

**SUPPLEMENTAL INFORMATION**

**Multimodal profiling of peripheral blood identifies proliferating circulating effector CD4<sup>+</sup> T cells as predictors for response to integrin  $\alpha$ 4 $\beta$ 7-blocking therapy in patients with inflammatory bowel disease**

Veronika Horn<sup>1,2,13,14</sup>, Camila Cancino<sup>1,2,13</sup>, Lisa Steinheuer<sup>3,13</sup>, Benedikt Obermayer<sup>4,13</sup>, *et al*

<sup>1</sup>Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Gastroenterology, Infectious Diseases and Rheumatology, 12203 Berlin, Germany

<sup>2</sup>Deutsches Rheuma-Forschungszentrum, ein Institut der Leibniz-Gemeinschaft, 10117 Berlin, Germany

<sup>3</sup>Institute for Experimental Oncology, Biomathematics Division, University Hospital Bonn, Bonn, Germany

<sup>4</sup>Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Core Unit Bioinformatics, 10117 Berlin, Germany

<sup>13</sup>Equal contribution

**Corresponding author:** Prof. Ahmed N. Hegazy, M.D, Ph.D., Department of Gastroenterology, Infectious Diseases and Rheumatology, Charité - Universitätsmedizin Berlin, Campus Benjamin Franklin, Hindenburgdamm 30, 12203 Berlin, Germany; and Deutsches Rheuma-Forschungszentrum, ein Institut der Leibniz-Gemeinschaft, 10117 Berlin, Germany. Phone: +49 30 450 614 378; E-mail: [ahmed.hegazy@charite.de](mailto:ahmed.hegazy@charite.de) and [ahmed.hegazy@drfz.de](mailto:ahmed.hegazy@drfz.de)

## SUPPLEMENTARY EXPERIMENTAL PROCEDURES

**Sample collection and banking.** Peripheral blood samples were collected from both IBD patients and healthy donors (HD) at defined time points before and during therapy (**Figure 1A**) using heparin and serum tubes. Samples were collected before the first infusion of vedolizumab (week 0), and before the following infusions (week 2, or 6: before the 2<sup>nd</sup> or 3<sup>rd</sup> infusion at week 2 or 6). For peripheral blood mononuclear cells (PBMCs) isolation, blood was diluted 1:1 in 1x PBS and subjected to density gradient centrifugation using Pancoll human (PAN-Biotech, Aidenbach, Germany; P06-1391500) at 800 x g for 20 minutes at room temperature. The resulting PBMC fraction was then resuspended in a freezing medium composed of 90% FCS and 10% DMSO. Subsequently, the PBMCs were frozen using a Cryo-freezing container (Nalgene, Rochester, NY) and stored in liquid nitrogen for long-term preservation. A proteomic stabilizer PROT-1 (Smart Tube Inc., Las Vegas, NV) was used for whole blood cryopreservation according to the manufacturer's instructions, and cryopreserved whole blood samples were stored at -80°C. Patient serum samples were obtained by centrifugation at 2000 x g for 7 minutes at room temperature. The resulting serum was immediately frozen at -80°C for subsequent analysis. We did not observe any impact of freezing and thawing on the expression of the integrins of interest (**Technical Supplementary Figure 1**).

**PBMC thawing for stimulation.** Frozen PBMCs were thawed in a 37°C water bath and then transferred to a preheated thawing medium consisting of RPMI 1640 Medium GlutaMAX™ Supplement, 10% FCS, 1% penicillin/streptomycin, 50 µM 2-mercaptoethanol, and 50 U/mL DNase I (Sigma-Aldrich, St Louis, MO). Cells were centrifuged at 350 x g for 10 minutes at room temperature. The cells were then resuspended in RPMI 1640 Medium GlutaMAX™ Supplement, 10% FCS, 1% penicillin/streptomycin, and 50 µM 2-mercaptoethanol, counted, and 1 million cells were used for flow cytometry.

**Stimulation of PBMCs with PMA/Ionomycin.** Thawed PBMCs were incubated with 100  $\mu$ l of a stimulation mixture containing 5 ng/ml PMA (Sigma-Aldrich, St Louis, MO; P1585-1MG) and 500 ng/ml ionomycin (Sigma-Aldrich, St Louis, MO; I0634-1MG,) at 37°C for 30 minutes. After stimulation, 50  $\mu$ l of brefeldin A at a concentration of 5  $\mu$ g/ml (Cayman Chemical, Ann Arbor, MI; Cay11861-25) was added, and the cells were incubated for further 3.5 hours at 37°C. After a total of four hours of stimulation, the cells were harvested for intracellular cytokine (IC) staining.

**Flow cytometry.** PBMCs were stained with surface antibodies in FACS buffer (1x PBS, 0.05% BSA, 0.01% NaN<sub>3</sub>, 2 mM EDTA; **Suppl. Table 8**) for 30 minutes. We did not observe an impact of Vedolizumab binding on the detection of Integrin  $\alpha$ 4 or Integrin  $\beta$ 7 (**Technical Supplementary Figure 2**). Cells were then fixed and permeabilized with the eBioscience™ Foxp3/Transcription Factor Staining Buffer Set (Invitrogen, Waltham, MA) to allow intracellular staining for Ki67 and transcription factors. For intracellular cytokine staining, PBMCs were fixed with 1x BD FACS Lysing Solution (BD Biosciences, Franklin Lakes, NJ) following stimulation with PMA/ionomycin, and surface staining, and live/dead staining (**Suppl. Table 8**). Subsequently, intracellular cytokines were stained after permeabilization with 0.05% saponin (Sigma-Aldrich). The gating strategy to identify CD4<sup>+</sup> memory T cells and the expression of various surface markers and staining controls are shown in **Technical Supplementary Figures 3, 4, and 5**. Cell suspensions were supplemented with Precision Count Beads (BioLegend, San Diego, CA) and then acquired on a BD FACSymphony (Configuration 5B 8V 3R 5YG 7UV). Daily quality control was performed using Sphero Rainbow Calibration Particles (BD Biosciences, Franklin Lakes, NJ). Flow cytometry data was analyzed using FlowJo software (BD, V10.4), and subsequent analyses were performed using R and Prism v8 (GraphPad Software, La Jolla, CA). Clinical information regarding the patients included in the flow cytometry analysis can be found in **Suppl. Tables 7 and 9**.

**Cell sorting for scRNAseq.** Thawed PBMCs were sorted using the SONY MA900 sorter with a 100-micron nozzle. CD45<sup>+</sup> cells were sorted based on CD45 expression while eliminating DAPI-positive cells as shown in **Technical Supplementary Figure 8**. The sorted CD45<sup>+</sup> cells were stained with the TotalSeq™ antibody mixture according to manufacturer's protocol (BioLegend, **Suppl. Table 12**). CD4<sup>+</sup> memory T cells, extracted from the same donors as those used for total PBMCs scRNAseq, were sorted on the BD FACSAria II with a 70-micron nozzle. Sorting criteria included CD3, CD4, and CD45RA expression, with the exclusion of non-viable cells by DAPI staining (**Technical Supplementary Figure 8**). To differentiate between individual donors, CD4<sup>+</sup> memory T cells were labeled with hashtags after sorting (BioLegend, **Suppl. Table 12**).

**Inhibition of VCAM-1 Binding in memory CD4<sup>+</sup> T Lymphocytes.** The assay was done as previously described by *Soler et al.*<sup>1</sup>. A 100 µl volume of fresh blood, supplemented with 4 mM MnCl<sub>2</sub>, was incubated for 1 hour at room temperature with a concentration curve of Vedolizumab (VDZ) or Natalizumab (NTZ) independently. Following this, 10 µg/ml of human VCAM-1, tagged with His, was added to a final volume of 120 µl and incubated for 1 hour at room temperature. The cells were then washed with assay buffer (25 mM Tris, 4 mM MnCl<sub>2</sub>, 2.7 mM KCl, 150 mM NaCl, 0.5% BSA, pH 7.2) and stained with anti-His PE (BioLegend) for 15 minutes at room temperature. After washing, the cells were incubated with a mix of antibodies (Live/Dead, CD3, CD4, CD45RA, CD49d, β7, CD103, CD29, CD38, and HLA-DR) for 30 minutes at 4°C. The cells were washed again, fixed with BD FACS lysing solution, washed once more, and resuspended in FACS buffer (1x PBS, 0.05% BSA, 0.01% NaN<sub>3</sub>, and 2 mM EDTA). Finally, the cells were analyzed using flow cytometry.

**Cellular barcoding and mass cytometry.** Whole blood samples were fixed with a proteomic stabilizer PROT1 (Smart Tube, Inc.) and stored at -80°C. Then, samples were thawed, stained, and acquired in batches of 15 samples that included all time points from the same patient and

matched healthy controls. An anchor sample was included in each run to control for batch effects. The thawing process and mass cytometry staining were performed following as previously described<sup>2</sup>. After thawing and erythrocyte lysis, total cell counts were obtained using a MACSquant cytometer. Subsequently, 1.5 million cells per sample were barcoded using the Cell-ID Palladium Barcoding Kit (Standard BioTools, South San Francisco, CA) according to the manufacturer's instructions and pooled into a single batch. To minimize nonspecific binding, cells were incubated with 0.2 mg/ml Beriglobin (CSL Behring, King of Prussia, PA) for 10 minutes to block Fc receptors. After washing, the cell suspensions were treated with heparin (Ratiopharm, Ulm, Germany) at a concentration of 100 U/ml for 15 minutes to reduce eosinophil background artifacts. Cell surface staining was performed for 30 minutes in the presence of heparin using the antibody panel listed in **Suppl. Table 10**. After washing, cells were incubated with the appropriate secondary antibody for 15 minutes at room temperature. Antibodies were either purchased pre-labeled (Standard BioTools), labeled in-house with platinum, or labeled using MAXPAR X8 labeling kits (Standard BioTools) according to the manufacturer's instructions. Antibody cocktails were prepared in advance and thawed immediately prior to staining.

After staining, cells were washed and suspended in 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA) overnight. The next day, cells were washed again and counterstained with an iridium-based DNA intercalator (Standard BioTools). Cell suspensions were repeatedly washed with deionized water, passed through a 30 µm cell strainer (Corning, Corning, NY), and spiked with 10% (v/v) EQ Four Element Calibration Beads (Standard BioTools). Daily quality control and adjustment of the Helios mass cytometer (Standard BioTools) were performed prior to acquisition. A total of 300,000 events per sample were acquired at a rate of 200-300 events per second. The samples were acquired in two cohorts and

normalized to account for any batch-dependent effects. Clinical information regarding the patients included in the mass cytometry analysis can be found in **Suppl. Table 11**.

#### **Mass cytometry data processing and algorithm-based high-dimensional analysis.**

Normalization to the spiked-in beads was performed immediately after acquisition using Fluidigm software (Standard BioTools) based on EQ Four Element Calibration Beads (passport P13H2302). Archsinh transformation with cofactor = 5, debarcoding and matrix-based compensation were performed in the R environment using the CATALYST workflow<sup>3</sup>. Events were then manually pre-gated and gate-cleaned in FlowJo with the exclusion of neutrophils, as shown in (**Technical Supplementary Figure 6**). In all conducted experiments, we acquired 13,311,287 cells derived from 154 samples, which were subsequently analyzed. Marker expression between the two cohorts was normalized to the 95th percentile using the batch adjust function in R<sup>4</sup>. The resulting data set showed no significant difference between the two cohorts (**Technical Supplementary Figure 7**).

Subsequent algorithm-based high-dimensional analysis was performed on the High Performance Computing cluster of the Berlin Institute of Health (BIH) using the CATALYST (v 1.18.1) workflow in R<sup>3</sup>. Briefly, all events were clustered according to lineage marker expression using the *FlowSOM/ConsensusPlus* algorithm<sup>5</sup>. The number of meta-clusters was set at 30 to provide optimal granularity across the dataset. Clusters were then manually merged and annotated according to lineage marker expression and distribution of FlowSOM codes. Eighty percent of total CD45<sup>+</sup> cells were neutrophils (CD66b<sup>+</sup>Siglec8<sup>-</sup>) that did not express integrin  $\beta$ 7 (Extended Data Fig. 3f). To mitigate abundance bias, we excluded them from subsequent analyses. For visualization, dimensionality reduction was computed using the UMAP algorithm with 100 cells per sample. Cluster frequencies and marker expression were then exported for downstream analysis in R. Additionally, mass cytometry data were analyzed using FlowJo v10.4 (FlowJo).

**Single-cell sequencing of PBMCs.** Following the manufacturer's instructions, single-cell RNA-seq libraries were generated using the Chromium Next GEM Single Cell 5' Reagent Kits v2 from 10x Genomics (Pleasanton, CA; CG000330 Rev D). Briefly, a droplet emulsion was generated in a microfluidic chip, followed by barcoded cDNA and surface tag generation within the droplets. Purified and amplified cDNA and surface tags were then subjected to NGS library construction. The manufacturer's instructions generated T and B cell receptor sequencing libraries from amplified cDNA. Sequencing was performed on a NovaSeq 6000 instrument (Illumina). Read depth was targeted at 30K reads per cell for GEX and ~8K for surface and TCR-BCR libraries. Clinical information regarding the patients included in the CITE-seq analysis can be found in **Suppl. Table 13**.

**scRNA-seq data analysis.** Sequencing libraries for gene expression and TCR/BCR were processed together using Cell Ranger multi (v5.0.0) and the GRCh38 genome annotation and analyzed using Seurat v4.0.11<sup>6</sup>. We next used Seurat's reference mapping workflow to jointly transfer cell-type labels at different granularity ("levels") and embedding coordinates from a PBMC reference<sup>6</sup> after filtering out cells with more than 10% mitochondrial gene content, less than 250 or more than 5000 genes, those with a level 1 cell type prediction score of less than 0.75, and doublets called by DoubletFinder<sup>7</sup>. After QC, we were left with a total of 191,578 cells for the PBMC analysis and 48,374 memory CD4<sup>+</sup> T cells for downstream studies. We used scRepertoire v1.7.4 to process Cell Ranger VDJ output. ADT data were normalized using CLR normalization, and expression thresholds were determined using the Binarize R package (v1.3).

Pooled scRNAseq data for FACS-sorted CD3<sup>+</sup> CD4<sup>+</sup> CD45RA<sup>-</sup> memory T cells from multiple donors were demultiplexed using Seurat's HTODemux function combined with SNP-based demultiplexing using cellSNP-lite<sup>8</sup> over common variants from the 1000 Genomes Project and Vireo<sup>9</sup> and otherwise processed similarly. We then selected all level 2 cell types (CD4 CTL,

173 CD4 proliferating, CD4 TCM, CD4 TEM or Treg) from each dataset and performed a joint  
174 UMAP and clustering (at resolution 0.5) using Seurat's SCTransform and IntegrateData  
175 workflows<sup>10,11</sup>. Finally, two smaller MKI67<sup>+</sup> clusters were merged. Differential expression  
176 between cells in cluster 1 cells and others was performed using logistic regression, with sample  
177 type (PBMC or CD4<sup>+</sup> sorted) as covariate.

178 **RNA isolation from tissue samples (IBDome cohort).** RNA was isolated from biopsies  
179 taken during routine endoscopy or from resected tissues at the First Department of Medicine,  
180 Friedrich-Alexander Universität Erlangen-Nürnberg (Germany), and at the Department of  
181 Gastroenterology, Infectious Diseases and Rheumatology at the Charité – Universitätsmedizin  
182 Berlin (Germany) by using a single-use biopsy forceps (Olympus). Samples were incubated in  
183 RNA protect reagent (RNAprotect Tissue Reagent, Qiagen) and stored at -80°C. One biopsy  
184 was thawed on ice and homogenized in RLT buffer (Qiagen) employing the TissueLyser LT  
185 (Qiagen) for RNA isolation. RNA was isolated using the RNeasy kit (Qiagen) and RNA Clean  
186 & Concentrator kit (Zymo Research). The concentration was measured at NanoDrop One/One  
187 (Thermo Fisher Scientific) and the quality (RNA integrity number, RIN) at Tape Station  
188 (Agilent). The RNA was used for bulk RNA sequencing at the NGS Competence Center  
189 Tübingen (NCCT).

190 **Bulk RNA-seq analysis.** Paired sequencing reads were processed with the nf-core/rnaseq  
191 pipeline version 3.4<sup>12</sup>. In brief, reads were aligned to the GRCh38 reference genome with  
192 GENCODE v33 annotation using STAR<sup>13</sup>. Transcripts per million (TPM) were quantified  
193 using Salmon<sup>14</sup> and transformed to log<sub>10</sub>(TPM+1) for visualization in R.

194 **Proteomics assay.** Proteomics analysis of patient serum samples was performed using the  
195 Olink "Target 96 Inflammation panel" platform ([https://www.olink.com/products/](https://www.olink.com/products/inflammation/)  
196 inflammation/). Frozen patient sera were randomly distributed on plates and spiked-in with  
197 controls according to the manufacturer's instructions (Olink, Uppsala, Sweden). Clinical

information regarding the patients included in the CITE-seq analysis can be found in **Suppl. Table 13**. The panel is a high-throughput, multiplex immunoassay that allows the simultaneous analysis of 92 inflammation-related protein biomarkers in 88 samples using Proximity Extension Assay (PEA) technology. In brief, each protein is bound by a matched pair of antibodies coupled to unique and partially complementary oligonucleotides. It hybridizes when the DNA coupled to the two antibodies is brought into close proximity. Only the tags that hybridize correctly are extended into an amplicon by quantitative real-time PCR, with a unique sequence for each protein. This dual antibody binding requirement and DNA barcoding provide exceptional readout specificity. The software then reports the relative protein concentrations of the proteins. The proteomics data is published in a normalized protein expression (NPX) format, Olink's arbitrary unit, on the log<sub>2</sub> scale. NPX data allows users to identify changes in individual protein levels throughout the sample and to identify protein signatures. The higher the NPX value, the higher the protein concentration. As NPX is in a log<sub>2</sub> scale, a 1 NPX difference means a doubling in protein concentration. NPX values that did not pass internal quality control were excluded from further analysis. Downstream data analysis was performed in R using the packages limma, ggplot2 and ComplexHeatmap. The assays were done in Olink Proteomics AB, Uppsala, Sweden. Several serum proteins were different in serum when comparing IBD patients with healthy controls (HC), and were increased after vedolizumab treatment (**Supplementary Fig. 9**).

**Reanalysis of microarray data:** Expression data for GSE73661 was obtained using the GEOquery R package (v2.58.0) and averaged over all probes per gene. Differential expression was performed using limma (v3.46.0). A heatmap was created for Olink panel genes using ComplexHeatmap (v2.6.2), using only control samples or patients treated with Veduzimab at 0w or 6w and with annotation about their responder status.

**Machine learning - prediction of treatment efficacy.** We aimed to identify markers across four different data modalities for predicting vedolizumab treatment-efficacy in patients with IBD. We only considered patient data with clear response outcomes, removing patients in remission. The exact number of patients used for each run can be found in **Suppl. Table 15**. Using logistic regression, we performed a 2-fold cross-validation approach for each classification run, which was repeated five times. Performance was evaluated using the mean area under the curve (AUC) and the standard error of the mean (SEM). Empirical p-values were calculated via permutation tests ( $n=1000$ ), with AUC as the final evaluation metric. After training the classifier on the full data sets, including all parameters, we extracted the feature coefficients from the classification models. These coefficients were then used to rank and select the top 10 most influential features. All AUC curves were generated using the *matplotlib* package (version 3.5.1), bar plots and dot plots were generated using the *ggplot2* R package (version 3.3.3)

**Machine learning - prediction of treatment response.** As mentioned above, we aimed to identify markers across four different data modalities, and the analysis was performed on the same curated patient cohort (**Suppl. Table S15**). Here, we employed a set of machine-learning techniques, including both linear (logistic regression, linear support vector machines) and non-linear methods (radial SVM, polynomial SVM, random forest) and functionalities embedded in the *scikit-learn*-package (version 1.0.2). Logistic regression was selected as the preferred method due to its robust performance and model simplicity (cf. Section Results). After a 2-fold cross-validation run with 5 repetitions as above, we extracted the top 10% features based on feature coefficients, and a new classification model was trained using these top 10% selected features. As a control, we performed the same analysis on a reduced patient set including only completely profiled patients, so that the number of patients was unchanged across all models. Next, the four models with the highest predictive capacity were identified based on the AUC

value, and they were validated using iterative model validation (see below). From the validated models, we selected the five features with the highest absolute feature coefficients, in line with the square-root-of-n criterion with respect to the number of patients postulated by Hua et al.<sup>15</sup>. Using the same 2-fold cross validation scheme with 5 repetitions, we also employed two automatic feature-regularizers, Lasso and Elasticnet, on the validated models. (**Figure S12C, Supplementary Table 5**). The predictive capacity of selected marker combinations was analysed on the validated models, again using the reduced patient set of completely profiled patients as control. The three most frequent FACS markers were further assessed on an external validation cohort.

**Machine-learning - iterative model validation.** To test the four best predictive models and identify a robust feature panel, we implemented a repeated hold-out validation strategy<sup>16</sup>. We repeatedly held out 20% of the patients per model (3-5 patients), while maintaining the observed ratio of responders to non-responders. For each of those iterations (50 in total), we trained and validated the model on the remaining 80% of the patients as described above. Specifically, within each iteration, we defined a feature panel by extracting the top 10% of features based on the absolute feature coefficients. Next, we assessed how often the top  $n$  features in that set were included in the top 10%, 20%, ..., 100% of feature panels derived in all other iterations. We calculated the robustness of the randomly selected validation set in each iteration in terms of the AUC across the obtained inclusion rates, and the set with the highest robustness score was selected for further analysis (see **Technical Supplementary Figure 9**). Finally, the models were tested using the held-out patient set to assess the predictive capacity on unseen data.

**Code availability.** All original code necessary to replicate our analyses, including scRNAseq analyses and machine learning code, are deposited in a GitHub repository ([https://github.com/VeroHo/vedo\\_paper](https://github.com/VeroHo/vedo_paper)).

272 **Data availability.** All source data relevant to understanding and reproducing the results  
273 presented in this paper are provided in the supplementary data. Sequencing data (scRNAseq,  
274 CITEseq and TCR sequencing) can be accessed at the following GEO accession number:  
275 GSE261334.

276 **Statistical analysis.** Statistical analyses and visualizations were created in R using ggpubr (v  
277 0.6.0), ggplot2 (v 3.4.3), limma (v 3.50.3), corrplot (v 0.92), and ComplexHeatmap (v 2.10.0)  
278 packages or with the Prism Software (GraphPad Software). *P* values were calculated using the  
279 Wilcoxon test. Paired analyses are indicated by connecting lines.

## SUPPLEMENTARY REFERENCES

1. Soler D, Chapman T, Yang L-L, et al. The Binding Specificity and Selective Antagonism of Vedolizumab, an Anti- $\alpha 4\beta 7$  Integrin Therapeutic Antibody in Development for Inflammatory Bowel Diseases. *J Pharmacol Exp Ther* 2009;330:864–875.
2. **Burns M, Ostendorf L**, Biesen R, et al. Dysregulated CD38 Expression on Peripheral Blood Immune Cell Subsets in SLE. *Int J Mol Sci* 2021;22:2424.
3. Nowicka M, Krieg C, Crowell HL, et al. CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. *F1000Research* 2019;6:748.
4. **Schuyler RP, Jackson C**, Garcia-Perez JE, et al. Minimizing Batch Effects in Mass Cytometry Data. *Front Immunol* 2019;10:2367.
5. Gassen SV, Callebaut B, Helden MJV, et al. FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytom Part A* 2015;87:636–645.
6. **Hao Y, Hao S**, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell* 2021;184:3573–3587.e29.
7. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* 2019;8:329–337.e4.
8. Huang X, Huang Y. Cellsnp-lite: an efficient tool for genotyping single cells. *Bioinformatics* 2021;37:4569–4571.
9. Huang Y, McCarthy DJ, Stegle O. Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference. *Genome Biol* 2019;20:273.
10. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* 2019;20:296.
11. **Stuart T, Butler A**, Hoffman P, et al. Comprehensive Integration of Single-Cell Data. *Cell* 2019;177:1888–1902.e21.

305 12. **Ewels PA, Peltzer A**, Fillinger S, et al. The nf-core framework for community-curated  
306 bioinformatics pipelines. *Nat Biotechnol* 2020;38:276–278.

307 13. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner.  
308 *Bioinformatics* 2013;29:15–21.

309 14. Patro R, Duggal G, Love MI, et al. Salmon provides fast and bias-aware quantification of  
310 transcript expression. *Nat Methods* 2017;14:417–419.

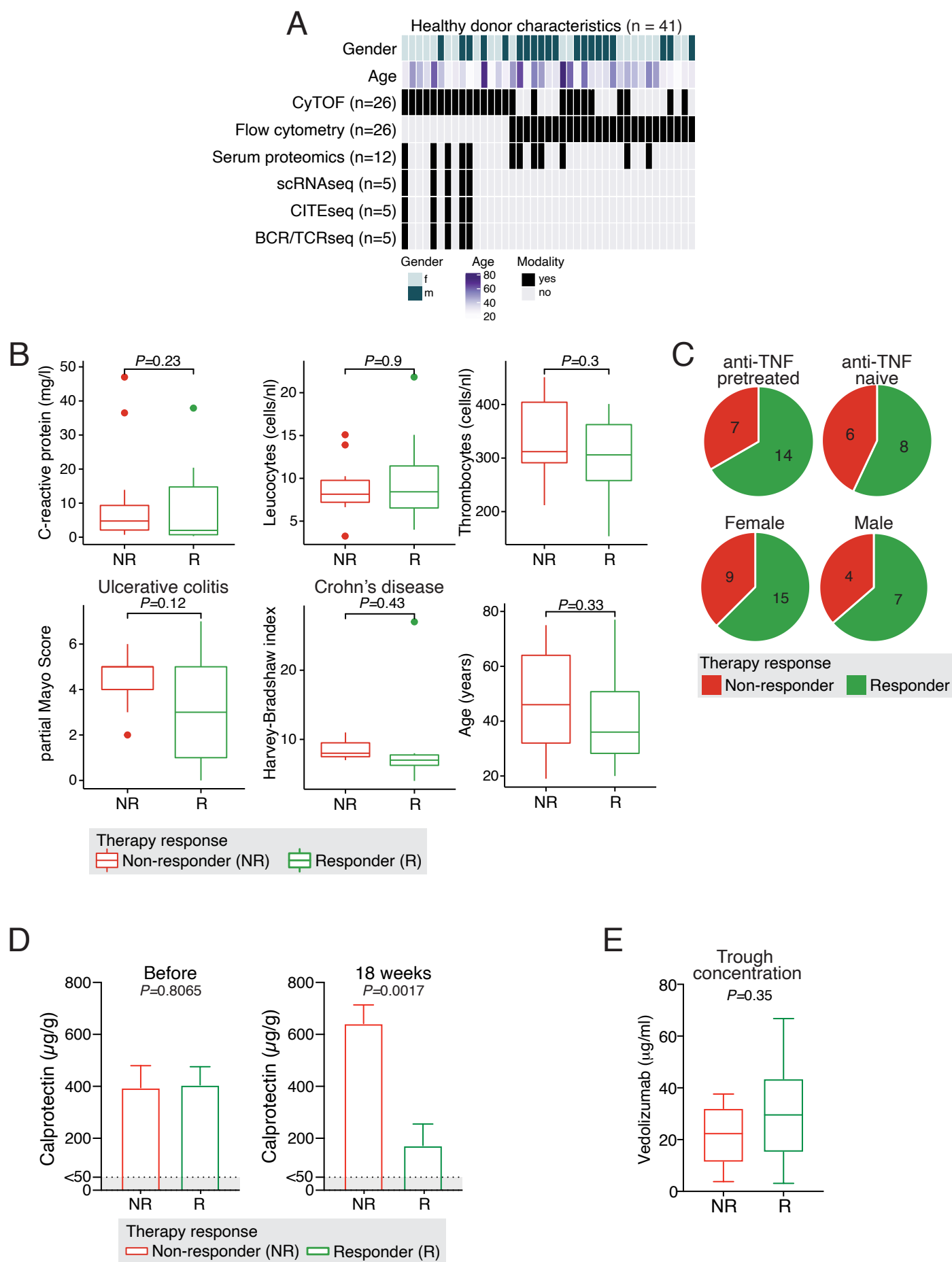
311 15. Hua J, Xiong Z, Lowey J, et al. Optimal number of features as a function of sample size  
312 for various classification rules. *Bioinformatics* 2004;21:1509–1515.

313 16. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable  
314 prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and  
315 Elaboration. *Ann Intern Med* 2015;162:W1–W73.

316

317 **EXTENDED DATA FIGURES 1-16 AND TECHNICAL SUPPLEMENTARY**  
318 **FIGURES 1-10**





Supplementary Figure 1

320 **Supplementary Figure 1. Characteristics of healthy donors and (para-)clinical data of**  
321 **IBD patients.**

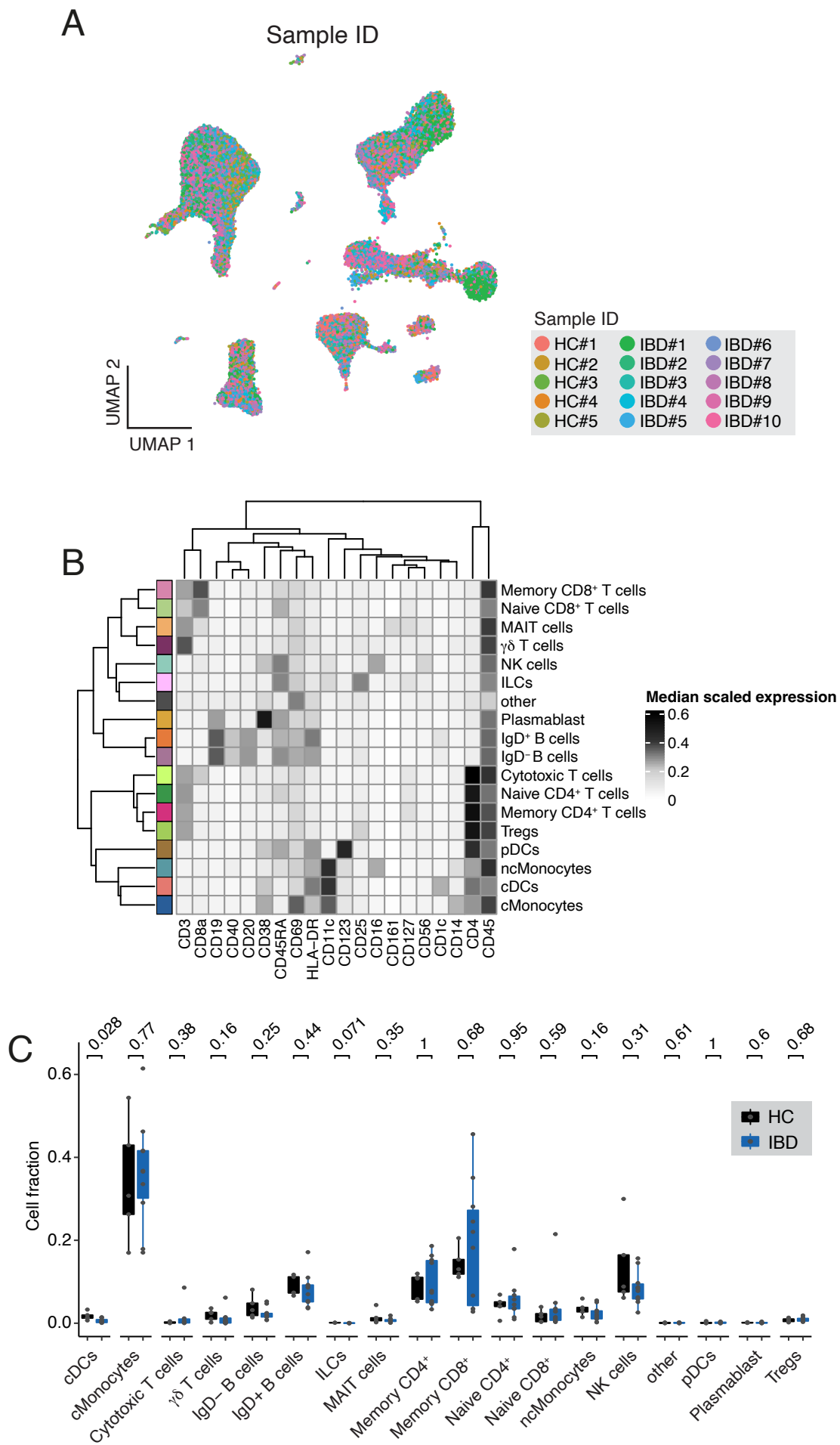
322 (A) Graph depicting the age and gender distribution within the healthy control cohort (n=41),  
323 along with the techniques applied to each sample.

324 (B) Box plots illustrating the comparison of (para-)clinical parameters between non-responders  
325 and responders prior to therapy (responders n=22, non-responders n=13).

326 (C) Pie charts illustrating the indicated characteristics in IBD patients before therapy.

327 (D) Bar graph illustrating stool calprotectin levels before initiation of treatment and  
328 approximately 18 weeks after treatment commencement.

329 (E) Trough concentrations in IBD patients receiving vedolizumab at week 6 of treatment. (b,  
330 d) Mann-Whitney test.



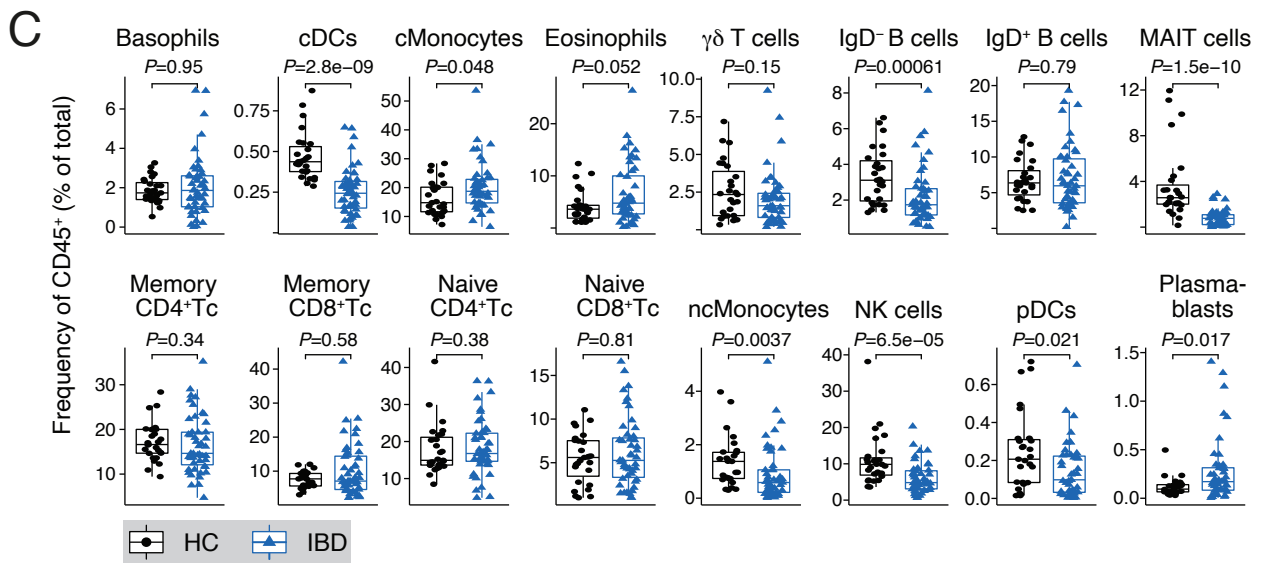
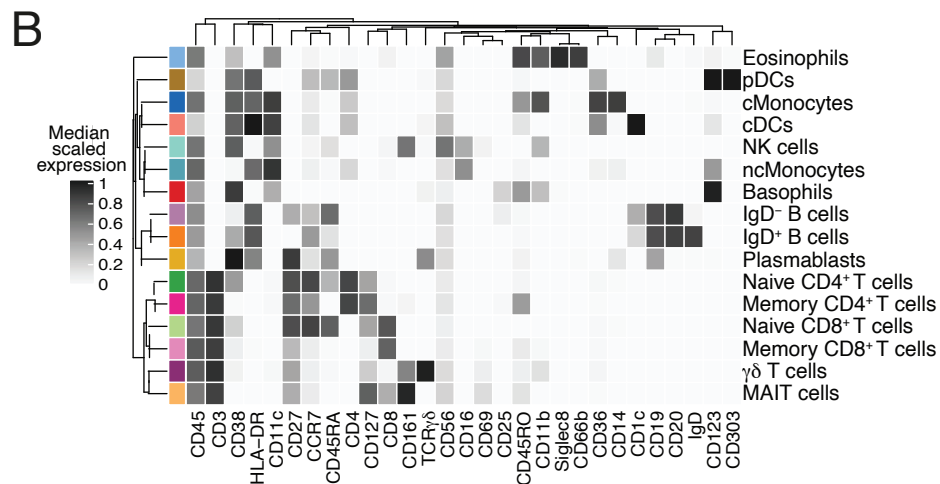
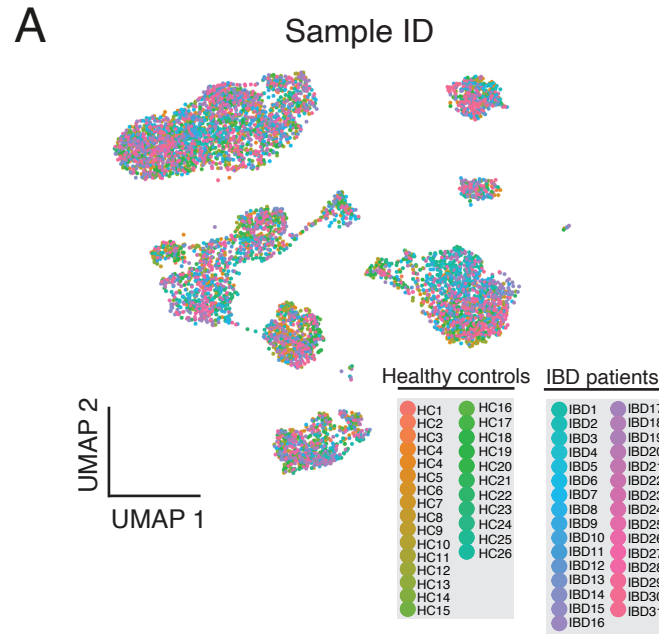
Supplementary Figure 2

**Supplementary Figure 2. Immune cell subset identification and quantification in peripheral blood of healthy controls and IBD patients using CITE-seq.**

(A) UMAP plots depicting CITE-seq profiles of peripheral blood cells, with color-coded differentiation indicating distinct donors (n=191,578 cells derived from a total of 25 samples: HC, n=5; IBD, n=10 matched before and after treatment). Sorted CD45<sup>+</sup> cells from the peripheral blood of 5 healthy controls and 10 ulcerative colitis patients before and 6 weeks after vedolizumab (VDZ) treatment initiation were subjected to CITE-seq analysis.

(B) Heatmap of surface antigen expression on identified cell types in peripheral blood cells derived from CITE-seq data.

(C) Box plot of cell type proportions comparing healthy controls and IBD patients before treatment. Mann Whitney test.



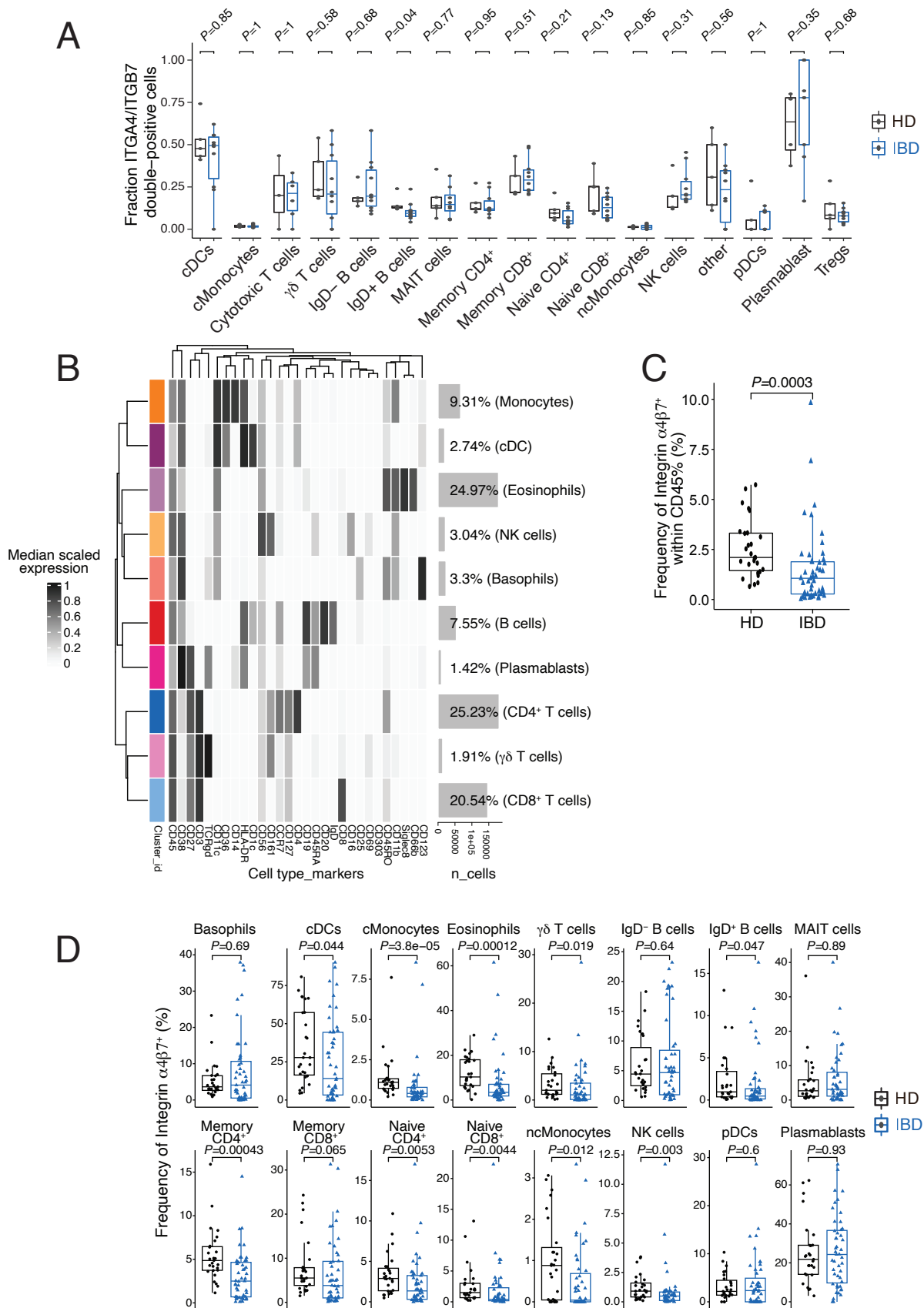
Supplementary Figure 3

**Supplementary Figure 3. Immune cell subset identification and quantification in peripheral blood of healthy controls and IBD patients using mass cytometry.**

(A) UMAP plots of peripheral blood cells extracted from the CyTOF dataset, with color representation based on patient IDs. Results derived from FlowSOM/ConsensusPlus clustering analysis performed on 13,311,287 cells from 154 samples. This dataset includes both healthy controls and IBD patients both before and during treatment.

(B) Median scaled lineage marker expression in the indicated cluster resulting from FlowSOM clustering.

(C) Comparison of cluster frequencies between healthy control (n=27) and IBD patients (n=31) before therapy. Mann Whitney test.



Supplementary Figure 4

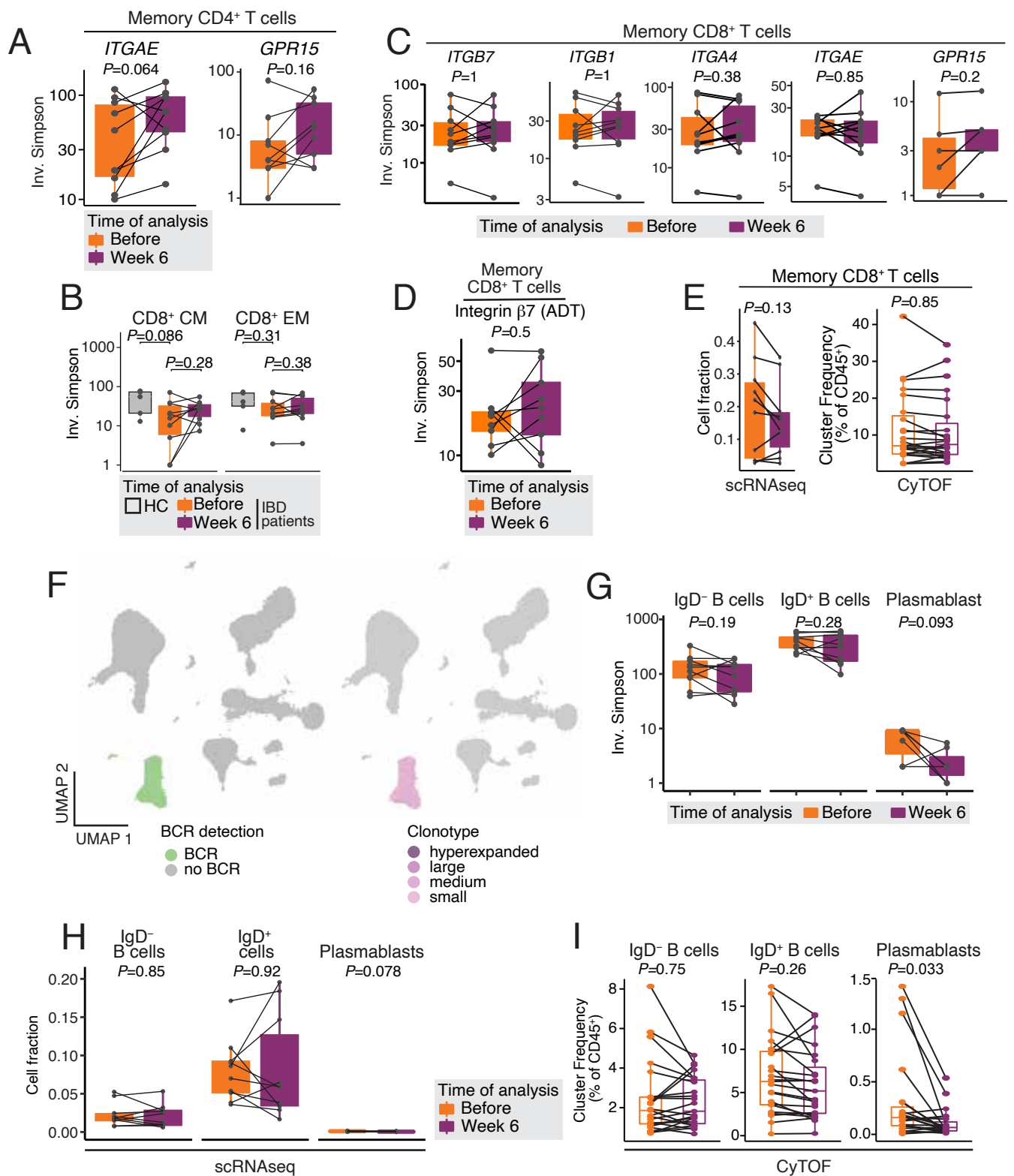
**Supplementary Figure 4. Expression and distribution of integrin  $\alpha 4\beta 7$  within CD45<sup>+</sup> cells in scRNA-seq and mass cytometry data.**

(A) Box plot of the proportion of *ITGA4/ITGB7* double-positive cells in scRNA-seq data, comparing healthy controls and IBD patients. Mann-Whitney test.

(B) Median scaled lineage marker expression in the indicated cluster resulting from FlowSOM clustering of  $\alpha 4\beta 7^+$  cells in the mass cytometry dataset related to Figure 2C.

(C) Comparison of integrin  $\alpha 4\beta 7$  expression within CD45<sup>+</sup> cells in peripheral blood between healthy control (n=27) and IBD patients (n=31) before therapy. Mann Whitney test.

(D) Comparison of integrin  $\alpha 4\beta 7$  expression within the indicated clusters between healthy control (n=27) and IBD patients (n=31) before therapy. Mann Whitney test.



Supplementary Figure 5

**Supplementary Figure 5. Vedolizumab treatment does not affect the diversity of circulating CD8 T or B cells.**

(A) Clonal diversity in memory CD4<sup>+</sup> T cell subpopulations defined by expression of the indicated marker genes.

(B) Clonal diversity quantified by the inverse Simpson index in central and effector memory CD8<sup>+</sup> T cell subpopulations.

(C) Box plot of TCR repertoire diversity in memory CD8<sup>+</sup> T cells positive for the indicated markers.

(D) Box plot of TCR repertoire diversity in memory CD8<sup>+</sup> T cells positive for integrin  $\beta$ 7 (ADT) determined by CITE-seq.

(E) Box plot depicting the abundance of CD8<sup>+</sup> memory T cells before and 6 weeks after vedolizumab treatment using both scRNA-seq data (left) and CyTOF data (right).

(F) UMAP plots of scRNA-seq data showing cells with detected BCR sequence (left) and expansion status of the associated clonotype (right; small:  $< 10^{-3}$ , medium:  $< 10^{-2}$ , large:  $< 10^{-1}$ , hyperexpanded:  $> 10^{-1}$ ).

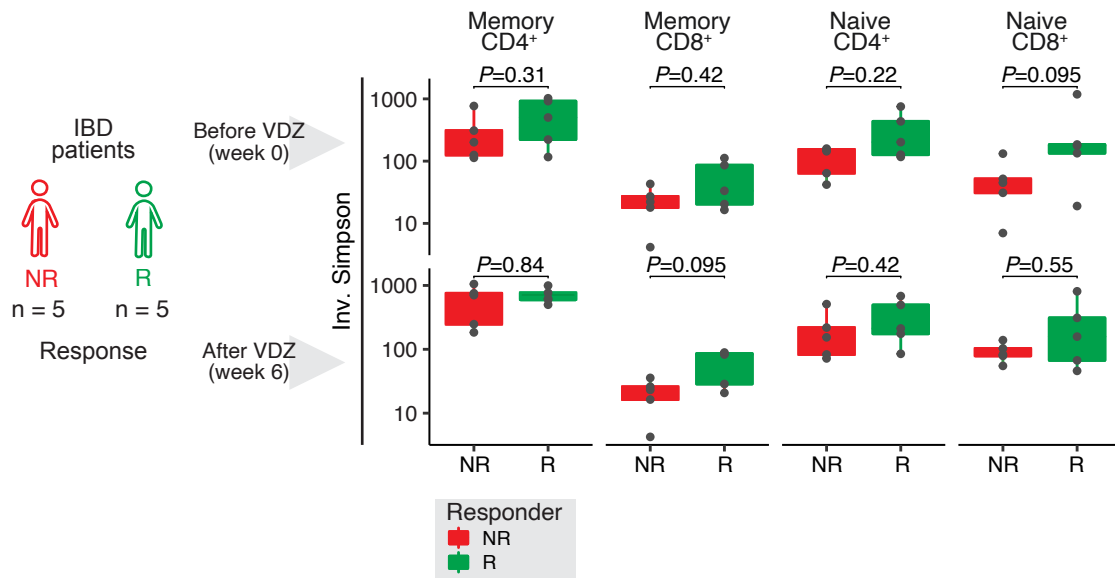
(G) Box plot of BCR repertoire diversity quantified by the inverse Simpson index in the indicated B cell subpopulations before and 6 weeks after starting vedolizumab treatment.

(H) Box plot of the proportions of the indicated B cell types before and after vedolizumab treatment.

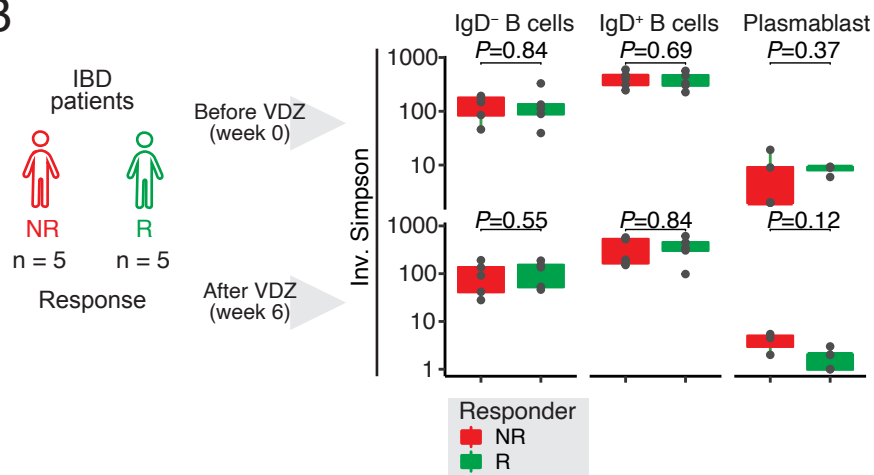
(I) Box plot depicting the abundance of B cell subtypes before and 6 weeks after vedolizumab treatment using CyTOF data.

(B-E, G-I) Wilcoxon test.

A



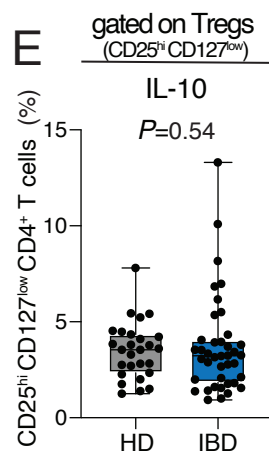
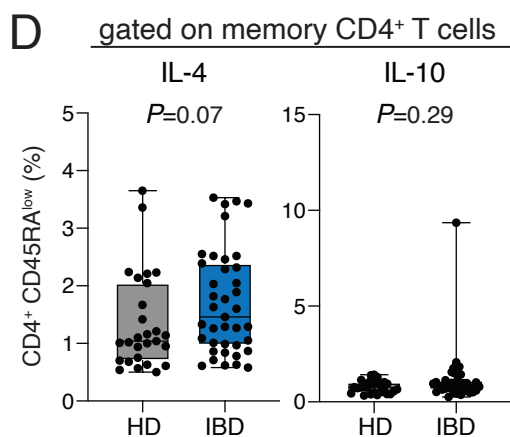
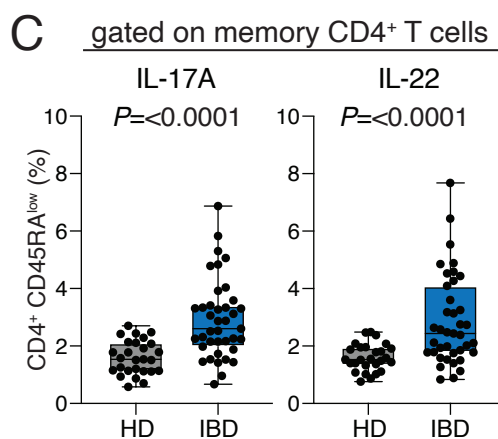
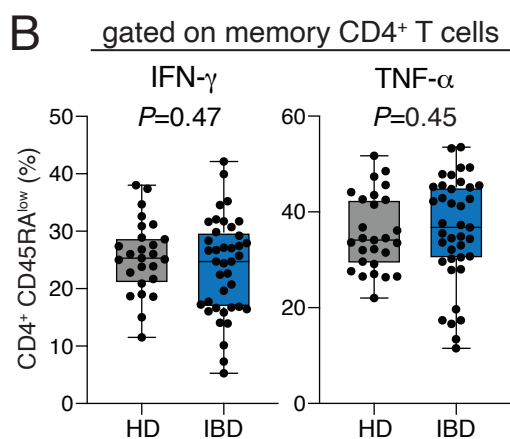
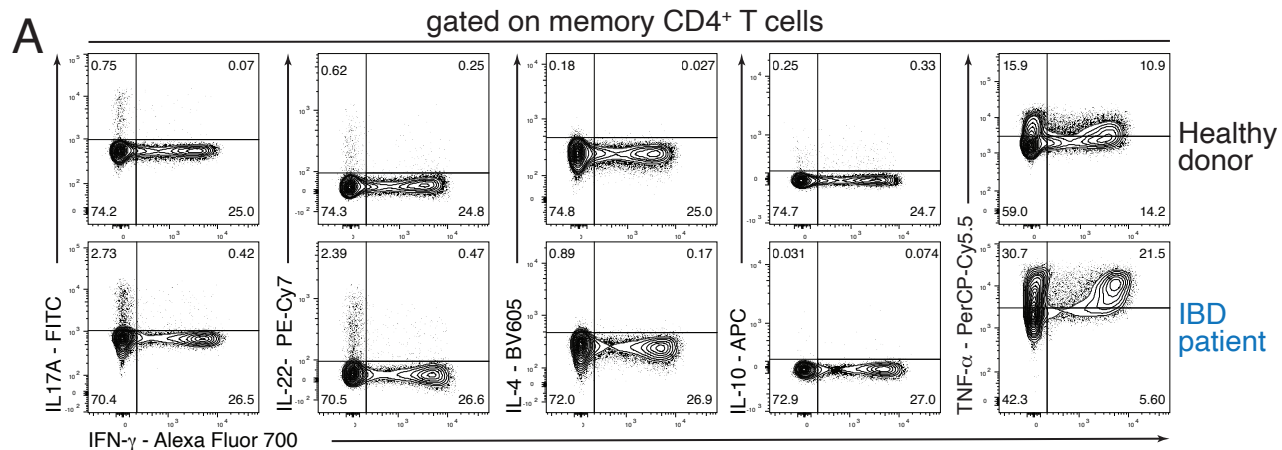
B



**Supplementary Figure 6. The alterations in circulating T and B cell diversity before and after vedolizumab don't predict therapy outcome**

(A) Box plot of TCR repertoire diversity (measured by the inverse Simpson score) in different T cell subsets before treatment and 6 weeks after treatment in Vedolizumab responders (R) and non-responders (NR). Paired Wilcoxon test.

(B) Box plot of BCR repertoire diversity (measured by the inverse Simpson score) in the indicated B cell subpopulations before and 6 weeks after Vedolizumab treatment in Vedolizumab responders (R) and non-responders (NR). Paired Wilcoxon test.



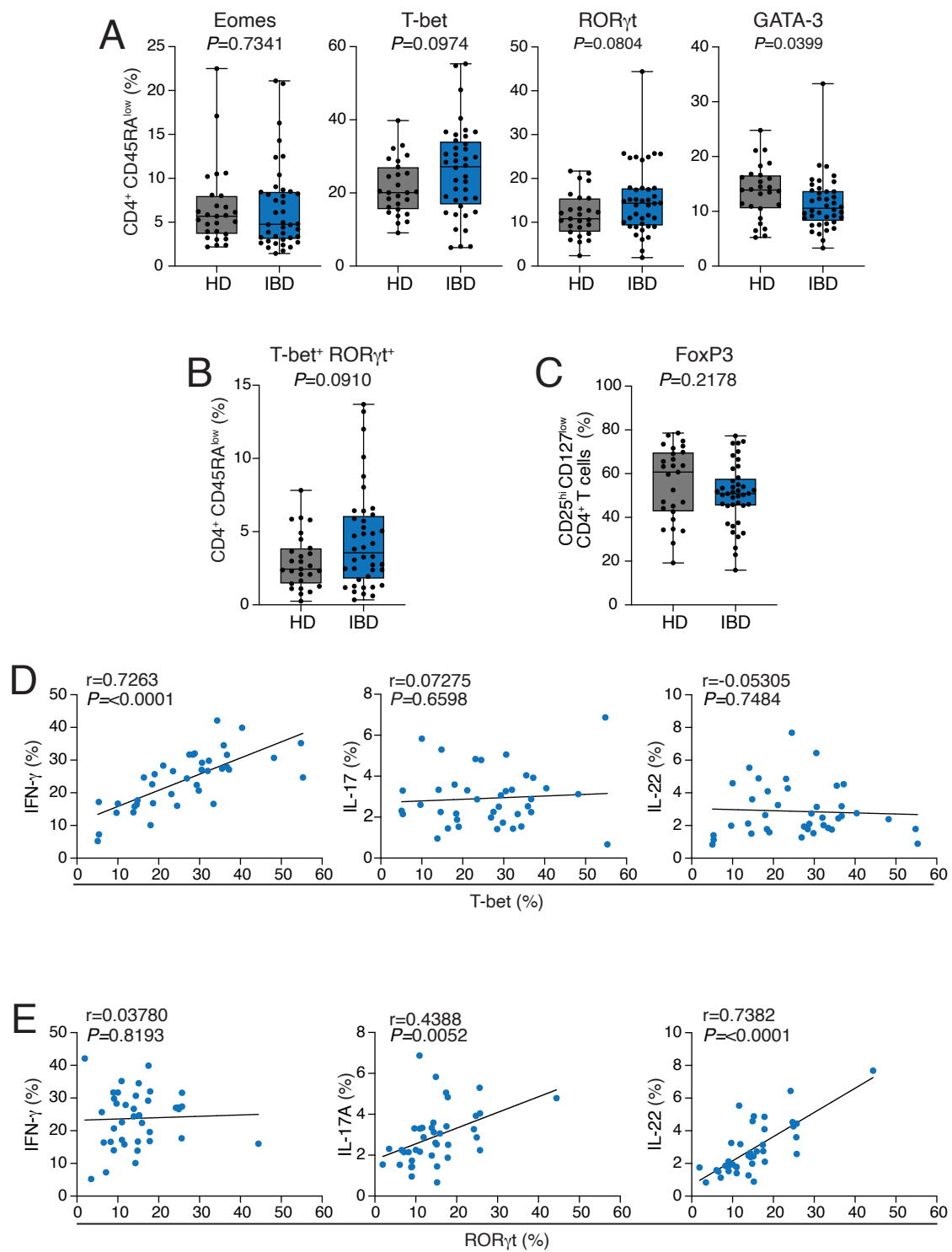
Supplementary Figure 7

392 **Supplementary Figure 7. Cytokine expression in circulating CD4<sup>+</sup> T cells in healthy**  
393 **donors and IBD patients before Vedolizumab treatment.**

394 (A) Dot plots of intracellular cytokine staining of memory CD4<sup>+</sup> T cells from healthy donors  
395 and IBD patients before Vedolizumab treatment.

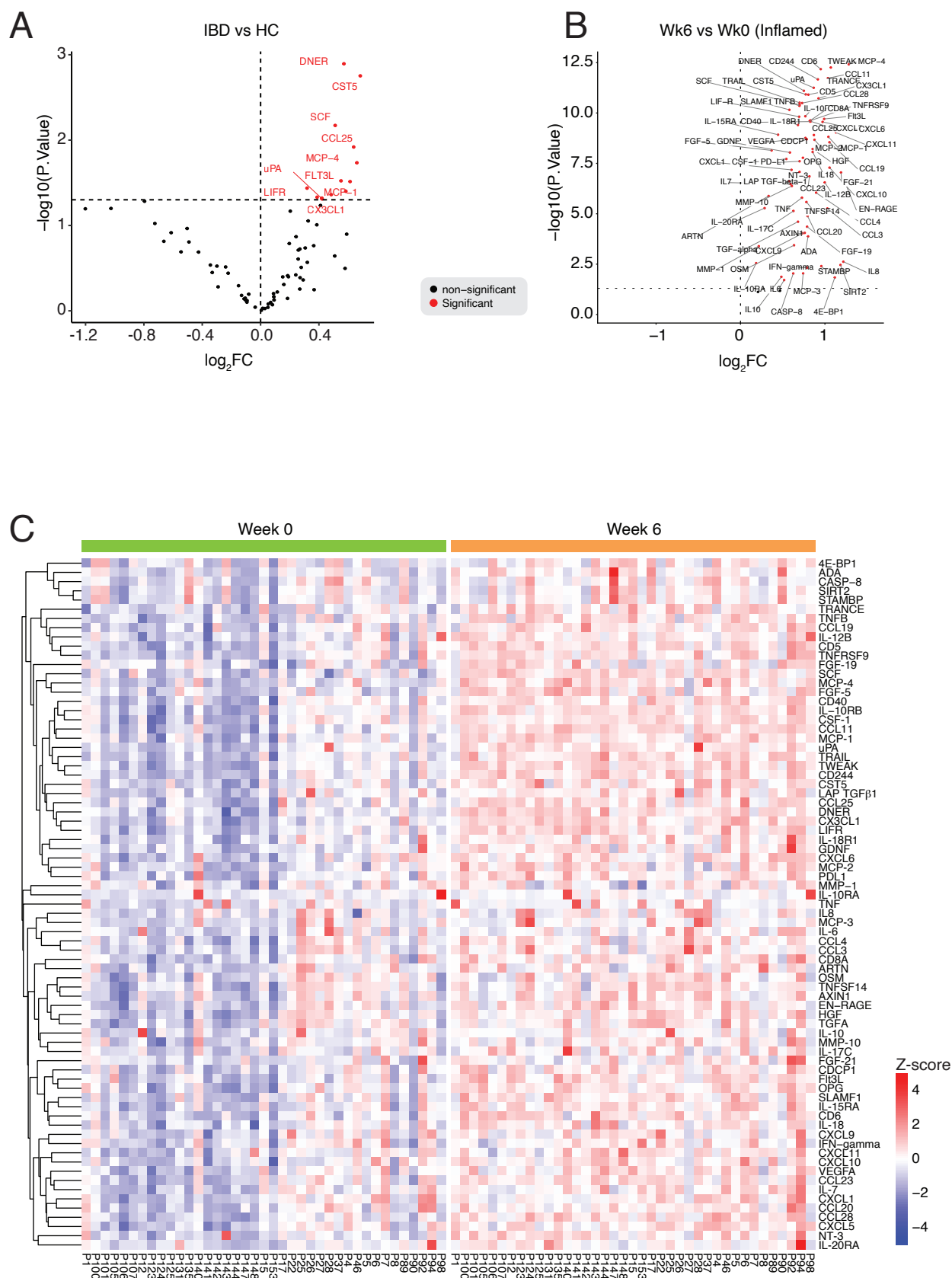
396 (B-D) Box plots showing the percentage of the indicated cytokines within memory CD4<sup>+</sup> T  
397 cells.

398 (E) Box plot showing the percentage of IL-10 expression within regulatory CD25<sup>high</sup> CD127<sup>low</sup>  
399 CD4<sup>+</sup> T cells.



Supplementary Figure 8

400 **Supplementary Figure 8. Transcription factor expression in circulating CD4<sup>+</sup> T cells in**  
401 **healthy donors and IBD patients before Vedolizumab treatment.**  
402 (A-B) Box plots showing the percentage of the indicated transcription factors within memory  
403 CD4<sup>+</sup> T cells.  
404 (C) Box plot showing the percentage of FoxP3 expression within regulatory CD25<sup>high</sup>  
405 CD127<sup>low</sup> CD4<sup>+</sup> T cells.  
406 (D, E) Spearman correlation between T-bet or RORγt expression and the expression of  
407 indicated cytokines within memory CD4<sup>+</sup> T cells.



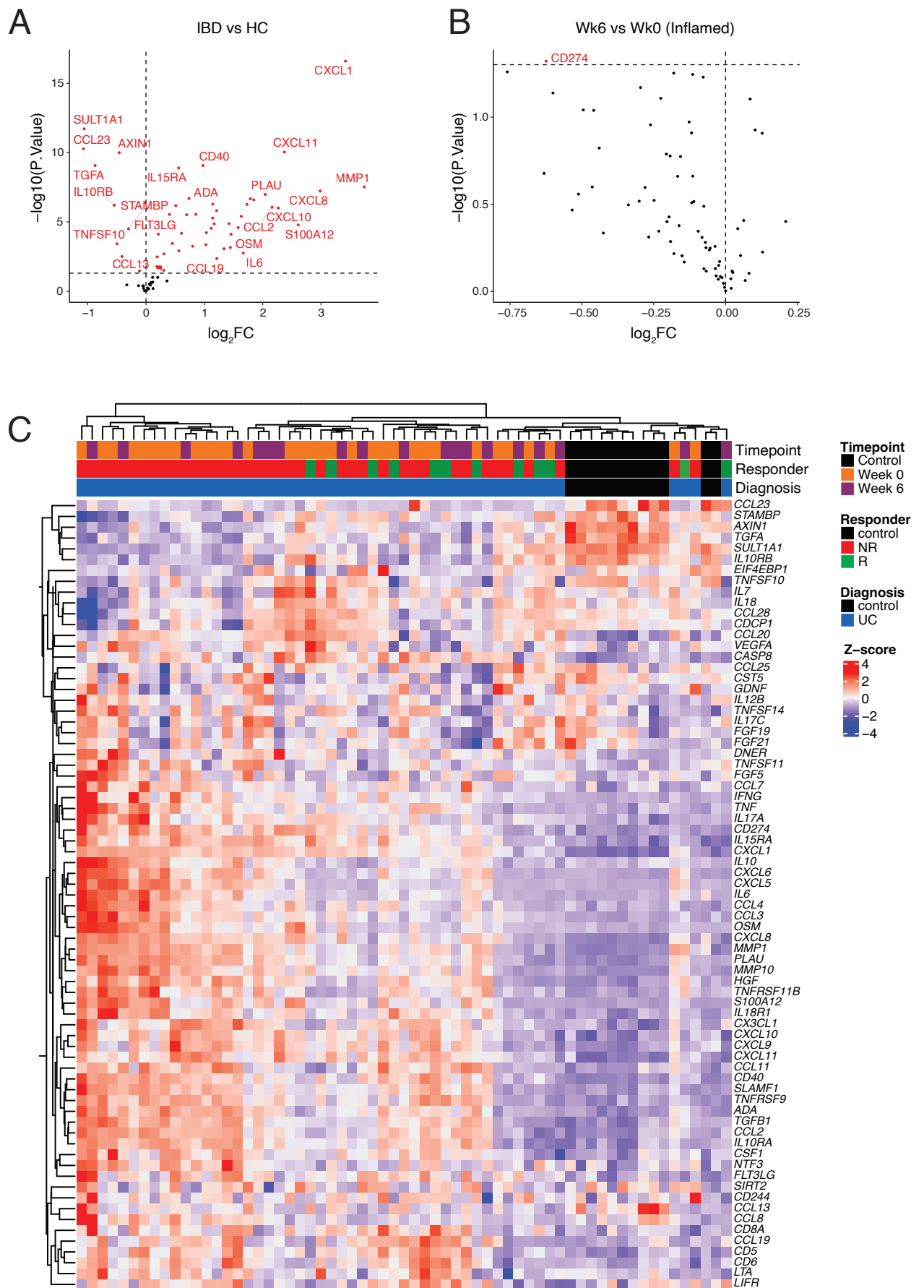
Supplementary Figure 9

408 **Supplementary Figure 9. Serum proteomics from healthy controls and IBD patients.**

409 (A) Volcano plot showing differential expression of serum protein concentrations (NPX  
410 values) between HD vs. IBD patients before therapy.

411 (B) Volcano plot showing differential expression of serum protein concentrations in IBD  
412 patients before and 6 weeks after treatment with vedolizumab.

413 (C) Heatmap of serum proteins in IBD patients before (week 0) and after treatment (week 6).



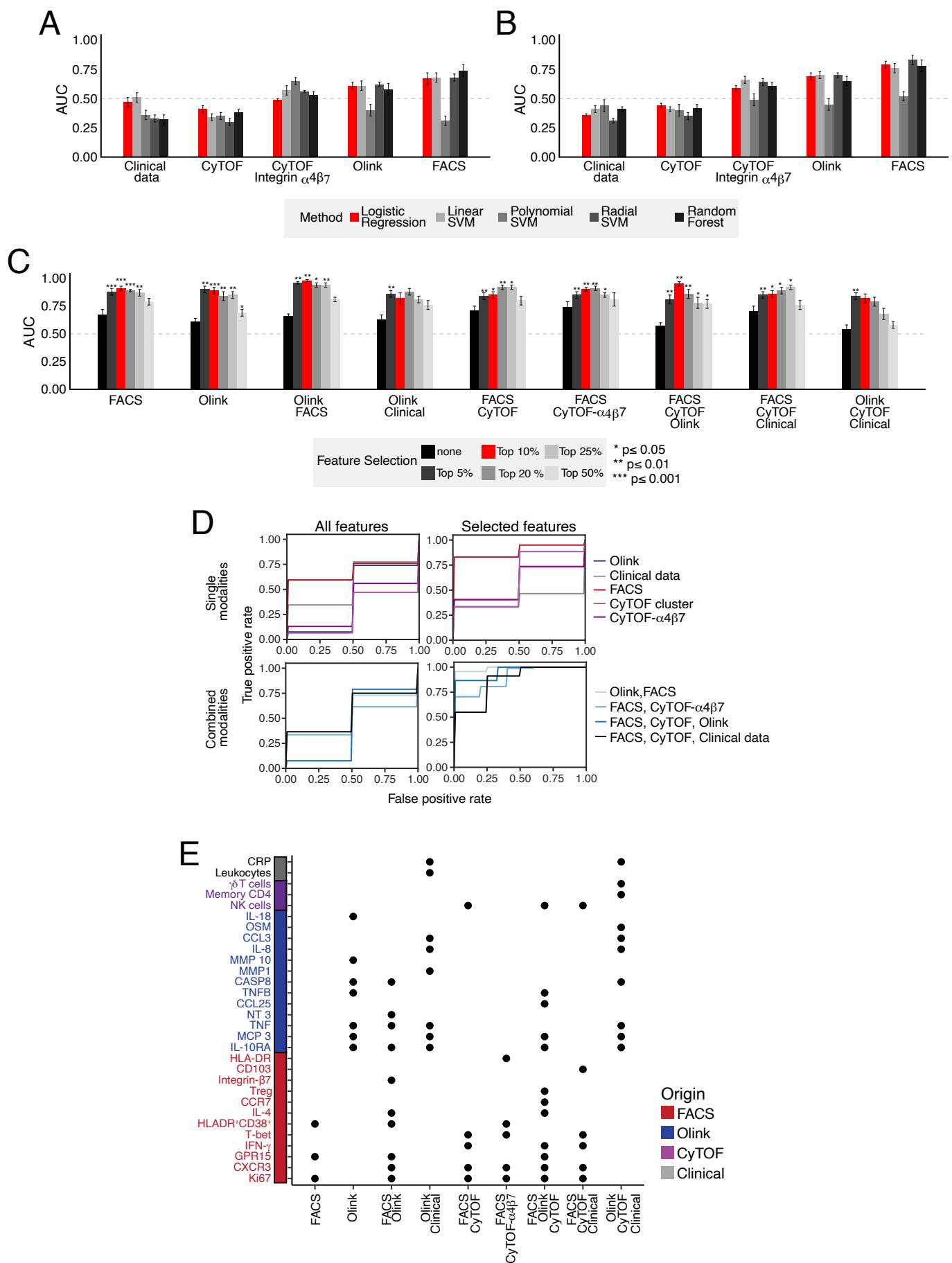
Supplementary Figure 10

**Supplementary Figure 10. Expression of inflammatory genes in mucosal tissue before and after vedolizumab treatment.**

(A) Volcano plot displaying differential gene expression of serum protein-associated genes from the Target 96 Inflammation Olink panel in mucosal samples before therapy induction, comparing controls (HC) to IBD patients. The data is derived from the RNA-seq data from the GEMINI trial (GSE73661).

(B) Volcano plot illustrating differential gene expression of genes corresponding to serum proteins analyzed by the Olink inflammatory panel. This comparison is made between mucosal samples collected from IBD patients before and 6 weeks after vedolizumab treatment. The data is derived from GSE73661.

(C) Heatmap of genes corresponding to serum proteins analyzed by the Target 96 Inflammation Olink panel in HC and IBD patients before (week 0) and after treatment (week 6).



Supplementary Figure 11

**Supplementary Figure 11. Data integration and optimization of classification tools.**

(A) Comparison of five different classification tools within individual data sets. Shown are AUC values with SEM.

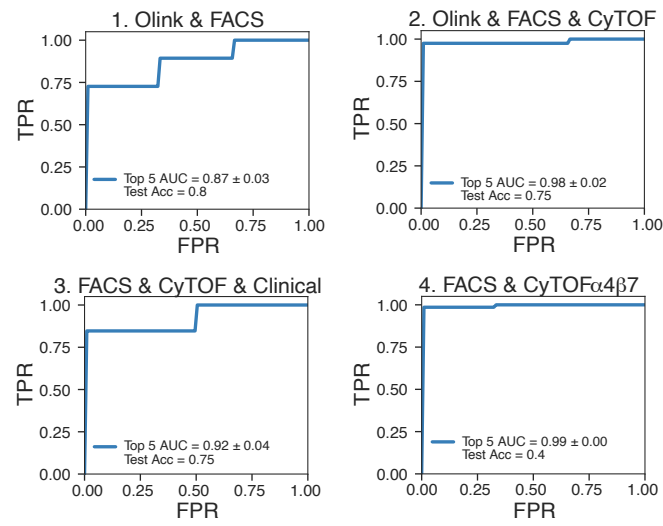
(B) Comparison of five classification tools using selected features. Shown are AUC values with SEM. Feature coefficients were determined using logistic regression, and only those in the top 50% were retained. The color code is the same as in panel A. Due to similar mean AUC scores across data sets and the simplicity of the model, logistic regression (red bars) was selected as the preferred classification tool in subsequent analyses.

(C) Predictive capacity before and after application of feature-coefficient thresholds as indicated. Per model, all features were ranked according to their absolute feature coefficients. Models that did not have an AUC of at least 0.5 (average AUC minus SEM) as indicated by the dashed line are not shown. Empirical p-values were calculated using permutation tests (n=1000). Based on the robustness of the obtained AUCs, we used the top 10% cut-off (red bars).

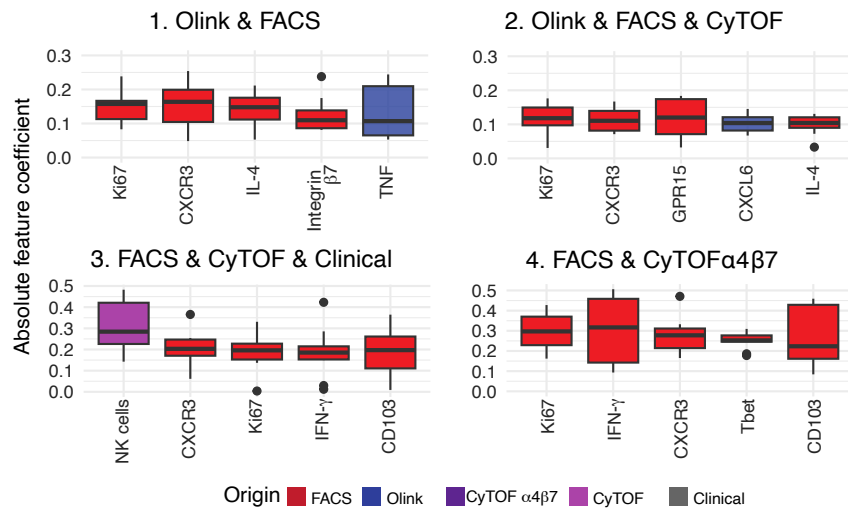
(D) Predictive capacity of models using the reduced set of completely profiled patients (n=14) as control, analogous to Figure 5B.

(E) Overview of the most influential features in the classification models shown in (B). For each data set, a feature is selected if its absolute feature coefficient scored within the top 10% of all features.

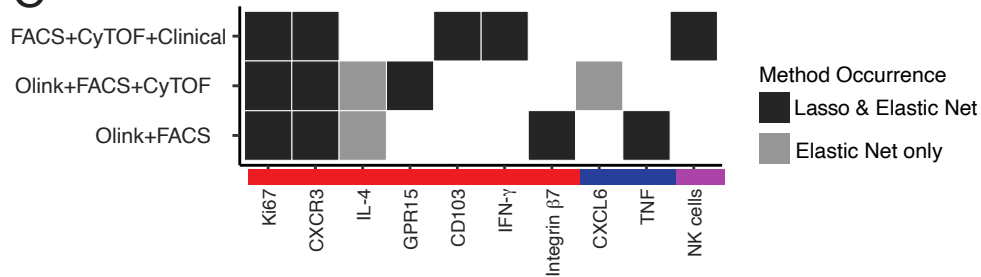
A



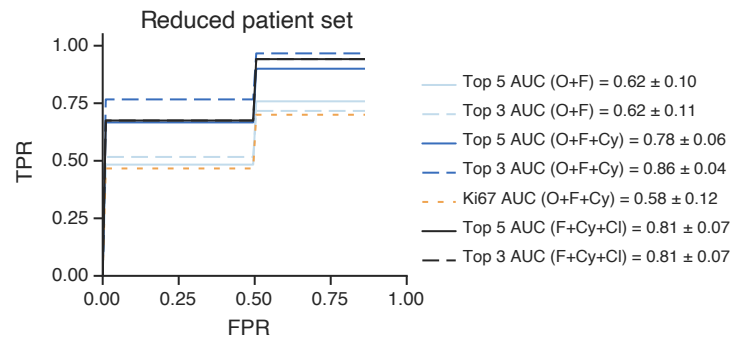
B



C



D



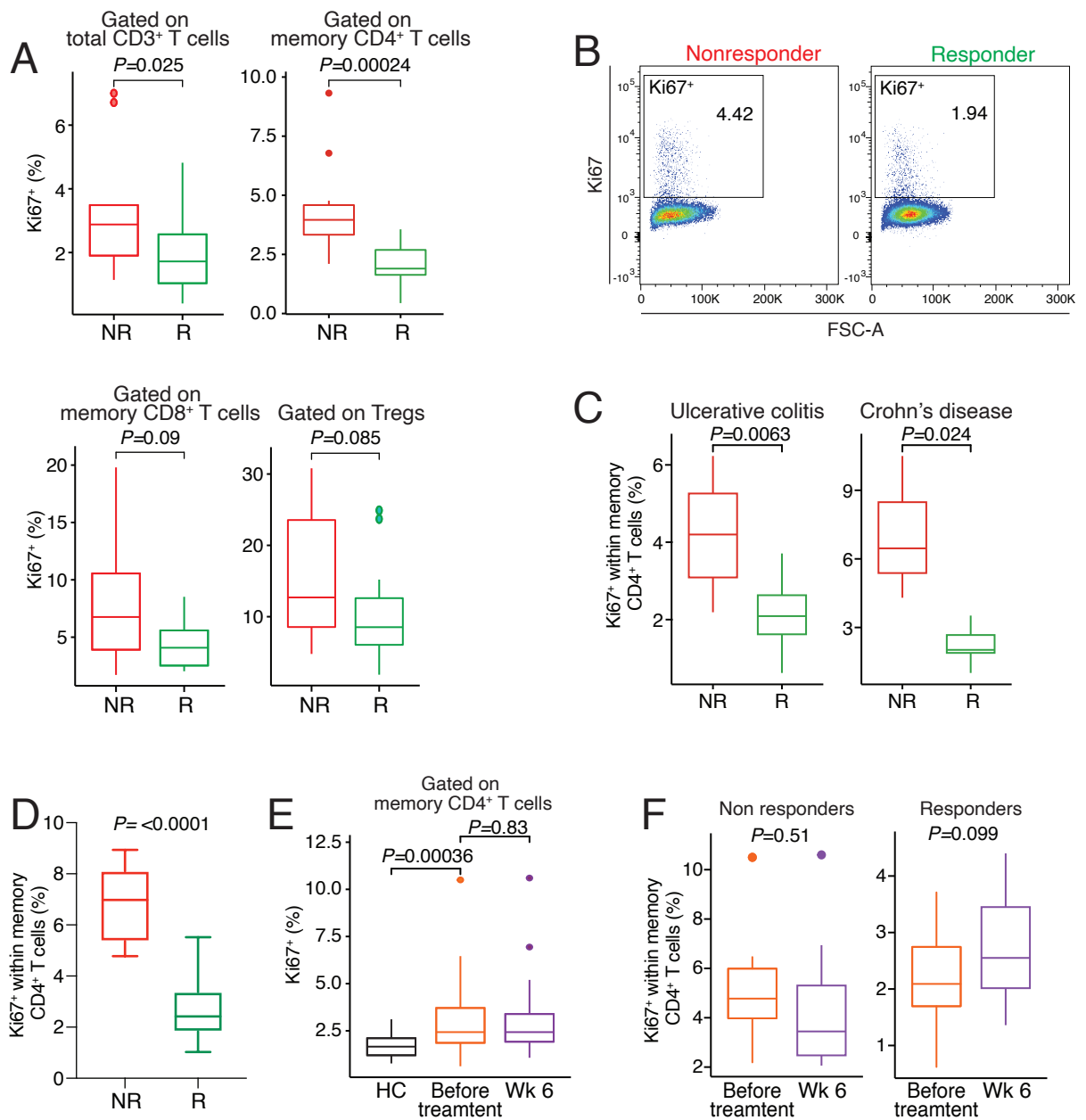
## **Supplementary Figure 12 Model validation and optimization**

(A) Validation of the best-performing data combination models (cf. Figure 5B), using an iterative approach. Based on the low test-accuracy of 0.4 in the FACS&CyTOF $\alpha$ 4 $\beta$ 7 model, we excluded it from further analysis.

(B) Overview of the top five features with highest absolute feature coefficient within the four best-performing data combination models (cf. Figure 5B). Shown are box-plots across repeated 2-fold cross validation runs.

(C) Analysis of the feature space derived from the iterative selection approach using automatic regularizers. The overlap in the feature space between our selection strategy and both regularizers is shown, similar to the presentation in Figure 5C. All features identified by our strategy and using Lasso were also selected by ElasticNet. The full analysis is included in **Suppl. Table 4**.

(D) Predictive capacity of selected marker combinations from the validated top-performing models in the reduced patient, analogous to Figure 5D.



Supplementary Figure 13

**Supplementary Figure 13. Distinctive features of proliferating effector CD4<sup>+</sup> T cells in vedolizumab non-responsive patients.**

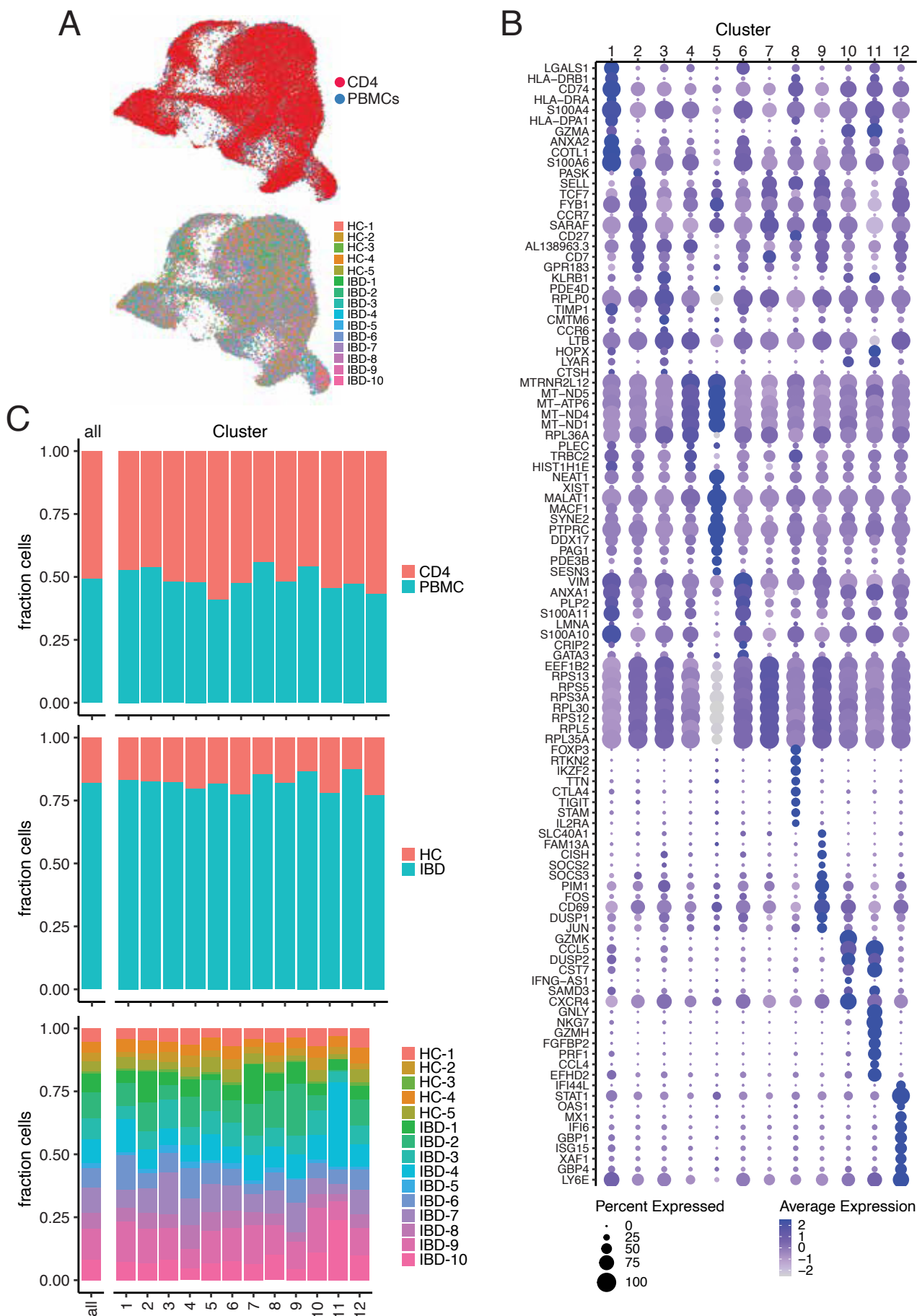
(A) Box plots illustrating the percentage of Ki67<sup>+</sup> cells within the indicated T cell subsets among IBD patients prior to initiating therapy (responders n=19, non-responders n=10).

(B) FACS plot showing Ki67 expression in memory CD4<sup>+</sup> T cells in representative donors.

(C) Box plots depicting the percentage of Ki67<sup>+</sup> cells within memory CD4<sup>+</sup> T cells in UC and CD patients. Data related to (A).

(D) Biomarker validation in an independent cohort of IBD patients treated with vedolizumab (n=26). Shown is the percentage of Ki67<sup>+</sup> within memory CD4<sup>+</sup> cells. Related to Figure 5F.

(E, F) Box plots depicting the percentage of Ki67<sup>+</sup> cells within memory CD4<sup>+</sup> T cells across three groups: healthy controls (HC, n=26), IBD patients before treatment (n=41), and IBD patients 6 weeks after treatment initiation (n=35). (F) Data segregated according to therapy response. (A-F) Mann-Whitney test.



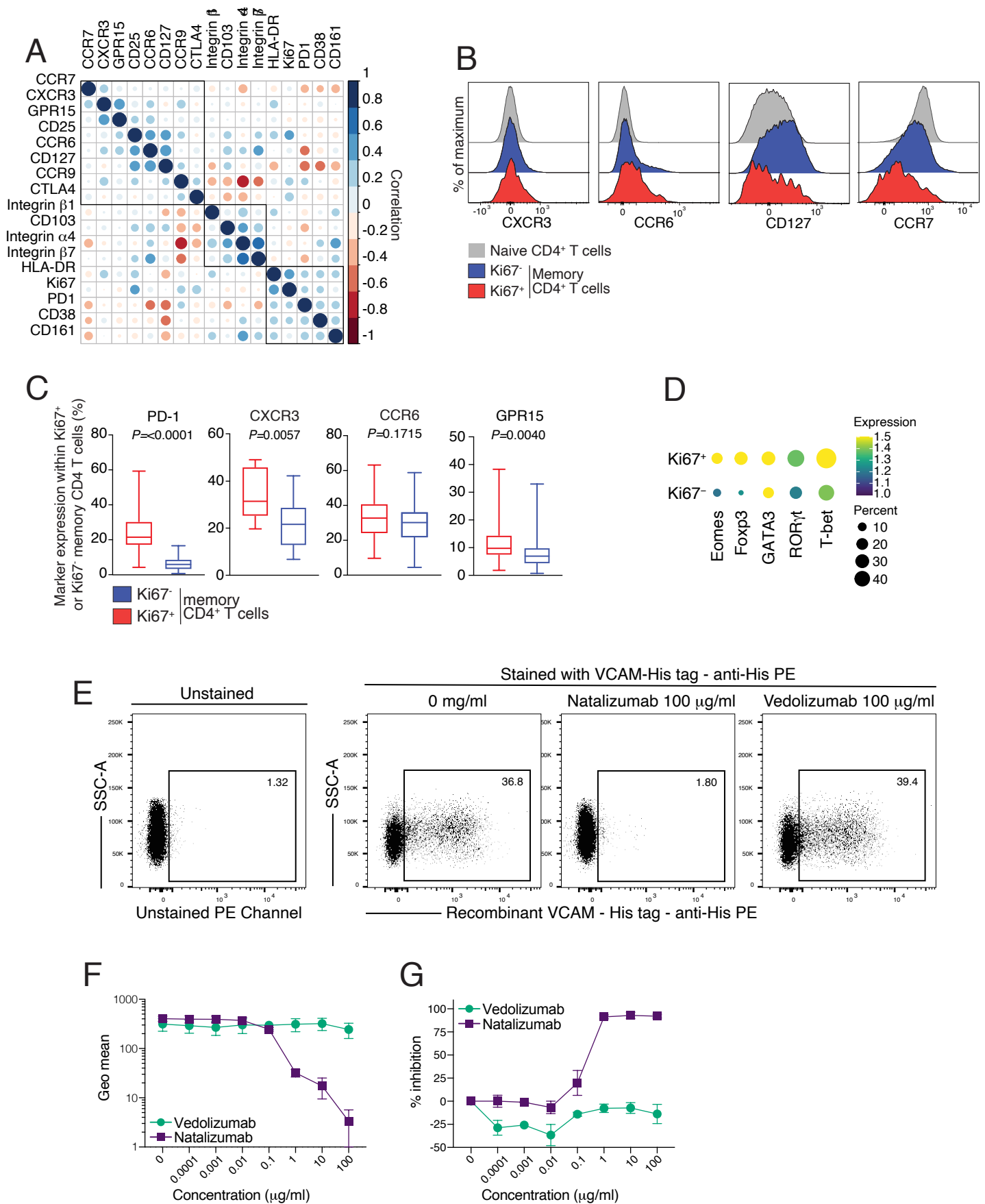
Supplementary Figure 14

**Supplementary Figure 14. Characterization of CD4<sup>+</sup> memory T cells through single-cell RNA sequencing (scRNA-seq).**

(A) UMAP plots of scRNA-seq profiles of memory CD4<sup>+</sup> T cells derived from total CD45<sup>+</sup> or sorted CD3<sup>+</sup>CD4<sup>+</sup>CD45RA<sup>low</sup> cells with color-coded representation of cell source (CD or PBMC) and patients (data derived from 25 samples: HC, n=5; IBD, n=10 matched before and after treatment). Sorted CD45<sup>+</sup> and CD3<sup>+</sup>CD4<sup>+</sup>CD45RA<sup>low</sup> cells from matched peripheral blood of 5 healthy controls and 10 ulcerative colitis patients before and 6 weeks after vedolizumab (VDZ) treatment initiation were subjected to scRNA-seq analysis.

(B) Dot plot of cluster marker genes for the CD4<sup>+</sup>-memory T cell subset in scRNA-seq data from PBMCs and memory CD4<sup>+</sup>-sorted cells. Dot size indicates percentage of positive cells, color scale indicates scaled average expression.

(C) Proportions of cells derived from PBMCs or CD4<sup>+</sup> memory T cells in healthy versus IBD patients and different donors.



Supplementary Figure 15

**Supplementary Figure 15. Characterization of proliferating effector CD4<sup>+</sup> T cells in peripheral blood.**

(A) Dot plot showing correlations between FACS surface markers within CD4<sup>+</sup> memory T cells in IBD patients before therapy (n=41).

(B) Histograms showcasing FACS surface marker expression in a representative IBD patient. The plot displays the expression of specified markers within naïve, Ki67<sup>+</sup>, and Ki67<sup>-</sup> memory CD4<sup>+</sup> T cells (n=29-44).

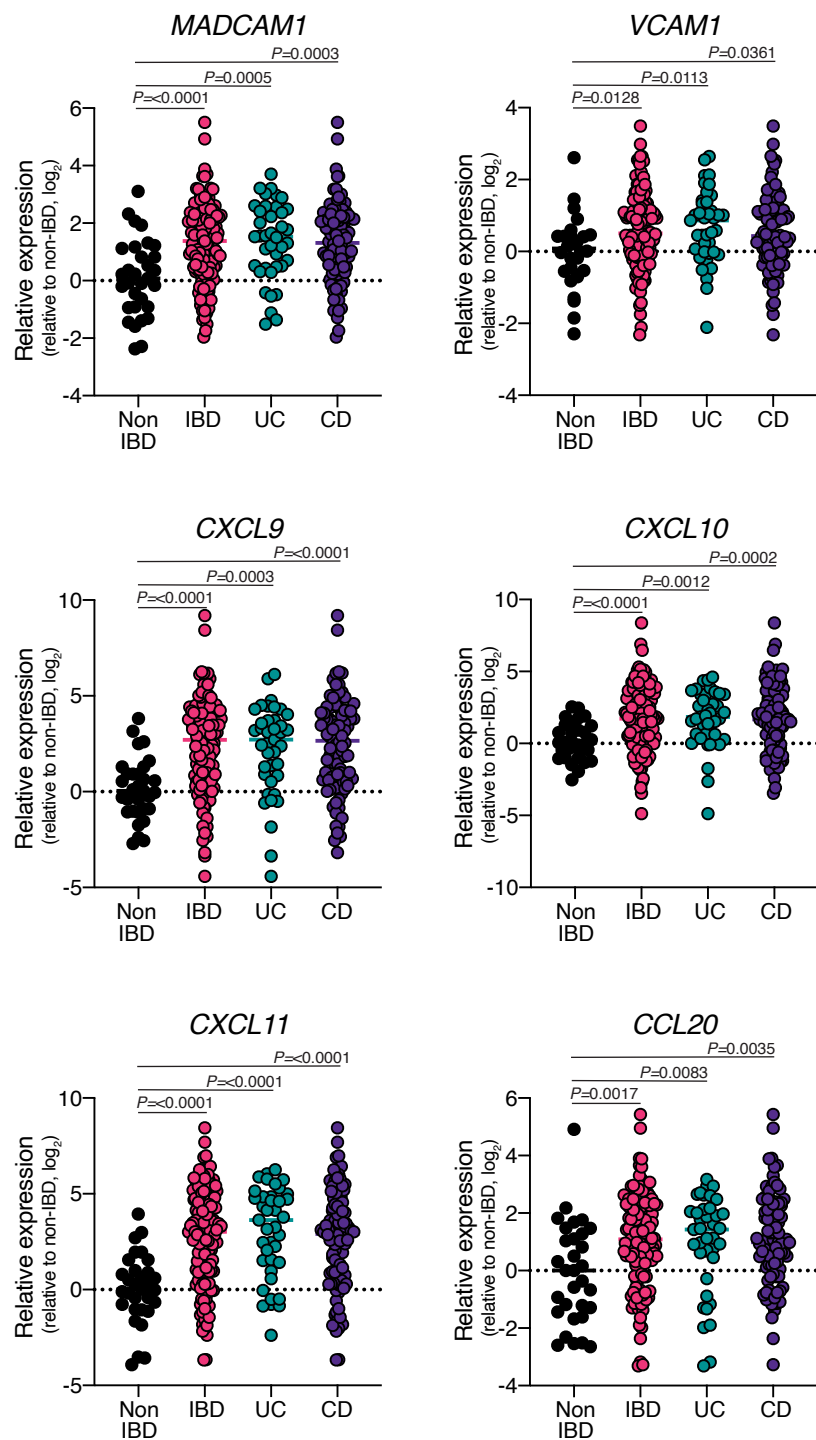
(C) Percentage of indicated surface marker expression within Ki67<sup>+</sup> and Ki67<sup>-</sup> memory CD4<sup>+</sup> T cells in IBD patients. n=29-44. Mann-Whitney test.

(D) Dot plot showing geomean expression of FACS transcription factors on Ki67<sup>+</sup> or Ki67<sup>-</sup> memory CD4<sup>+</sup> T cells in IBD patients before therapy (n=41). Expression is normalized to naïve CD4<sup>+</sup> T cells.

(E) Detection of recombinant VCAM His tag on memory CD4<sup>+</sup> T cells following incubation with natalizumab or vedolizumab (0 or 100 µg/ml). An unstained control is shown on the left. Dot plots represent one exemplary donor.

(F) Geometric mean intensity (Geomean) of VCAM binding on memory CD4<sup>+</sup> T cells after incubation with increasing concentrations of Vedolizumab or Natalizumab.

(G) Percentage inhibition of VCAM binding on memory CD4<sup>+</sup> T cells following incubation with increasing concentrations of natalizumab or vedolizumab. (E-G) Data are derived from three independent donors (n=3).

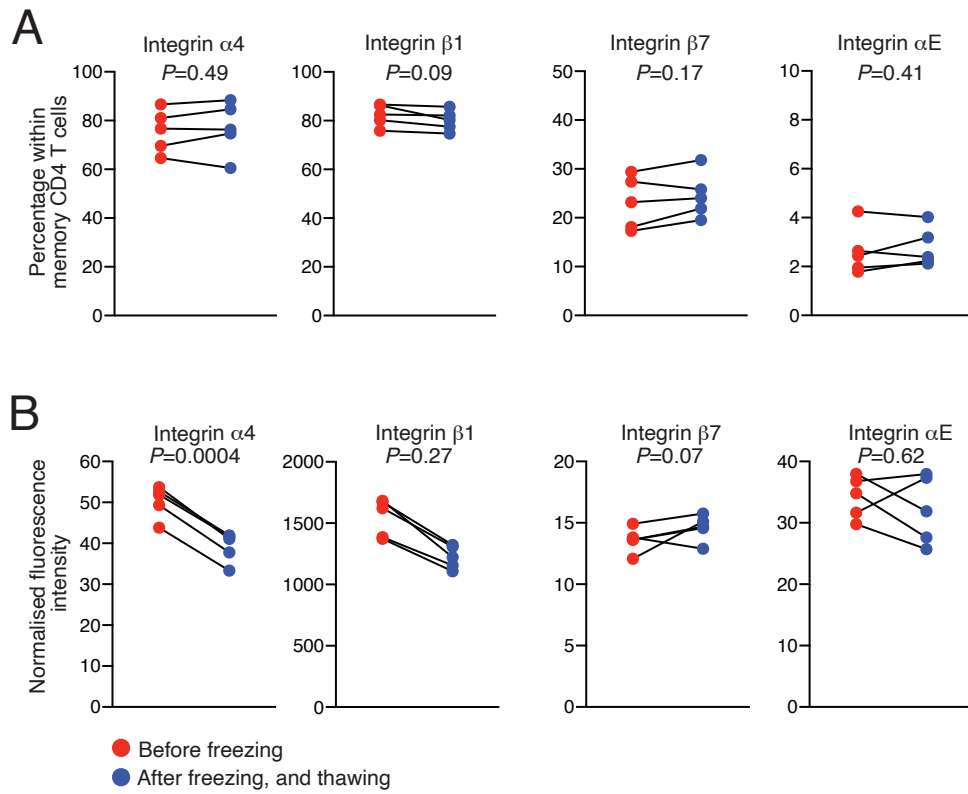


Supplementary Figure 16

**Supplementary Figure 16. Expression of chemokine ligands and adhesion molecules in mucosal tissues of non-IBD controls and IBD patients.**

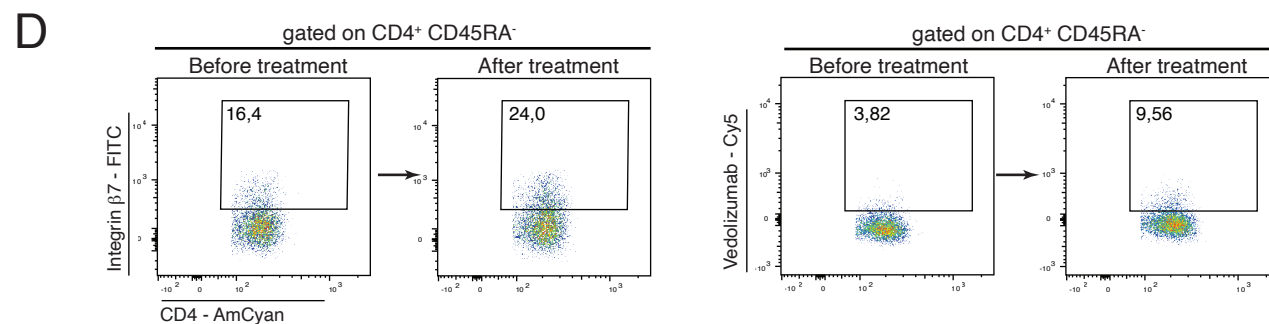
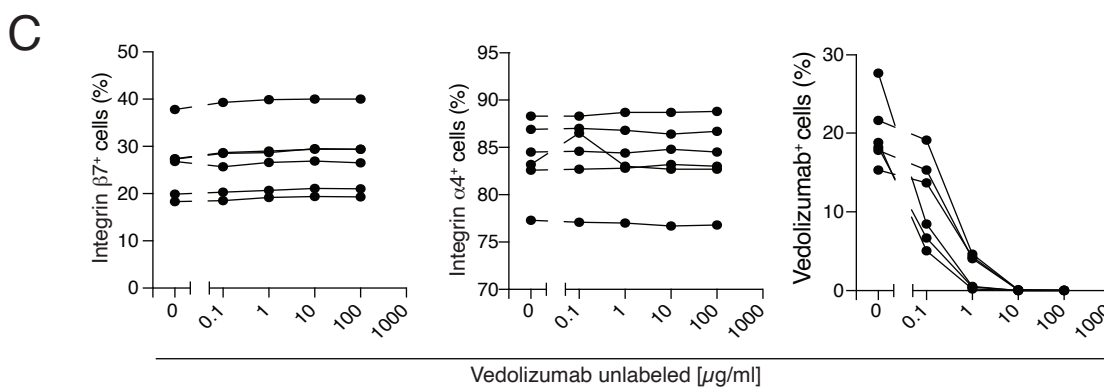
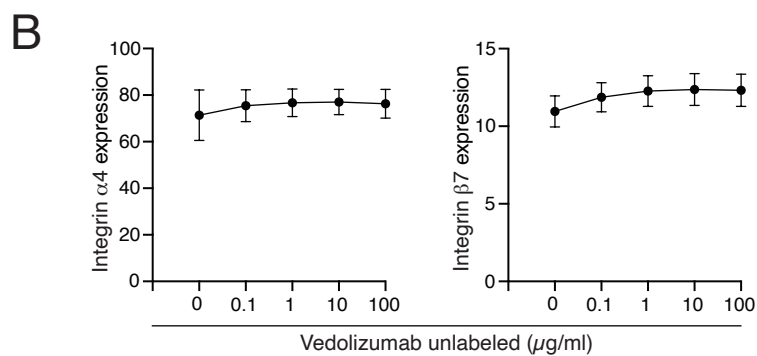
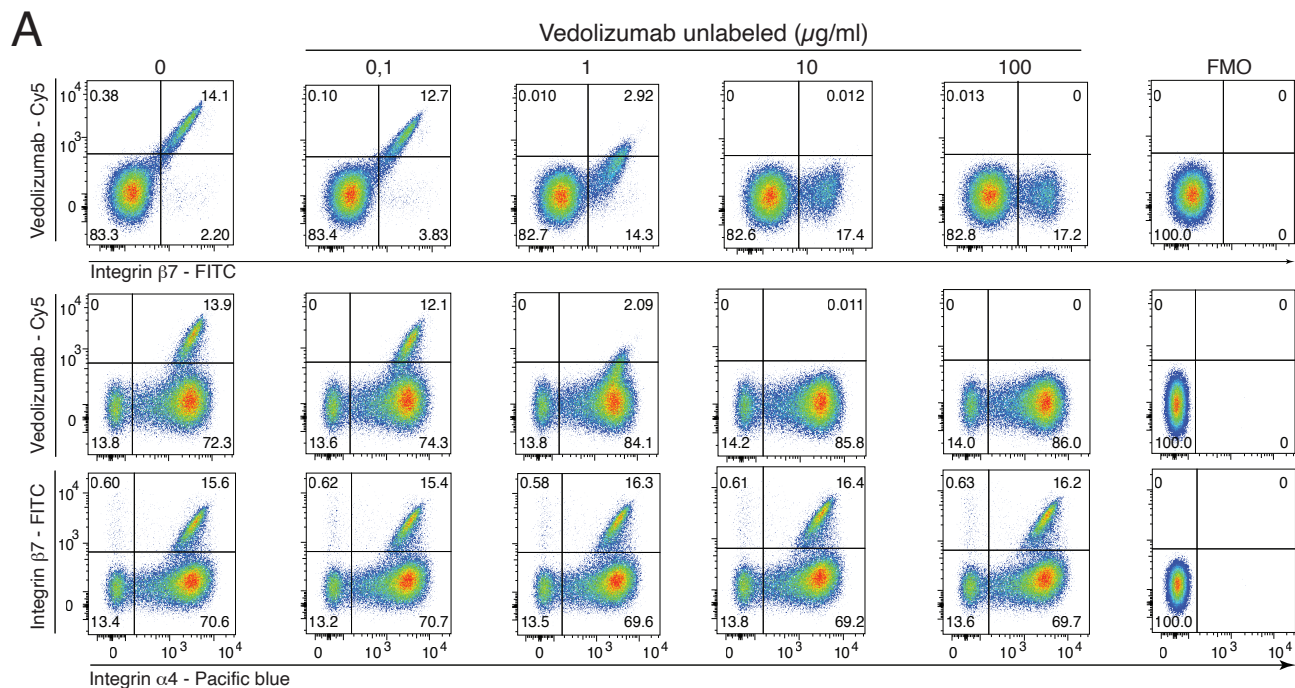
Expression of the indicated genes in mucosal samples obtained from non-IBD controls (n=32), UC patients (n=39), and CD patients (n=85). Endoscopic biopsies and surgical specimens were collected within the IBDome cohort and subjected to RNA sequencing. The gene expression in UC and CD samples were normalized to the median expression level in non-IBD control samples. Statistical comparisons were made using one-way ANOVA with Dunn's multiple-comparison tests.





Technical Supplementary Figure 1

514 **Technical Supplementary Fig. 1. Assessment of integrin stability after a freeze-thaw cycle**  
515 **of PBMCs.**  
516 (A, B) Percentage and expression intensity of the indicated molecules before and after freezing  
517 and thawing of PBMCs. Data gated on CD4<sup>+</sup> memory T cells (n=5 donors).

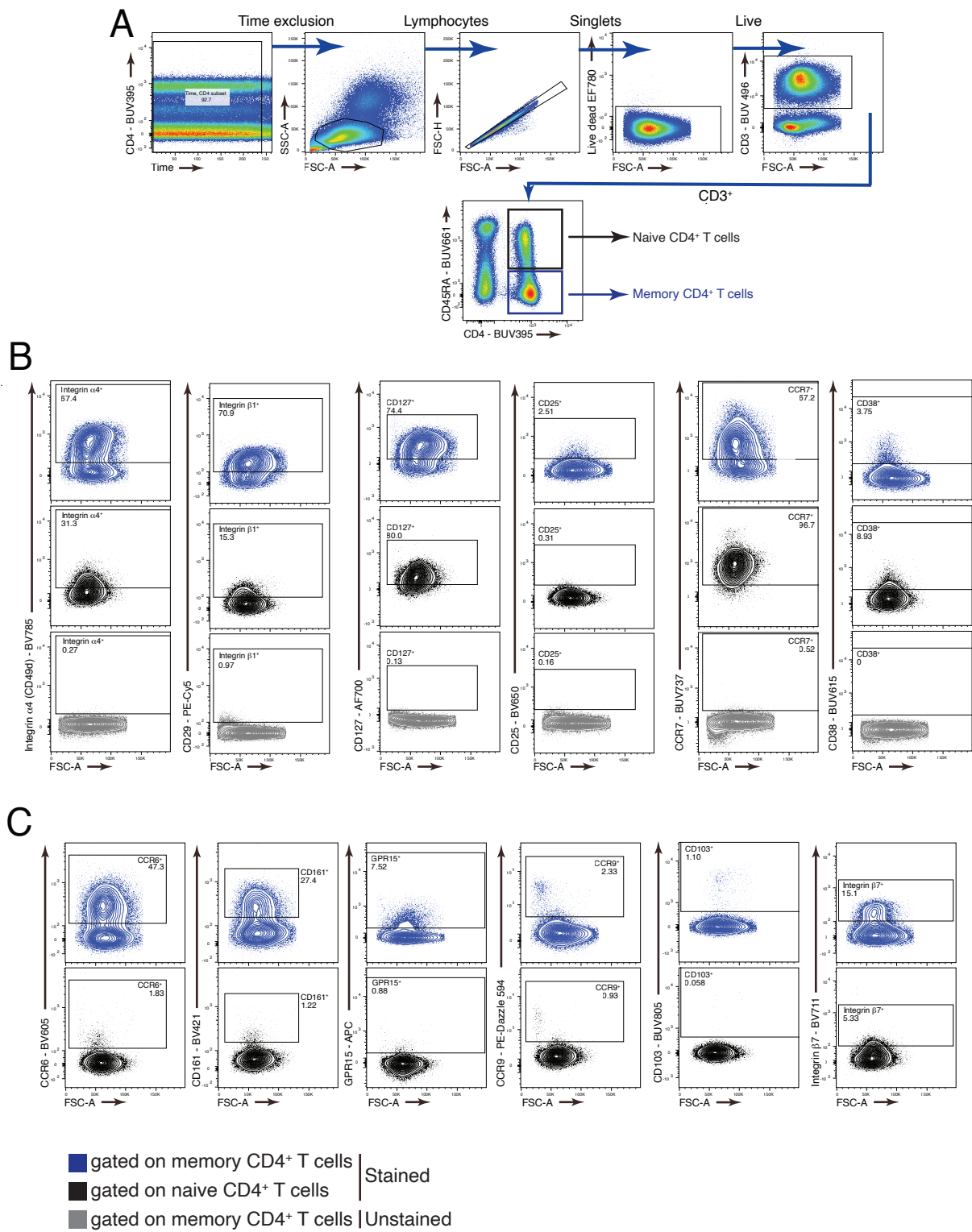


Technical Supplementary Figure 2

518 **Technical Supplementary Fig. 2. Impact of vedolizumab treatment on the detection of**  
519 **integrin  $\alpha 4\beta 7$  by flow and mass cytometry.**

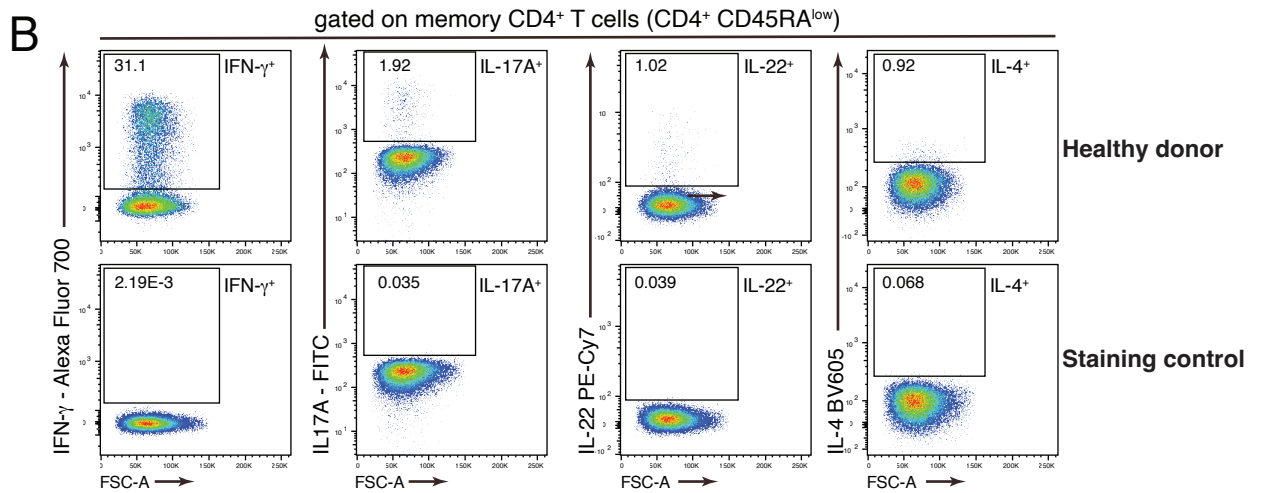
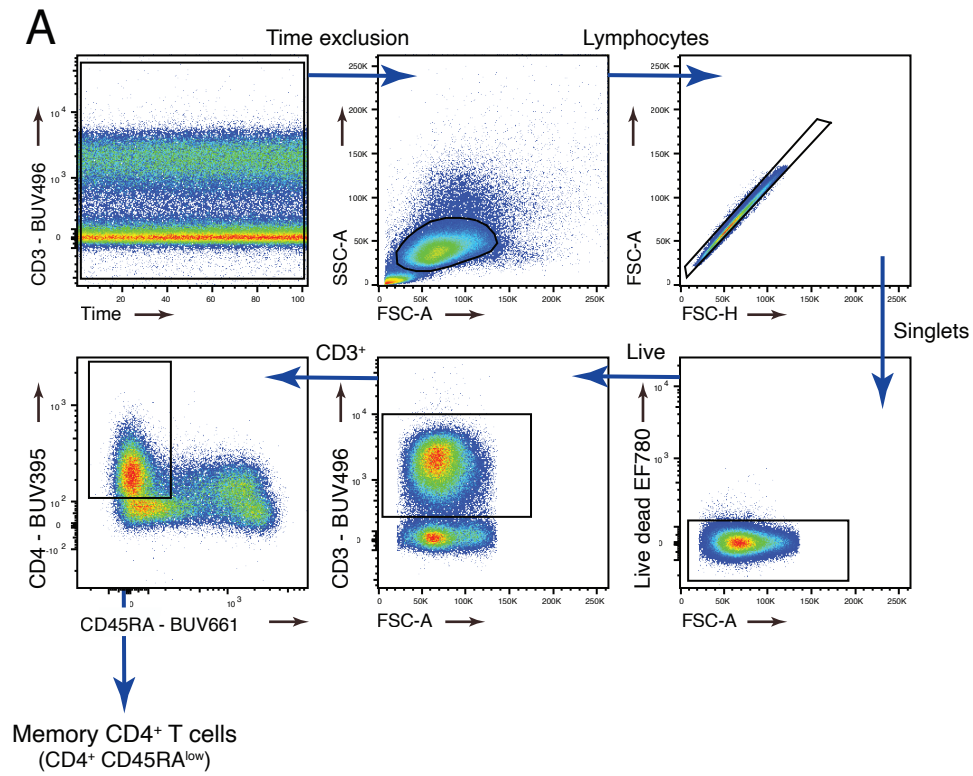
520 (A-C) Staining of integrin  $\alpha 4$  and integrin  $\beta 7$  after *in vitro* treatment of PBMCs with  
521 vedolizumab. PBMCs were isolated and incubated with unlabelled vedolizumab with the  
522 indicated concentrations. Integrin  $\alpha 4\beta 7$  was detected by Cy5-labelled vedolizumab (DRFZ) or  
523 FITC-labelled anti-integrin  $\beta 7$  antibody (clone FIB504). (B) The expression intensity (Geo  
524 mean index, expression normalized to naïve  $CD4^+$  T cells) of integrin  $\alpha 4$  and integrin  $\beta 7$  is  
525 shown. (C) Frequency of integrin  $\alpha 4^+$  and integrin  $\beta 7^+$   $CD4^+$  memory T cells are shown after  
526 treatment with vedolizumab with the indicated concentrations. Data representative of 3  
527 experiments with a total n=6.

528 (D) Integrin  $\beta 7^+$  staining on SMART-tube fixed whole blood from an IBD patient before (0  
529 weeks) and after (6 weeks) of treatment. A representative donor is shown.



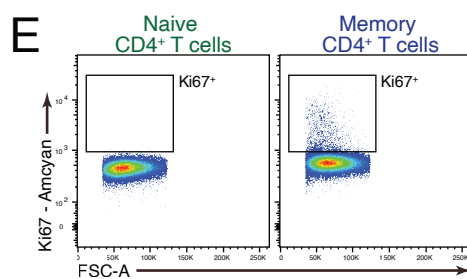
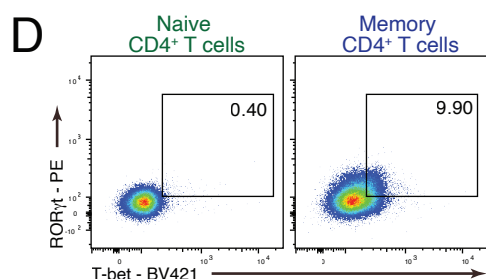
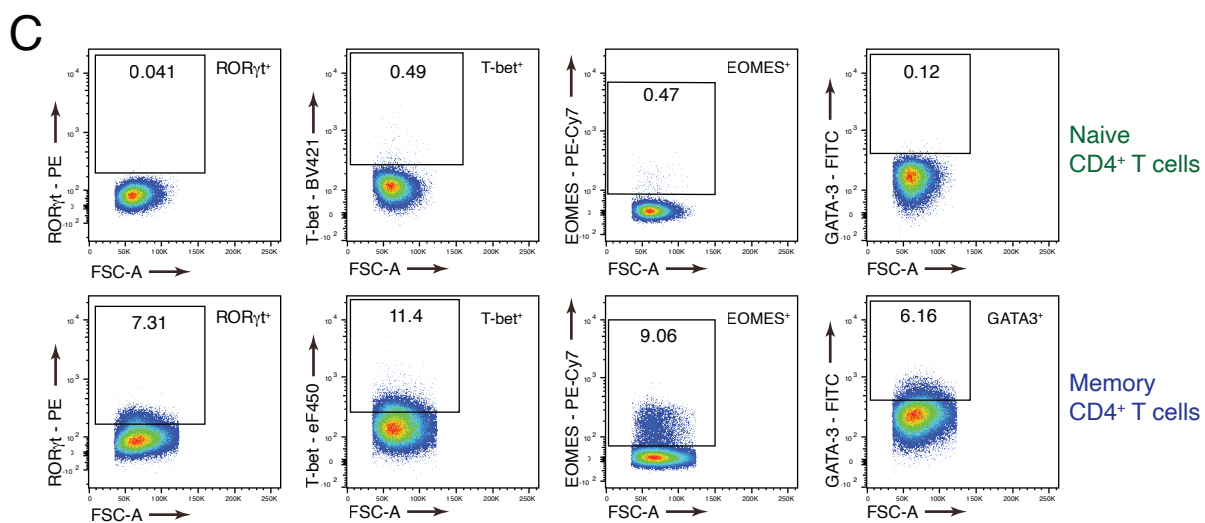
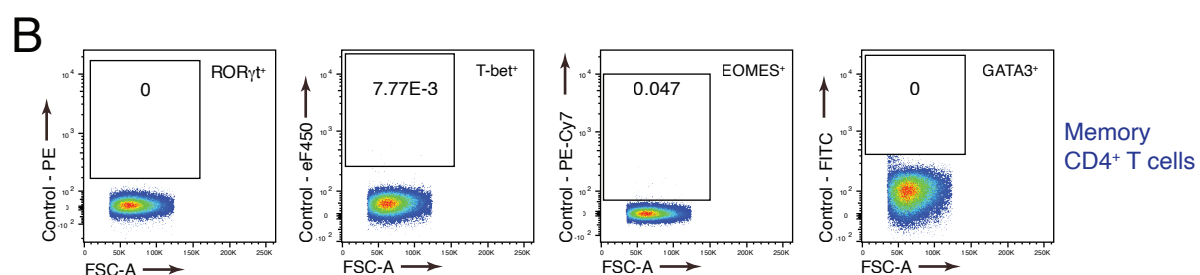
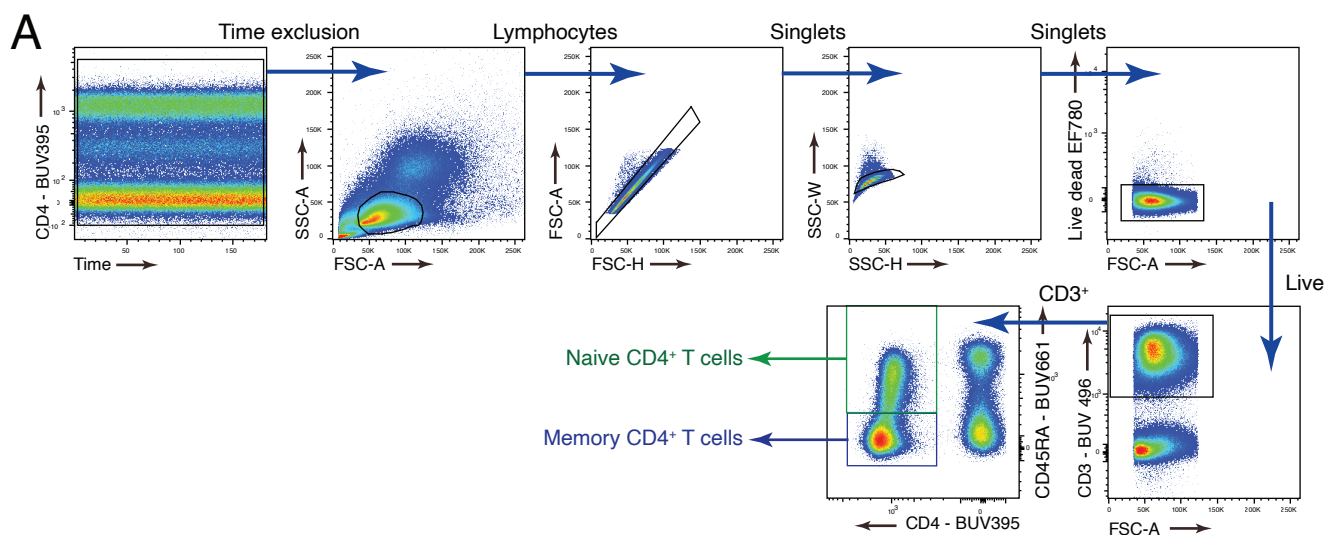
Technical Supplementary Figure 3

530 **Technical Supplementary Fig. 3. Identification of peripheral blood CD4<sup>+</sup> memory T cells**  
531 **via flow cytometry.**  
532 (A) Gating strategy for CD4<sup>+</sup> memory T cells shown for one representative donor.  
533 (B, C) Dot plots depicting the expression of the indicated surface markers. Naïve, memory  
534 CD4<sup>+</sup> T cells, and staining control are shown for each surface marker for one representative  
535 donor.



Technical Supplementary Figure 4

536 **Technical Supplementary Fig. 4. Exploring cytokine expression in CD4<sup>+</sup> memory T cells**  
537 **using intracellular cytokine staining and flow cytometry.**  
538 (A) Gating strategy for CD4<sup>+</sup> memory T cells after stimulation with PMA/ionomycin and  
539 intracellular staining.  
540 (B) Dot plots depicting the expression of the indicated cytokines. CD4<sup>+</sup> memory T cells, and  
541 staining control (no stimulation control) are shown for each cytokine.



Technical Supplementary Figure 5

542 **Technical Supplementary Fig. 5. Evaluating the expression of transcription factors in**  
543 **CD4<sup>+</sup> memory T cells using flow cytometry.**

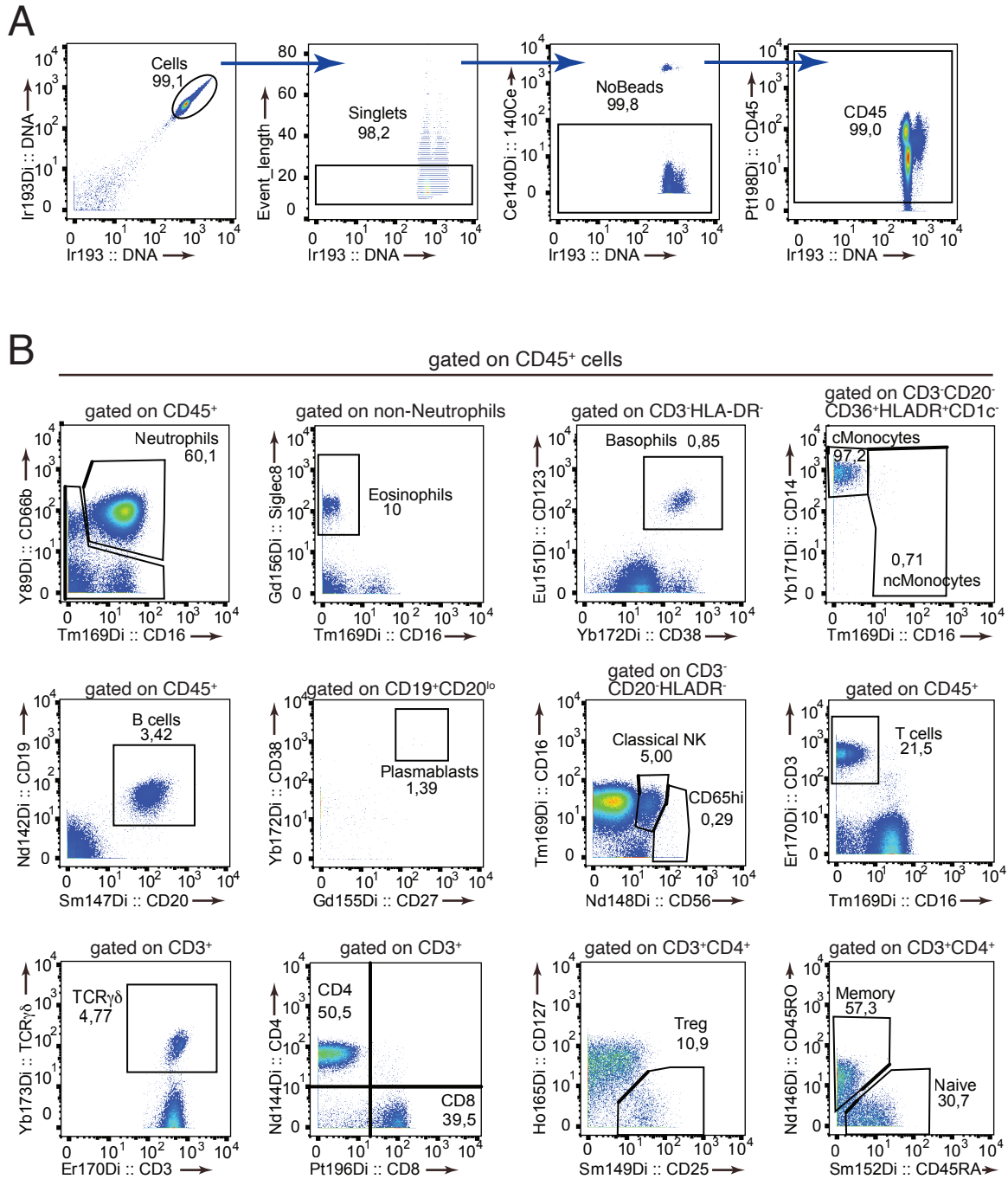
544 (A) Gating strategy for memory CD4<sup>+</sup> T cells after fixation with eBioscience™  
545 Foxp3/Transcription Factor Staining Buffer Set and intracellular staining.

546 (B) Cells were stained with surface markers, and the indicated channels for the transcription  
547 factors were left empty. Gated on CD4<sup>+</sup> memory T cells.

548 (C) Dot plots depicting the expression of the indicated transcription factors. Cells were stained  
549 for the indicated transcription factors. Shown are naïve and memory CD4<sup>+</sup> T cells.

550 (D) Dot plots showing the co-expression of RORγt and T-bet in naïve and memory CD4<sup>+</sup> T  
551 cells.

552 (E) Dot plots showing the expression of Ki67 in naïve and memory CD4<sup>+</sup> T cells.

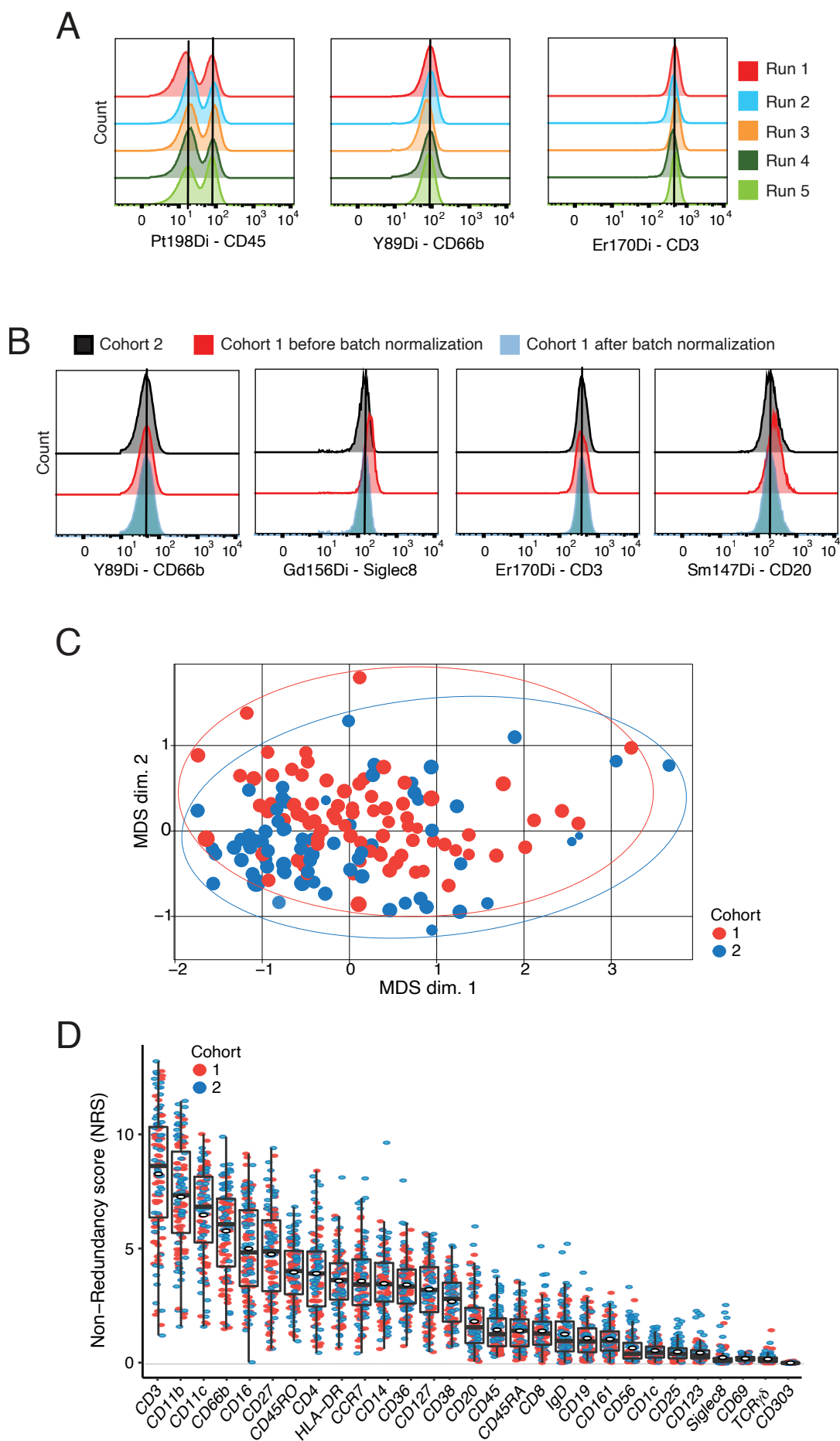


Technical Supplementary Figure 6

553 **Technical Supplementary Fig. 6. Immune cell subtype identification using mass**  
554 **cytometry.**

555 (A) Gating approach and identification of CD45<sup>+</sup> cells.

556 (B) Dot plots showing the expression of the indicated markers and the identification of the  
557 depicted cell populations.



Technical Supplementary Figure 7

558 **Technical Supplementary Fig. 7. Quality control for mass cytometry data acquisition.**

559 (A) Histograms depicting expression of exemplary lineage markers among anchor samples  
560 (from the same donor) across different runs within a single cohort.

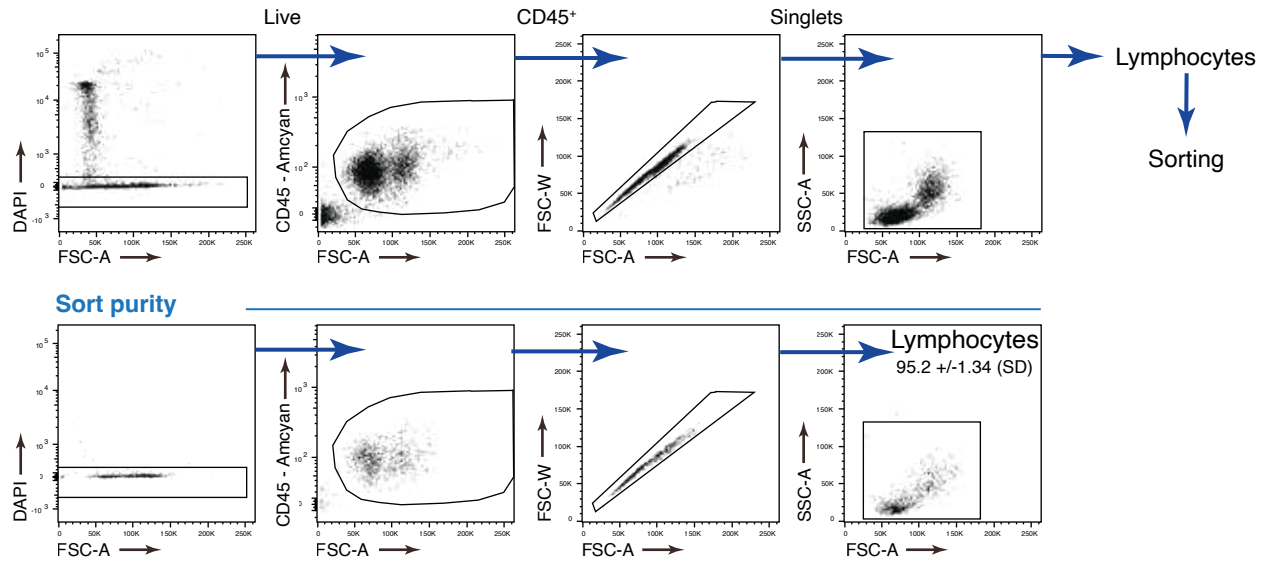
561 (B) Histograms depicting expression of exemplary lineage markers within the same donor  
562 across different cohorts, both before and after batch normalization.

563 (C) MDS plot showing the distribution of both cohorts after batch normalization. Each sample  
564 represents one plot.

565 (D) Boxplot of lineage marker distribution calculated by non-redundancy score (NRS) between  
566 cohorts 1 and 2.

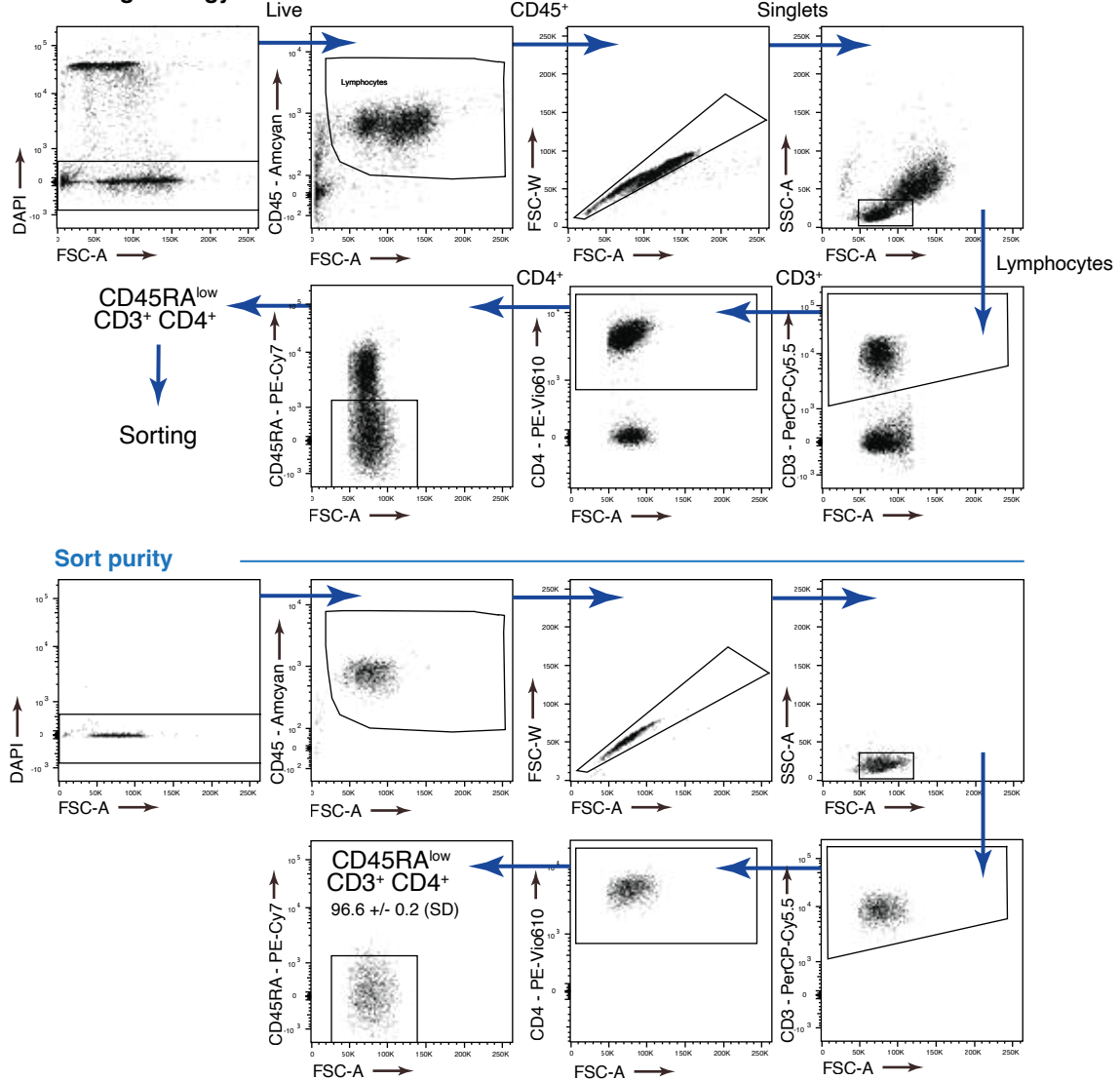
**A**

**Sorting strategy**



**B**

**Sorting strategy**



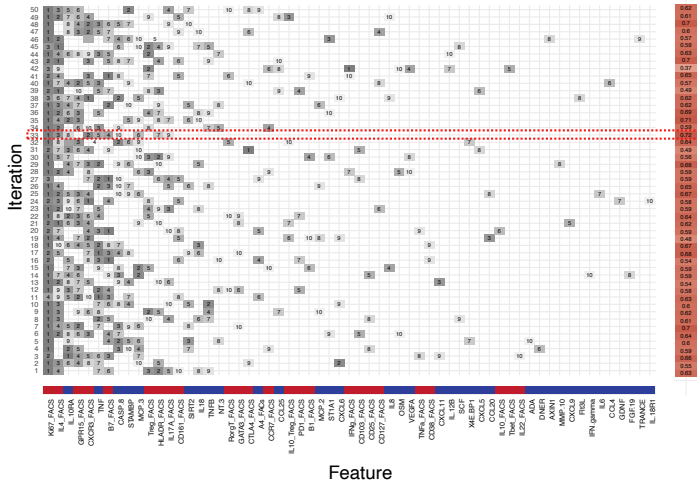
567 **Technical Supplementary Fig. 8. Fluorescence-activated cell sorting of peripheral blood**  
568 **mononuclear cells and CD4<sup>+</sup> memory T cells for single-cell sequencing.**

569 (A) Sorting strategy of PBMCs (upper panel) and purity control after sorting (lower panel) are  
570 shown.

571 (B) Sorting strategy of CD4<sup>+</sup> memory T cells (upper panel) and purity control after sorting are  
572 shown.

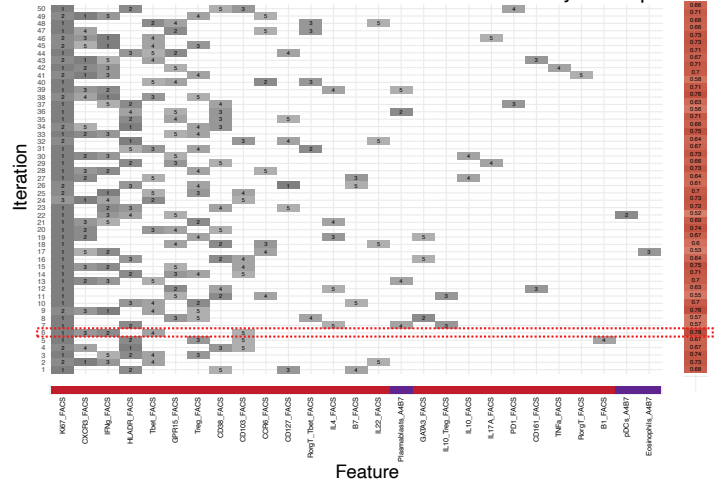
A

Olink + FACS



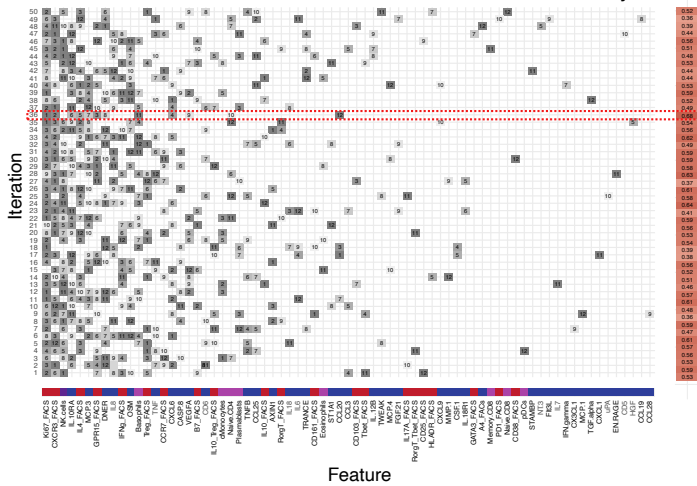
B

FACS + CyTOFα4β7



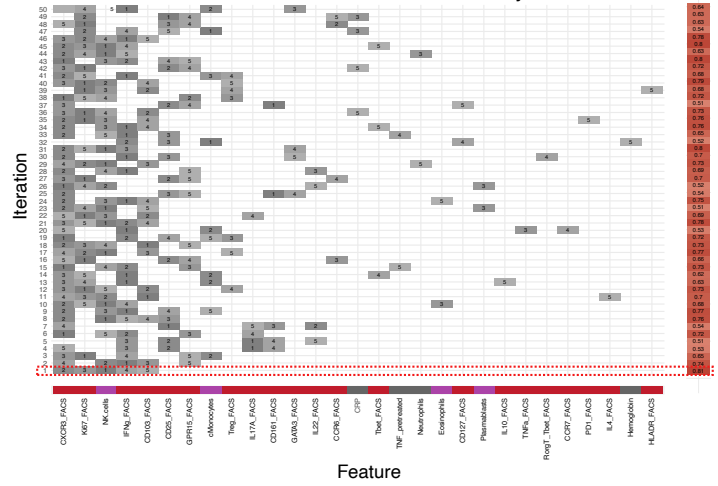
C

Olink &amp; FACS &amp; CyTOF



D

FACS &amp; CyTOF &amp; Clinical



573 **Technical Supplementary Figure 9. Overview of the iterative model validation workflow.**

574 (A-D) Selection of feature panels across iterations for the indicated models (cf. Figure 5B).

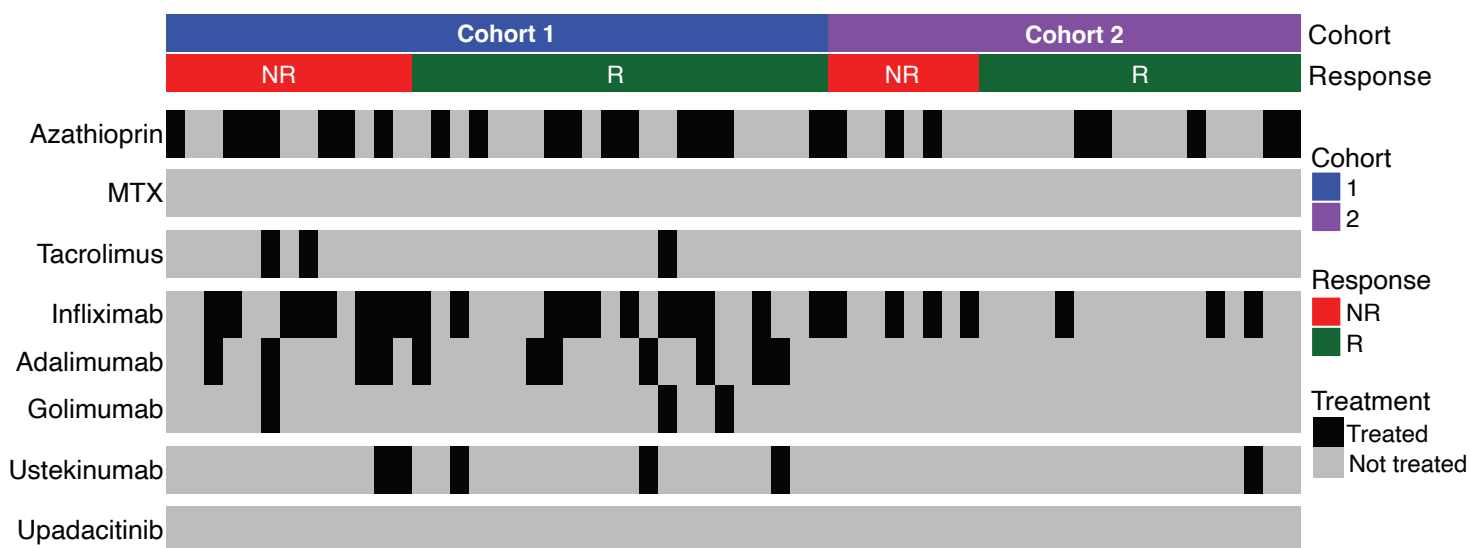
575 Feature panels are defined by the top 10% of features based on absolute feature coefficients in

576 each iteration (rows). Small numbers indicate feature ranks within each iteration. Robustness

577 was calculated in terms of the AUC across inclusion rates (heatmaps) using cut-offs from 10%

578 to 100%. The iteration with the highest robustness was selected for further analysis. See

579 Methods, Section ‘Machine-learning – iterative model validation’.



Technical Supplementary Figure 10

580 **Technical Supplementary Figure 10. Overview of previous lines of therapy for both**  
581 **cohort 1 and cohort 2.**

582 The heatmap indicates the various treatments administered to patients prior to vedolizumab  
583 therapy. The treatments include primarily anti-TNF, anti-IL-12p40, and the JAK inhibitor  
584 Upadacitinib. Both cohorts are shown, along with their respective therapy responses.  
585 Notably, no patients received anti-IL-23p19 therapy.