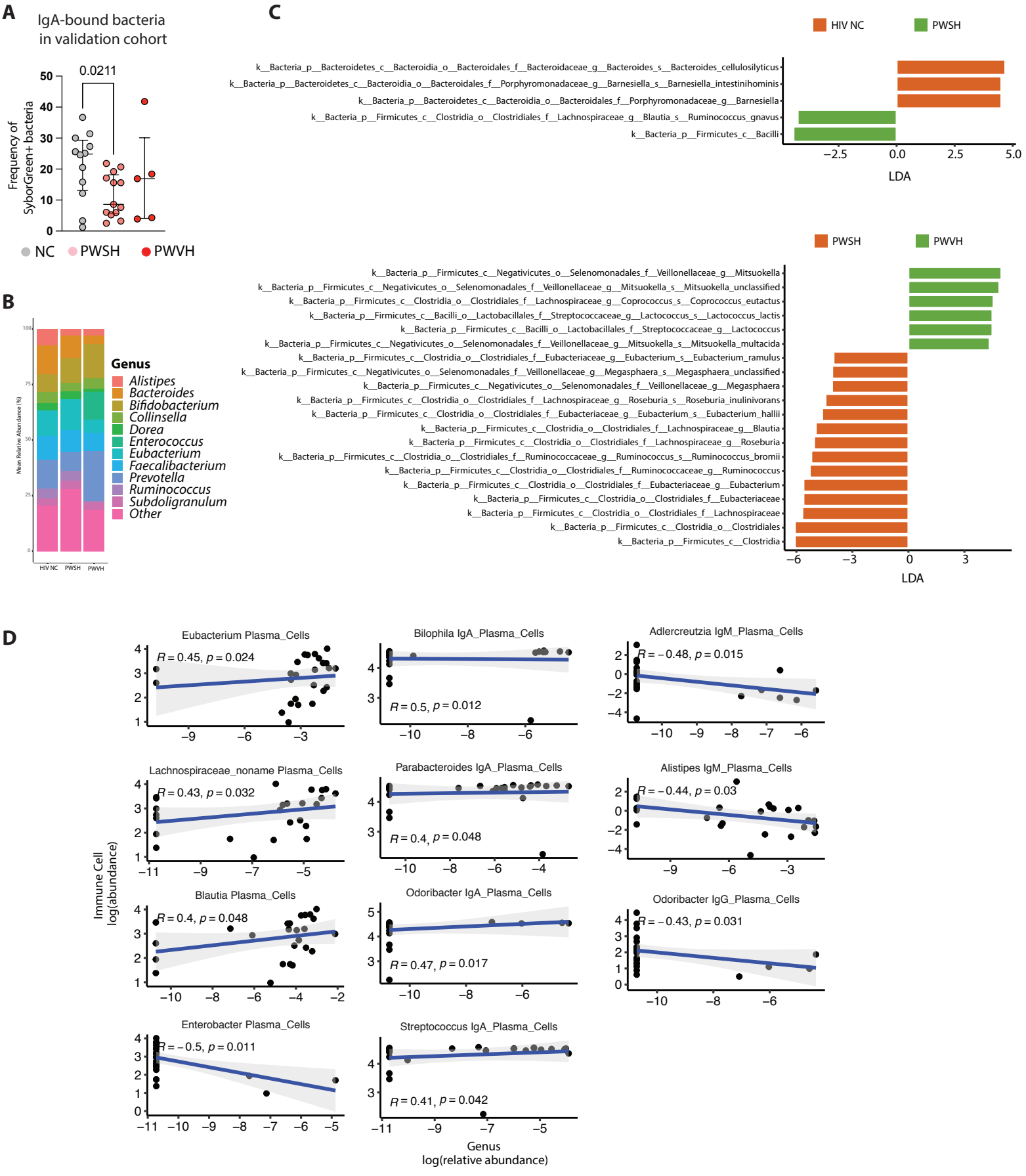

A
Colon
CD4+ T cells
CD8+ T cells
% of CD3+ T cells
<0.0001
0.0283
0.0001
<0.0001
0.0013
B
Ileum
0.0119
0.0286
0.0004
0.0871
C
CD4+ CD38+ T cells
CD8+ CD38+ T cells
% of CD4+ T cells
% of CD8+ T cells
0.0002
0.0021
0.0061
D
0.0401
0.0386
0.0188
E
CD4+ Ki67+ T cells
CD8+ Ki67+ T cells
0.0071
0.0217
F
0.0074
NC
PWSH
PWVH
G
CD8+ CTLs
Effector CD8+ T cells
Fraction of CD8+ T cells
0.69
0.73
H
Interferon alpha/beta signaling
Endosomal/Vacuolar pathway
WNT5A–dependent internalization of FZD2, FZD5 and ROR2
RHO GTPases activate IQGAPs
Regulation of IFNA signaling
Signaling by FGFR4 in disease
Mitotic Telophase/Cytokinesis
IRF3–mediated induction of type I IFN
Establishment of Sister Chromatid Cohesion
MASTL Facilitates Mitotic Progression
TRAF3–dependent IRF activation pathway
Interleukin–6 signaling
Signaling by cytosolic FGFR1 fusion mutants
Retrograde neurotrophin signalling
AMER1 mutants destabilize the destruction complex
AXIN mutants destabilize the destruction complex, activating WNT signaling
Regulation of necroptotic cell death
CASP8 activity is inhibited
Regulation by c–FLIP
Dimerization of procaspase–8
Folding of actin by CCT/TriC
mean PW.Z score
–0.2
–0.1
0
0.1
0.2

Figure S7


A
IgA-bound bacteria in validation cohort
0.0211
Frequency of SyborGreen+ bacteria
NC
PWSH
PWVH
B
Genus
Alistipes
Bacteroides
Bifidobacterium
Collinsella
Dorea
Enterococcus
Eubacterium
Faecalibacterium
Prevotella
Ruminococcus
Subdoligranulum
Other
Mean Relative Abundance (%)
HIV NC
PWSH
PWVH
C
HIV NC
PWSH
k__Bacteria_p__Bacteroidetes_c__Bacteroidia_o__Bacteroidales_f__Bacteroidaceae_g__Bacteroides_s__Bacteroides_cellulosilyticus
k__Bacteria_p__Bacteroidetes_c__Bacteroidia_o__Bacteroidales_f__Porphyromonadaceae_g__Barnesiella_s__Barnesiella_intestinihominis
k__Bacteria_p__Bacteroidetes_c__Bacteroidia_o__Bacteroidales_f__Porphyromonadaceae_g__Barnesiella
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Lachnospiraceae_g__Blautia_s__Ruminococcus_gnavus
k__Bacteria_p__Firmicutes_c__Bacilli
LDA
PWSH
PWVH
k__Bacteria_p__Firmicutes_c__Negativicutes_o__Selenomonadales_f__Veillonellaceae_g__Mitsuokella
k__Bacteria_p__Firmicutes_c__Negativicutes_o__Selenomonadales_f__Veillonellaceae_g__Mitsuokella_s__Mitsuokella_unclassified
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Lachnospiraceae_g__Coprococcus_s__Coprococcus_eutactus
k__Bacteria_p__Firmicutes_c__Bacilli_o__Lactobacillales_f__Streptococcaceae_g__Lactococcus_s__Lactococcus_lactis
k__Bacteria_p__Firmicutes_c__Bacilli_o__Lactobacillales_f__Streptococcaceae_g__Lactococcus
k__Bacteria_p__Firmicutes_c__Negativicutes_o__Selenomonadales_f__Veillonellaceae_g__Mitsuokella_s__Mitsuokella_multacida
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Eubacteriaceae_g__Eubacterium_s__Eubacterium_ramulus
k__Bacteria_p__Firmicutes_c__Negativicutes_o__Selenomonadales_f__Veillonellaceae_g__Megasphaera_s__Megasphaera_unclassified
k__Bacteria_p__Firmicutes_c__Negativicutes_o__Selenomonadales_f__Veillonellaceae_g__Megasphaera
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Lachnospiraceae_g__Roseburia_s__Roseburia_inulinivorans
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Eubacteriaceae_g__Eubacterium_s__Eubacterium_hallii
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Lachnospiraceae_g__Blautia
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Lachnospiraceae_g__Roseburia
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Ruminococcaceae_g__Ruminococcus_s__Ruminococcus_bromii
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Ruminococcaceae_g__Ruminococcus
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Eubacteriaceae_g__Eubacterium
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Eubacteriaceae
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Lachnospiraceae
k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales
k__Bacteria_p__Firmicutes_c__Clostridia
LDA
D
Eubacterium Plasma_Cells
R = 0.45, p = 0.024
Bilophila IgA_Plasma_Cells
R = 0.5, p = 0.012
Adlercreutzia IgM_Plasma_Cells
R = −0.48, p = 0.015
Lachnospiraceae_noname Plasma_Cells
R = 0.43, p = 0.032
Parabacteroides IgA_Plasma_Cells
R = 0.4, p = 0.048
Alistipes IgM_Plasma_Cells
R = −0.44, p = 0.03
Blautia Plasma_Cells
R = 0.4, p = 0.048
Odoribacter IgA_Plasma_Cells
R = 0.47, p = 0.017
Odoribacter IgG_Plasma_Cells
R = −0.43, p = 0.031
Enterobacter Plasma_Cells
R = −0.5, p = 0.011
Streptococcus IgA_Plasma_Cells
R = 0.41, p = 0.042
Immune Cell log(abundance)
Genus log(relative abundance)