Mass cytometry reveals systemic and local immune signatures that distinguish inflammatory bowel diseases

Samuel J.S. Rubin1,2, Lawrence Bai1,2, Yeneneh Haileselassie2, Gotzone Garay3, Chohee Yun2,4, Laren Becker2, Sarah E. Streett2, Sidhartha R. Sinha2 & Aida Habtezion1,2

Inflammatory bowel disease (IBD) includes Crohn’s disease and ulcerative colitis. Each disease is characterized by a diverse set of potential manifestations, which determine patients’ disease phenotype. Current understanding of phenotype determinants is limited, despite increasing prevalence and healthcare costs. Diagnosis and monitoring of disease requires invasive procedures, such as endoscopy and tissue biopsy. Here we report signatures of heterogeneity between disease diagnoses and phenotypes. Using mass cytometry, we analyze leukocyte subsets, characterize their function(s), and examine gut-homing molecule expression in blood and intestinal tissue from healthy and/or IBD subjects. Some signatures persist in IBD despite remission, and many signatures are highly represented by leukocytes that express gut trafficking molecules. Moreover, distinct systemic and local immune signatures suggest patterns of cell localization in disease. Our findings highlight the importance of gut tropic leukocytes in circulation and reveal that blood-based immune signatures differentiate clinically relevant subsets of IBD.
Inflammatory bowel disease (IBD) is increasingly prevalent\(^1\,^2\). Despite available therapies, response remains challenging largely due to heterogeneity of clinical phenotypes\(^3\). IBD is divided into Crohn’s disease (CD) and ulcerative colitis (UC), both chronic relapsing/remitting immune-mediated conditions thought to be triggered by environmental factors in genetically predisposed individuals\(^4\). Within IBD, there is significant heterogeneity, especially amongst CD patients. CD can affect any part of the gastrointestinal tract, and disease lesions can present as ulcers, strictures, or penetrate from the lumen to the fat wrapping the intestines\(^5\). CD can be classified by disease behavior—inflammatory (B1; non-stricturing nonpenetrating), stricturing (B2), or penetrating (B3)\(^6\). CD can also be categorized by intestinal region(s) affected, grouped commonly by ileal, ileocolonic, or colonic disease. The majority of CD patients with colonic and/or rectal involvement develop perianal disease\(^7\). However, there are no objective measures of these disease phenotypes for use as markers of disease, tools to follow disease course, and/or enhance understanding of pathogenesis.

Compared to CD, UC is a comparatively homogenous disease restricted to the colon, generally characterized by ulcerations and pseudopolyps\(^8\). In the absence of treatment effects, UC always involves the rectum, and disease is typically classified as proctitis (E1; restricted to the rectum), left-sided (E2; involvement from the rectum to the splenic flexure), or pan colitis (E3; extensive disease beyond the splenic flexure)\(^8\,^9\). Conversely, CD is typically patchy with skip lesions\(^10\). Thus, UC includes a more uniform set of disease phenotypes.

CD and UC can be distinguished clinically by disease location (UC is restricted to the colon, while CD can involve any part of the gastrointestinal tract) and disease phenotype (stricturing or penetrating can indicate CD), using endoscopy and imaging\(^5\). However, gathering this information can be invasive and costly, and still up to 20% of IBD cases with colonic disease are indeterminate\(^10\). About 10% of patients who undergo colectomy for perceived UC are subsequently diagnosed with CD previously confined to the colon\(^10\,^11\). Because CD and UC can each present with a variety of disease phenotypes, identification of better markers using less invasive methods would enable earlier diagnosis as well as improved monitoring and treatment of disease.

Despite roles of host genetics, the immune system, the microbiota, and the environment in the pathogenesis of IBD\(^4\), factors responsible for the breadth of disease manifestations between patients (heterogeneity) are not well defined. Identifying and monitoring the symptoms, severity, behavior, and therapeutic response that characterize each disease manifestation (phenotype) is critical for improving health, yet frequently requires the use of invasive procedures, such as endoscopy.

In the tissue, IBD is characterized by intestinal barrier breakdown, allowing microbiota to prime the immune system. The ensuing inflammatory response involves recruitment of leukocytes from the periphery to the gut and is associated with dissemination of activated cells in circulation\(^12\,^13\). Since leukocytes that traffic to the gut from the periphery mediate this inflammatory response, we hypothesized that, by enriching for gut tropic cells in the blood, we could detect CD- and UC-specific as well as phenotype-specific signatures and investigate their relationship to the tissue. Since blood is easily accessible, identification of immune signatures in circulation that represent intestinal immunity would be highly informative and provide less invasive tools to assess disease.

We use mass cytometry (CyTOF) to resolve single cells and characterize their lineage, gut tropism, and function\(^14\). Consistent with our hypothesis, we identify blood signatures of disease diagnosis, phenotype, and state (flare/remission), heavily represented by gut tropic cell populations and reflective of the greater clinical heterogeneity observed for CD than UC. We report a multi-parameter generalized linear model to classify patients by disease diagnosis (CD/UC), suggesting that blood-based assays could reduce the delay, risk, discomfort, and cost associated with more invasive procedures.

**Results**

**Study design.** We collected blood and biopsies from IBD patients (Fig. 1a) and selected 68 subjects with gastroenterologist-confirmed IBD diagnoses or matched healthy controls (HC) (Table 1). Cohort 1 included blood samples from 56 subjects (Supplementary Table 1), and cohort 2 had paired blood and biopsy samples from 12 subjects (Supplementary Table 2). By cohort, disease-group demographics were matched as best as possible for sex and reports of extraintestinal manifestations (Table 1; Supplementary Tables 1, 2). There was no significant difference in disease-state proportions (flare/remission for cohort 1), sex, age, age at onset, disease duration, extraintestinal manifestations, or tissue-state proportions (inflamed/uninflamed for cohort 2) between CD and UC patients by cohort (Table 1). Thus, it was unnecessary to adjust data based on these parameters. Disease phenotypes were specific to each diagnosis and could not be matched. IBD samples included subjects in clinical remission or flare (determined by IBD specialists; see the Methods section), and a spectrum of disease activity scores were represented in each group.

We analyzed viably cryopreserved leukocytes from blood and tissue by CyTOF using panels with surface and intracellular antigens (Supplementary Table 3; Supplementary Figs 1, 2). We used four trafficking molecules to identify gut tropic cells: α4β7, a pan-gut-trafficking molecule and target of the therapeutic antibody vedolizumab\(^13\), CCR1, a trafficking molecule identified in GWAS studies and a marker of activity in CD\(^15\,^16\); CCR9, a lymphocyte trafficking molecule associated with small intestine tropism\(^13\); and GPR15, a T cell trafficking molecule that we and others showed to be important for trafficking to the colon\(^13\,^17\,^18\). While our CyTOF panels included phosphoproteins and functional markers, we found in pilot studies that cell stimulation was unnecessary to resolve differences in phospho-signaling between sample groups. Trafficking receptor expression patterns in tissue and blood shed light on local and peripheral immune responses since little is known about leukocyte trafficking to the gut in human, especially in the context of disease.

**Blood leukocytes demonstrate increased heterogeneity in CD.** We conducted targeted analysis of CyTOF data by manually gating and calculating median protein expression levels to compile 2208 parameters per sample, as well as unbiased analysis using viSNE, CITRUS, and Spade algorithms. Coefficients of variation (CVs) for each parameter were used as a proxy for disease group population diversity\(^19\), supporting clinical observations that CD includes more heterogeneous disease manifestations than UC (Fig. 1b). Samples from CD remission had significantly higher CVs than samples from UC remission or HCs, suggesting greater heterogeneity in CD remission compared to UC remission and HCs. Samples from CD flare had higher CVs but exhibited a much higher standard deviation and were not significantly different than from CD remission, UC flare, or HCs, suggesting an overall increase in heterogeneity amongst CD flare. Conversely, samples from UC flare had significantly higher CVs than samples from UC remission or HCs, suggesting distinct states of flare and remission in UC where flare is inherently more heterogeneous than remission regardless of diagnosis.
Gut-homing molecule expression distinguishes CD from UC. Manually calculated parameters were compared between all CD and UC blood samples in cohort 1, which revealed statistically significant differences in abundance of CCR9+GPR15+ CD56+ B cells and a4β7+CCR1+ CD56+ plasmablasts after correction for multiple testing (Fig. 1c). These data show that gut destined, or tropic, B cell subsets, including CCR9 and GPR15 co-expressing cells (Supplementary Fig. 6), appear to be relevant for disease distinction. Although expression of CD56 (neural cell adhesion molecule) is typically associated with NK cells, the protein is also expressed on other leukocytes, including activated B cells in lymphoma patients20 and T helper 1 (Th1) responses21. CD56 expression is consistent with elevated CD56+ B cells and plasmablasts here in CD, where Th1 responses are known to contribute to disease22.

Based on gene expression deconvolution analysis23 (Supplementary Fig. 8) and knowledge of IBD, we conducted a refined set of additional hypothesis-driven tests to compare CD and UC. This approach revealed six parameters in the blood significantly different between all CD and UC samples (Fig. 1d), as well as four significantly different for flare (Supplementary Fig. 3). The increase of basophils in CD compared with UC is consistent with gene expression deconvolution data (Supplementary Fig. 8). Moreover, median phospho-cyclic AMP-responsive element-binding protein (pCREB) expression by basophils was significantly higher in CD compared with UC, which is consistent with the association between CD activity and blood expression of the total CREB24 (Fig. 1d). Circulating pDCs, a4β7+ mDCs, and memory effector CD4+ T cells were also significantly increased in CD compared with UC (Fig. 1d). IgD−CD27− B cells were significantly increased among the total CD19+CD20+ B cells in UC as compared with CD, suggesting an antigen-driven B cell response (Fig. 1d). These IgD−CD27− memory B cells are larger and more granular than IgD+CD27− naive B cells25.
class switched, and somatically hypermutated. Expansion of IgD−CD27− B cells was reported in systemic lupus erythematosus, HIV, and rotavirus infections. In summary, we identified eight cellular features from a combination of unbiased and hypothesis-driven analyses that were significantly different between all CD and UC patients.

Blood heterogeneity is conserved in CD flare and remission. Next, we compared IBD patients in remission to HCs to identify signatures of subclinical disease persistent in remission, as well as patients in flare to those in remission to identify features specific to disease activity. We found few significant trends across CD patients (Fig. 2a, b). CD remission compared with HC revealed an increase in two gut tropic B cell populations identified by GPR15 and CCR9 expression (Fig. 2a). We compared CD flare to remission and found an increased abundance of highly activated CD38+HLA-DR+ CD8 T cells (Fig. 2b). CD flare was associated with increased median total 1kBa expression in IgD+ memory B

Table 1 Summary of demographic and clinical characteristics of the patients

|                         | Cohort 1 | Cohort 1 CD vs. UC p-value | Cohort 2 | Cohort 2 CD vs. UC p-value |
|-------------------------|----------|---------------------------|----------|---------------------------|
| Cohort size (number of patients) | 56       | n/a                       | 12       | n/a                       |
| Diagnoses (number of patients)   |          |                           |          |                           |
| CD                       | 24       | n/a                       | 6        | n/a                       |
| UC                       | 20       |                           | 6        |                           |
| HC                       | 12       |                           | 0        |                           |
| Disease status (number of patients in clinical flare/remission) |          | >0.9999                   |          | >0.9999                   |
| CD                       | 13/11    |                           | 0/6      |                           |
| UC                       | 10/10    |                           | 0/6      |                           |
| Sex (number of M/F patients)    |          |                           |          |                           |
| CD                       | 11/13    |                           | 1/5      |                           |
| UC                       | 7/13     |                           | 5/1      |                           |
| HC                       | 8/4      |                           | 0/0      |                           |
| Age (years)†             |          | 0.6172                    |          | 0.7259                    |
| CD                       | 39 [23–65] |                           | 44 [22–65] |                           |
| UC                       | 37 [19–75] |                           | 38 [25–60] |                           |
| HC                       | 51 [24–63] |                           | n/a      |                           |
| Age at onset (years)†     |          | 0.1422                    |          | 0.5996                    |
| CD                       | 18 [5–55] |                           | 31 [19–52] |                           |
| UC                       | 27.5 [12–60] |                          | 29.5 [21–43] |                           |
| Disease duration (years)† |          | 0.2547                    |          | 0.8495                    |
| CD                       | 13.5 [0.08–39] |                         | 9 [1–17] |                           |
| UC                       | 7.5 [1–42]  |                           | 10.5 [4–17] |                         |
| Reported extraintestinal manifestations (number of patients) |          | 0.2591                    |          | >0.9999                   |
| CD                       | 6        |                           | 3        |                           |
| UC                       | 2        |                           | 3        |                           |
| Crohn’s disease phenotype |          |                           |          |                           |
| Score (Harvey-Bradshaw Index) | 2.5 [0–7] |                           | 0.5 [0–6] |                           |
| Location (number of patients) |          |                           |          |                           |
| Ileum only               | 5        |                           | 1        |                           |
| Ileum and colon          | 15       |                           | 4        |                           |
| Colon only               | 4        |                           | 1        |                           |
| Behavior (number of patients) |          |                           |          |                           |
| Inflammatory             | 8        |                           | 0        |                           |
| Fistulizing              | 7        |                           | 2        |                           |
| Stricture                | 7        |                           | 2        |                           |
| Fistulizing and stricture| 2        |                           | 2        |                           |
| Reported perianal disease (number of patients) |          |                           |          |                           |
| CD                       | 7        |                           | 1        |                           |
| UC                       |          |                           |          |                           |
| Ulcerative colitis phenotype |          |                           |          |                           |
| Score (Partial Mayo score) | 4 [0–9] |                           | 0 [0–3] |                           |
| Location (N patients)     |          |                           |          |                           |
| Left-sided               | 7        |                           | 3        |                           |
| Pan colonic              | 12       |                           | 3        |                           |
| Proctitis                | 1        |                           | 0        |                           |
| Biopsies collected per patient (N) |          |                           |          |                           |
| CD                       | 0        |                           | n/a      |                           |
| UC                       | 0        |                           | 4 [3–5] |                           |
| Tissue state (number of inflamed/uninflamed biopsies) |          |                           |          |                           |
| CD                       | n/a      |                           | 11/12    |                           |
| UC                       | n/a      |                           | 5/13     |                           |

Cohort 1 contained blood samples, and cohort 2 contained paired blood and tissue biopsy samples. All clinical data reflects the time of sample collection. P-values are shown where applicable for CD vs. UC disease groups in cohorts 1 and 2. Statistics: unpaired two-tailed Student’s t test (cohort 1 age, t = 0.5036, df = 42; cohort 2 age, t = 0.3607, df = 10; cohort 1 age at onset, t = 1.496, df = 42; cohort 2 age at onset, t = 0.5421, df = 10; cohort 1 disease duration, t = 1.155, df = 42; cohort 2 disease duration, t = 0.1947, df = 10; cohort 2 biopsies collected per patient, t = 2.712, df = 10) and two-sided Fisher’s exact test (disease status; sex; reported extra-intestinal manifestations; tissue state). Sample sizes are shown in the table for each comparison. († = median [range]; CD = Crohn’s disease; UC = ulcerative colitis; HC = healthy control)
Fig. 2 Some circulating leukocyte signatures differentiate disease flare and remission while others persist in remission. Significant differences between CD remission (N = 11) and HC (N = 12) samples (a), CD flare (N = 13) and CD remission (N = 11) samples (b), UC remission (N = 10) and HC (N = 12) samples (c), and UC flare (N = 10) and UC remission (N = 10) samples (d) with p ≤ 0.05 after correction for multiple testing. On the right of (b, d) are Pearson correlations based on vectors that represent all significant feature values for individual patients shown in the heatmaps immediately to the left. Statistics: BH FDR-corrected unpaired two-tailed Student’s t test using Morpheus (see the Methods section; t-statistics shown in heatmap legends). Source data are provided as a Source Data file.

Fig. 3 Comparison of disease features between CD and UC. We compared CD patients with inflammatory versus fistulizing, inflammatory versus strictureing, and fistulizing versus strictureing disease (excluding two patients with history of both fistulizing and strictureting disease), we found one B cell signature (Fig. 3a), three signatures of monocyte and B cell subsets (Fig. 3b), and one dendritic cell (DC) signature (Fig. 3c), respectively, that distinguished these disease phenotypes. Interestingly, when we compared CD patients with and without perianal disease, we found a single difference in gut tropic Tregs (Fig. 3d). When we compared CD patients with only ileal versus only colonic disease, we noted a significant decrease in circulating a4b7+ CCR9+ subpopulations of mature NK and CD45RO+ NKT cells in ileal CD (Fig. 3e). Since there were no significant differences in the abundance of the total mature NK cells or total CD45RO+ NKT cells, differences in gut tropic subsets suggest enhanced tissue localization in ileal CD, although we did not have sufficient samples from patients with only ileal or only colonic CD in cohort 2 to investigate this further. In UC, only one highly activated GPR15 and CCR9 co-expressing CD4+ T cell population was increased in pan-colonic compared with left-sided disease (Fig. 3f). In summary, gut-tropic signatures in blood differentiated clinical disease subsets.

**Blood leukocytes reflect clinical stratifications of disease.** To identify blood signatures responsible for disease heterogeneity, we separated subjects by disease behavior and location. These stratifications are important for monitoring and treatment. When we compared CD patients with inflammatory versus fistulizing, inflammatory versus strictureing, and fistulizing versus strictureing disease (excluding two patients with history of both fistulizing and strictureting disease), we found one B cell signature (Fig. 3a), three signatures of monocyte and B cell subsets (Fig. 3b), and one dendritic cell (DC) signature (Fig. 3c), respectively, that distinguished these disease phenotypes. Interestingly, when we compared CD patients with and without perianal disease, we found a single difference in gut tropic Tregs (Fig. 3d). When we compared CD patients with only ileal versus only colonic disease, we noted a significant decrease in circulating a4b7+ CCR9+ subpopulations of mature NK and CD45RO+ NKT cells in ileal CD (Fig. 3e). Since there were no significant differences in the abundance of the total mature NK cells or total CD45RO+ NKT cells, differences in gut tropic subsets suggest enhanced tissue localization in ileal CD, although we did not have sufficient samples from patients with only ileal or only colonic CD in cohort 2 to investigate this further. In UC, only one highly activated GPR15 and CCR9 co-expressing CD4+ T cell population was increased in pan-colonic compared with left-sided disease (Fig. 3f). In summary, gut-tropic signatures in blood differentiated clinical disease subsets.
be curative for UC (total colectomy), but not CD. We compared samples from patients with colonic CD (Crohn's colitis) to UC. GPR15+ Tregs, naive IgD+ B cells, and GPR15+ naive IgD+ B cells were increased while CCR9+ GPR15+ plasmablasts were decreased in colonic CD compared with UC (Fig. 3g). Three of these four significantly different cell subsets were enriched for GPR15 while one cell subset was naive IgD+ B cells, demonstrating that trafficking molecule expression by blood leukocytes facilitates complex disease differentiation. These four blood-based features could help address the unmet need to differentiate colonic CD from UC.

**Blood and tissue cells are largely but not entirely distinct.** We analyzed paired blood and biopsy samples from cohort 2 (Fig. 4a, b) and found two signatures previously identified in cohort 1 blood samples consistent with tissue for all CD versus UC (Fig. 4c). The reduction of basophil frequency we observed in UC biopsies compared with CD suggests an overall reduction of basophils in UC, since this trend is consistent with the blood (Figs 1d, 4c). Moreover, tissue inflammation (identified by endoscopist and confirmed by blinded pathologist) appeared to decrease basophil frequency further in UC, but not in CD (Fig. 4c). Plasmablast frequency was significantly increased in UC tissue compared with CD, which was also observed in the blood when comparing the frequency of α4β7+ plasmablasts between CD and UC flare, suggesting an overall increase of plasmablast abundance in UC flare (Fig. 4c; Supplementary Fig. 3). Basophils and plasmablasts in blood and tissue were consistent with predicted cell frequencies for basophils and plasma cells from gene expression deconvolution of publicly available microarray data sets (Supplementary Fig. 8; plasmablasts were not identifiable via deconvolution, but our plasmablast gate included plasma cells).

Because B cells and plasmablasts were common among signatures identified, we analyzed CD3+ CD4+ CD45RO+ CXCR5- PD-1+ peripheral helper T (Tph) cells, which drive B cells and plasmablasts in autoimmunity. We found no Tph abundance differences in blood between CD and UC patients in cohort 2, but there were significantly higher frequencies in tissue compared with blood in both diseases (Fig. 4d). Although there was no difference in Tph cell frequency between CD and UC tissues, Tph cells had significantly higher median pSTAT3 expression in CD than UC tissues (Fig. 4d). Median pSTAT3 expression by Tph cells was not different between inflamed and uninfamed tissues, suggesting that higher pSTAT3 in this population may be a hallmark of CD regardless of inflammation state. In summary, Tph cell activation as indicated by pSTAT3 expression but not abundance was significantly higher in CD than UC tissue.

**Blood and tissue leukocyte correlations.** Due to expected differences in cell frequencies and median protein expression levels between blood and tissue, we investigated immune phenotype correlates. Since it was unlikely that marker expression or cell abundance would be equal between blood and tissue because of their distinct cellular environments, we studied paired blood and tissue samples from the ileum, colon, or rectum to identify blood signatures reflective of tissue. We found that 795 of 2145 targeted parameters were significantly different between blood and tissue for all patients in cohort 2. Of these 795 significant features, 55.35% were higher in blood and 44.65% were higher in tissue (Fig. 5a). Pearson correlation based on significant features
Tissue contains immune responses distinct from blood. a Schematic of the study conducted on cohort 2. Blood was drawn and biopsies were collected from study subjects, peripheral blood mononuclear cells (PBMCs) and tissue leukocytes were isolated and cryopreserved, and samples were analyzed in batches by CyTOF. Created with BioRender. b viSNE based on 15 core lineage antigens (CD11c, CD11b, CD56, CD16, CD8, CD3, CD123, CD27, CD24, CD14, CD19, CD4, CD20, TCyR6, and CD45R0) for samples from cohort 2. Sample sizes: CD blood = 6; UC blood = 6; CD inflamed tissue = 11; CD uninflamed tissue = 12; UC inflamed tissue = 5; UC uninflamed tissue = 13. c Significant differences between disease tissues consistent with trends observed in the blood from cohort 1. Statistics: unpaired two-tailed Student’s T test (Basophils: all CD vs. all UC tissue, t = 1.83, df = 39; UC inflamed vs. uninflamed tissue, t = 2.43, df = 16. Plasmablasts: all CD vs. all UC tissue, t = 2.51, df = 39). Sample sizes: CD blood = 6 (5 for α4β7+); UC blood = 6; all CD tissue = 23; UC inflamed tissue = 11; CD uninflamed tissue = 12; all UC tissue = 18; UC inflamed tissue = 5; UC uninflamed tissue = 13. Center lines = mean; whiskers = standard deviation. d T peripheral helper (Tph) cells, defined as CD3+CD4+CD45RO+CXCR5+PD-1+, in paired blood and tissue samples. Statistics: unpaired two-tailed Student’s T test (all blood vs. tissue, t = 2.65, df = 51; all CD vs. UC tissue, t = 2.25, df = 39; CD inflamed vs. uninflamed tissue, t = 0.75, df = 21). Sample sizes: all blood = 12; all tissue = 41; all CD tissue = 23; all UC tissue = 18; CD inflamed tissue = 11; CD uninflamed tissue = 12. Center lines = mean; whiskers = standard deviation. Source data are provided as a Source Data file.

Fig. 4 Tissue contains immune responses distinct from blood. a Schematic of the study conducted on cohort 2. Blood was drawn and biopsies were collected from study subjects, peripheral blood mononuclear cells (PBMCs) and tissue leukocytes were isolated and cryopreserved, and samples were analyzed in batches by CyTOF. Created with BioRender. b viSNE based on 15 core lineage antigens (CD11c, CD11b, CD56, CD16, CD8, CD3, CD123, CD27, CD24, CD14, CD19, CD4, CD20, TCyR6, and CD45R0) for samples from cohort 2. Sample sizes: CD blood = 6; UC blood = 6; CD inflamed tissue = 11; CD uninflamed tissue = 12; UC inflamed tissue = 5; UC uninflamed tissue = 13. c Significant differences between disease tissues consistent with trends observed in the blood from cohort 1. Statistics: unpaired two-tailed Student’s T test (Basophils: all CD vs. all UC tissue, t = 1.83, df = 39; UC inflamed vs. uninflamed tissue, t = 2.43, df = 16. Plasmablasts: all CD vs. all UC tissue, t = 2.51, df = 39). Sample sizes: CD blood = 6 (5 for α4β7+); UC blood = 6; all CD tissue = 23; CD inflamed tissue = 11; CD uninflamed tissue = 12; all UC tissue = 18; UC inflamed tissue = 5; UC uninflamed tissue = 13. Center lines = mean; whiskers = standard deviation. d T peripheral helper (Tph) cells, defined as CD3+CD4+CD45RO+CXCR5+PD-1+, in paired blood and tissue samples. Statistics: unpaired two-tailed Student’s T test (all blood vs. tissue, t = 2.65, df = 51; all CD vs. UC tissue, t = 2.25, df = 39; CD inflamed vs. uninflamed tissue, t = 0.75, df = 21). Sample sizes: all blood = 12; all tissue = 41; all CD tissue = 23; all UC tissue = 18; CD inflamed tissue = 11; CD uninflamed tissue = 12. Center lines = mean; whiskers = standard deviation. Source data are provided as a Source Data file.

Demonstrated relative homogeneity amongst blood samples compared with tissue diversity (Fig. 5a), suggesting distinct immune responses across regions of the gut and inflammation states. Since we did not have sufficient samples to separate uninflamed and inflamed tissues by region (ileum, colon, and rectum), we focused on correlations between blood and biopsies from each region regardless of inflammation due to expected variation by location. Since the heterogeneity of tissue and the breadth of parameters tested would preclude identification of significant correlates by directly comparing all pairs and correcting for all comparisons, we first compared subsets based on disease diagnosis, tissue location, and/or tissue inflammation state with adjustment for multiple comparisons. We combined this screen with biological insight from disease mechanisms to inform a refined set of hypothesis-driven tests without adjustment using all samples divided only by region. With this approach, paired blood and tissue samples revealed that CD3+ frequency of CD14+CD11c+HLA-DR+ cells, median pCREB expression by central memory CD4 T cells, and CCR9+ switched memory B cell frequency were correlated between blood and ileum (Fig. 5b). There were no significant correlations between blood and colon or rectum in cohort 2 amongst 2145 parameters tested, suggesting greater variance in samples from these regions. Surprisingly, all significant correlations were positive; inverse correlations between blood and tissue indicative of cellular trafficking were not found, possibly due to downregulation of trafficking molecules after homing to the tissue and/or the single collection timepoint and location of the samples analyzed here (biopsies represent a small area of the tissue). The limited number of blood and tissue pairs available for correlation analyses after stratifying by disease diagnosis, tissue location, and/or tissue inflammatory state as well as the breadth of parameters tested with stringent adjustment for multiple comparisons could account for the low number of parameters significantly correlated (Fig. 5b); in the future additional samples may reveal more subset-specific correlations. Upon closer inspection of blood/tissue correlates without correction for multiple testing, median pCREB expression by...


**Activated leukocytes distinguish tissue inflammation in CD.** We compared inflamed to uninflamed tissues from CD patients to explore what factors might contribute to the discontinuous nature of inflamed regions in CD unlike continuous patterns in UC. Increased frequencies of CD45RO+ B cell subsets and decreased frequencies of highly activated T cells, NK cells, and antigen-presenting B cells were associated with inflammation in CD tissues (Table 2c; Supplementary Fig. 5e). Since CD is characterized by heterogeneous, patchy disease, we evaluated four pairs of uninflamed and inflamed tissues from similar locations in CD by hypothesis-driven comparisons. We identified four

---

**Fig. 5** Mapping paired samples reveals blood correlates of tissue immunity. **a** Of 2145 manually gated cell frequencies and median expression levels, 795 were significantly different (adjusted \( p \leq 0.05; \) data not shown) between blood (\( N = 12 \)) and tissue (\( N = 41 \)). Of significantly different parameters, 55.35% were higher in blood (red) and 44.65% were higher in the tissue (purple). Significantly different parameters were used to construct the Pearson correlation map. Statistics: BH FDR-corrected unpaired two-tailed Student’s \( t \) test using Morpheus (see the Methods section). **b** Of 2145 parameters, three were correlated (green box and green wedge in pie chart) between blood and ileum tissue samples. Statistics: Pearson correlation coefficient (\( r \)) and \( p \)-values from Pearson correlation tests are shown. \( p \)-values for these three correlations were not corrected for multiple testing because hypothesis-driven tests were conducted based on biological insight after independent preliminary analysis. Sample sizes: ten blood/tissue pairs for CD3+ cells and central memory CD4 T cell median pCREB expression; nine blood/tissue pairs for CCR9+ switched memory B cells. Solid line = linear regression; dotted lines = 95% confidence interval. Source data are provided as a Source Data file.
significant differences between paired inflamed and uninflamed areas of the same anatomical region (i.e., ileum or colon) (Fig. 6b). There were higher frequencies of CD45RO+ memory CD4 T cells in paired inflamed tissues, which suggests chronic antigen-driven responses. Interestingly, there was greater median expression of pSTAT3 in CD27+CD45RO− CD4+ T cells and higher frequencies of activated CD43+ T cells in uninflamed tissues, which could reflect subclinical disease in uninflamed tissues, immune exhaustion and senescence in inflamed regions, and/or an anti-inflammatory effect of these cells. Consistent with the above potential anti- and pro-inflammatory roles of STAT3, we found greater median pSTAT3 expression by regulatory phenotype CD25+B cells in uninflamed tissues, which may indicate a protective role for these cells in CD. Overall, activated and antigen-driven T and B cell responses as well as pSTAT3 expression and NK cell expansion were associated with tissue inflammation in CD, which is consistent with trends identified in blood. However, we found no differences between inflamed and uninflamed UC tissues, highlighting distinctions between CD and UC tissues. Alternatively, these findings might be due to limited tissue sample sizes and the relatively mild disease of UC patients in cohort 2.

Blood-based signatures classify IBD patients. We utilized differences identified between total CD and UC blood for classification of samples by disease as a potential non-invasive blood-based diagnostic to reduce the required frequency of endoscopy (Fig. 7a). Unbiased approaches for biomarker discovery, such as CITRUS analysis using SAM, PAMR, or LASSO/GLMNET algorithms, did not produce robust results because dividing cohorts into cross-validation folds reduced disease group sample sizes. We constructed generalized linear models (GLMs) using blood signatures we identified throughout the study to classify patients into one of two disease groups (CD versus UC). Using eight blood features significantly different between all CD and UC samples in cohort 1 (Fig. 1c, d), we created a GLM to classify patients as CD or UC (Fig. 7b). Discriminatory performance of the GLM was assessed using receiver-operating characteristic (ROC) analysis, in which the true-positive rate (sensitivity) is plotted versus the false-positive rate (1-specificity). The area under the curve (AUC) reflects the probability that the model will rank a randomly chosen positive (CD) sample higher than a randomly chosen negative (UC) sample. This approach revealed an AUC of 0.845 (95% CI, 0.742–0.948) when using the model with all eight signatures combined to discriminate CD from UC amongst all blood samples in the study, which was higher than any single-signature model (Fig. 7b; Supplementary Table 4). A cutoff of 0.4874743 for the model identified with the Youden index method to maximize sensitivity and specificity (see Methods) yielded a sensitivity of 0.80 and a specificity of 0.85 (PPV = 0.86, NPV = 0.79).

We performed multivariate analysis to determine if subject characteristics confounded the eight parameters utilized for classification of CD and UC. We used analysis of covariance (ANCOVA) to assess the interaction between age, sex, or age and sex with each of the eight parameters (Supplementary Table 5). There were no significant interactions for cohort 1 or 2. Thus, it was unnecessary to adjust the data based on age or sex, and we did not include these in the classification model.

CD and UC patients are treated with different medications, often as combination therapies. Most subjects in this study were on multiple medications with distinct mechanisms of action. We sought to determine whether or not medications confounded the eight parameters utilized for disease classification. There was no standard method to correct for medications as confounding factors in the context of our sample sizes or the numerous combination and single-therapy regimens. We stratified all subjects on versus not on each therapy (Supplementary Fig. 7a).
and reevaluated the performance of the same classification model built using all samples. The most common medications were TNF antagonists, glucocorticoids, 5-aminosalicylates (5-ASA), and 6-mercaptopurines (6-MP). Methotrexate was used for a small number of subjects as a therapy for preventing immunogenicity of biologic therapy. IL-12/23 antagonists, glucocorticoids, 5-aminosalicylates (5-ASA), and 6-mercaptopurines (6-MP) were used for three and four subjects, respectively. We used the same GLM classification model. This is ideal for evaluating patients on corresponding medications and reevaluated the performance of the same classification model.

### Discussion

Animal models of IBD provide insight into disease pathogenesis and opportunities to study therapeutic leads, yet very few successfully treat the human disease and there remains a dramatic unmet clinical need for long-lasting treatments. Many cytokine-based therapies extrapolated from experimental models have failed in clinical trials. Translation of basic research into clinical practice is also hampered by the heterogeneity of human disease. Although the gut is accessible, endoscopy is invasive and has risks. We sought to address these challenges and improve future studies by identifying blood signatures that objectively differentiate disease type, state, and behavior; represent local immune responses; and classify patient subsets. We performed CyTOF with lineage, activation, and trafficking markers. This allowed us to identify cellular signatures and phenotypes associated with IBD subsets, highlighting gut tropic leukocytes. It will also be interesting for other investigators to apply this approach to other tissues with corresponding trafficking molecules.

We utilized a diverse cohort of IBD and UCs and reported findings that relate peripheral and tissue immune responses across distinct disease states. Prior knowledge of leukocyte trafficking in the field is largely from animals; our study demonstrates significance and opportunities for trafficking molecules and gut tropic cells. Decreases in gut tropic cells we observed in UC compared with remission and HC might indicate an increase in trafficking or tissue localization, which could explain why UC patients often respond better to the anti-α4β7 vedolizumab.

### Table 2 Tissue leukocytes that distinguish diseases and inflammation states reveal shared signatures

| a Cell subset | Metric | Higher in | p-value | t-statistic |
|---------------|--------|-----------|---------|-------------|
| CD14+ cells | % of non-basophils | CD tissue | 0.0138 | 3.16 |
| CD14+ DCs | % of DCs | CD tissue | 0.035 | 4.24 |
| CD14-CD11c+HLA-DR+CD19+ cells | Median pSTAT1 | CD tissue | 0.0444 | 3.07 |
| CCR1+CCR9+ DCs | % of DCs | CD tissue | 0.033 | 3.34 |
| CCR1+GPR15+ DCs | % of DCs | CD tissue | 0.0477 | 3.65 |
| CD38+ NK cells | Median pCREB | CD tissue | 0.0214 | 4.29 |
| CD25+ NK cells | Median pSTAT3 | CD tissue | 0.0138 | 3.27 |
| Effector memory CD4 T cells | Median pSTAT3 | CD tissue | 0.0286 | 3.38 |
| Effector memory CD8 T cells | Median pSTAT3 | CD tissue | 0.0444 | 3.01 |
| α4β7+GPR15+ effector memory CD8 T cells | % of effector memory CD8 T cells | UC tissue | 0.0214 | -4.43 |
| IgD-CD27+ B cells | Median CD25 | CD tissue | 0.0421 | 3.32 |
| CD38+ switched memory B cells | % of switched memory B cells | UC tissue | 0.0214 | -4.22 |
| CXCR5+ switched memory B cells | % of switched memory B cells | CD tissue | 0.0214 | 2.51 |
| CCR1+ IgD+CXCR5+ B2 cells | % of IgD+CXCR5+ B2 cells | CD tissue | 0.0477 | 3.23 |
| Plasmablasts | Median pSTAT3 | CD tissue | 0.0138 | 3.49 |
| CCR9+GPR15+ IgD+ plasmablasts | % of IgD+ plasmablasts | UC tissue | 0.0214 | -2.7 |

| b Cell subset | Metric | Higher in | p-value | t-statistic |
|----------------|--------|-----------|---------|-------------|
| DCs | Median pSTAT3 | CD inflamed tissue | 0.0394 | 3.4 |
| α4β7+CCR1+GPR15+ DCs | % of α4β7+ DCs | CD inflamed tissue | 0.0106 | 3.18 |
| CD25+ NK cells | Median CCR9 | CD inflamed tissue | 0.0106 | 3.19 |
| Naive CD4 T cells | Median pSTAT3 | CD inflamed tissue | 0.028 | 3.41 |
| Effector CD4 T cells | Median pSTAT3 | CD inflamed tissue | 0.0214 | 2.78 |
| CXCR5+PD-1+ CD45RO+CD4 T cells | % of CD45RO+CD4 T cells | CD inflamed tissue | 0.0155 | 4.96 |
| GPR15+PD-1+ B cells | % of PD-1+ B cells | CD inflamed tissue | 0.0482 | 2.78 |
| CCR9+CD45RO+ B cells | % of CD45RO+ B cells | CD inflamed tissue | 0.0056 | 6.67 |
| GPR15+CD45RO+ B cells | % of CD45RO+ B cells | CD inflamed tissue | 0.0056 | 11.2 |
| CCR9+GPR15+CD45RO+ B cells | % of CD45RO+ B cells | CD inflamed tissue | 0.0056 | 8.34 |
| CD25+CD19+CD20+B cells | Median pSTAT3 | CD inflamed tissue | 0.0238 | 3.27 |
| Switched memory B cells | Median pSTAT3 | CD inflamed tissue | 0.0319 | 3.7 |
| IgD−IgA−CD19+CD20+B cells | Median pSTAT3 | CD inflamed tissue | 0.0153 | 3.61 |
| IgD−CD27+ B cells | Median pSTAT3 | CD inflamed tissue | 0.0195 | 3.45 |
| IgD−CD27+ B cells | Median CD25 | CD inflamed tissue | 0.0195 | 3.66 |

| c Cell subset | Metric | Higher in | p-value | t-statistic |
|----------------|--------|-----------|---------|-------------|
| α4β7+CCR9+PD-1+ NK cells | % of PD-1+ NK cells | CD uninfamed tissue | 0.0165 | -1.46 |
| α4β7+CCR1+GPR15+CD38+HLA-DR+CD8 T cells | % of α4β7+CD38+HLA-DR+CD8 T cells | CD uninfamed tissue | 0.0165 | 1.79 |
| HLA-DR+ B cells | Median pSTAT1 | CD uninfamed tissue | 0.0172 | -1.71 |
| CCR9+ transitional B cells | % of transitional B cells | CD uninfamed tissue | 0.0165 | 2.15 |
| CD45RO+ B cells | CD tissue | CD uninfamed tissue | 0.0172 | 2.47 |

Features that distinguish all CD and UC tissue (A; see Supplementary Fig. 5c for plots and sample numbers), inflamed CD and UC tissue (B; see Supplementary Fig. 5d for plots and sample numbers), and inflamed and uninfamed CD tissue (C; see Supplementary Fig. 5e for plots and sample numbers). Statistics: BH FDR-corrected unpaired two-tailed Student’s t test using Morpheus (see Methods). Source data are provided as a Source Data file.
Furthermore, our flow-cytometric validation of CyTOF data demonstrates the potential for others to apply these results. Our findings are consistent with distinct responses in circulating and tissue-based CD vs. UC based on blood. Blood was drawn from study subjects, peripheral blood mononuclear cells (PBMCs) were isolated and cryopreserved, and samples were analyzed in batches by CyTOF. Created with BioRender. Generalized linear models (GLMs) were created for eight parameters significantly different between all CD and UC samples. Corresponding receiver-operating characteristic (ROC) curves are shown for single feature and eight feature models. All CD and UC samples were used to plot ROC curves. UC was used as baseline for the purposes of the GLMs, such that a true-positive indicates correct classification of a CD sample. Statistics: generalized linear models were constructed using glm in R (see the Methods section). Intercepts and parameter coefficients for each model are provided in Supplementary Table 4. Cutoffs and associated performance characteristics are discussed in the text. Source data are provided as a Source Data file.

Methods

Patient samples. All blood and tissue samples were collected using a study protocol approved by the Stanford Institutional Review Board. Informed consent was obtained from subjects aged 18–75 with an IBD-specializing gastroenterologist-confirmed diagnosis of inflammatory bowel disease (except healthy controls), excluding those who were pregnant, had other autoimmune or inflammatory diseases (except for extraintestinal manifestations of IBD), had a malignancy, had an active infection at the time of enrollment, had undergone surgery within 1 month of enrollment, had a blood transfusion within 1 month of enrollment, had received an organ or bone marrow transplant, or were unable to provide informed consent. All clinical data for subjects were current at the time of sample collection.

of disease diagnosis (CD vs. UC), state (flare vs. remission), and phenotypes (behavior and location). The prevalence of elevated pSTAT3 in CD suggests pSTAT3 as a marker of CD in subjects studied here. Elevated pSTAT3 in CD tissue Tph cells is consistent with increased B cell subsets. STAT3 also signals in the IL-23/Th17 pathway, which was identified in IBD GWAS studies and is a therapeutic target. Expression of pSTAT3 in CD25+ B cells was higher in uninfamed than paired inflamed CD tissue of the same region, consistent with the regulatory effects of these cells. Regulatory B cells are associated with IL-10 and TGF-β, both upstream of pSTAT3. Thus, observed pSTAT3 levels might indicate a protective role for CD25+ B cells in CD.

Gut tropic CD45RO+ B cell subsets were elevated in inflamed CD tissue compared with UC. These cells were previously described as a biomarker correlated with CD activity index and intestinal permeability. B cell expression of CD45RO is associated with late-stage differentiation and antigen stimulation, activation, and IgVH mutation, suggesting antigen-driven B cell responses in CD tissue. Further elucidating B cell subsets in the pathophysiology of IBD represents a significant advance in the field.

Our analysis has several limitations. First, some subclinical active disease processes may be ongoing in patients in clinical remission. We used IBD specialist-defined flare and remission for patients based on clinical criteria utilized for standard of care with intent to treat at the time of sample collection, although there was a spectrum of quantitative disease activities and phenotypes for each group. Some cohort 2 subjects in clinical remission exhibited histologic features of mild disease activity and might not fit deep remission states. That we identified distinct signatures between CD and UC patients even in remission suggests that while flares of either disease might respond to similar therapies as evidenced by their clinical utility, different maintenance therapies might be necessary for subsets of patients. Second, comparison with tissue samples from healthy subjects is lacking. Third, our findings were guided by our CyTOF panels, and can only suggest cell localization based on knowledge of trafficking largely from mouse studies. Short-term homing assays where labeled cells are transferred are currently not feasible in human.

Consistent with clinical observations, we found greater cellular heterogeneity among CD compared with UC. Blood signatures could be developed as non-invasive, cost-effective, and safe markers for disease diagnosis and monitoring. Although our classification model could be over specified and there exist cellular features not tested here, these limitations do not interfere with its potential utility. Future validation using larger independent cohort(s) will be necessary.

Since the vast majority of patients in academic and/or tertiary centers are not treatment naive, the breadth of drug regimens represented here reflects the reality encountered. Our future studies will address limitations of this work and focus on evaluating disease classification models in larger and more diverse independent cohorts, include HC samples and additional disease comparators, and track individuals longitudinally to examine phenotype stability over time.
Flare was determined by IBD-specializing physicians for standard of care and intent to treat using all available information, including bloodwork, clinical assessments, patient history, and/or endoscopy. When available (for most patients), recent endoscopy reports were reviewed to confirm disease state. Disease activity was defined using HBI (CD) and Partial Mayo (UC) scores.

Blood samples were collected by standard of care venipuncture; three vacutainers with sodium heparin (BD, cat. #366480) were filled with blood and kept at room temperature until processing, which occurred within 2 h. Tissue samples were collected during standard of care endoscopic procedures (within 2 h of paired blood sample collection) using biopsy forceps rinsed in sterile saline to remove any residual formalin. Two to four tissue biopsies per site were collected and deposited in 3 mL of sterile PBS without calcium or magnesium in a 5-ml Eppendorf tube and kept at room temperature if processed within 30 min or kept on ice if processed within 2 h. Biopsies were excluded if they were sampled from sites exposed to methylene blue during endoscopic procedures. Samples from inflamed and uninfamed tissues were initially identified by the endoscopist and subsequently confirmed by a blinded pathologist.

**Blood leukocyte isolation and cryopreservation.** Blood was centrifuged in vacutainers for collection at 2000 rpm for 10 min. Plasma was aspirated from the top and frozen at 20 °C according to the manufacturer’s instructions, stored at –80 °C for 1–7 days, and then transferred to liquid nitrogen for storage.

**Tissue leukocyte isolation and cryopreservation.** Combined biopsy bite tissue samples from the same site were washed in HBSS without calcium or magnesium supplemented with 2% BSA and then transferred to 5 mL of RPMI with HEPES, 5% BSA, collagenase IV at 0.7 mg/mL (Sigma-Aldrich, cat. #C5138), and quantified. Tissue samples with less than 1 × 10⁶ live cells recovered, thawed mouse splenocytes added to liquid nitrogen for storage.

**Flow cytometry.** Flow cytometry was conducted at the Stanford Shared FACS Facility (SSF) on a BD LSRII instrument in accordance with standard methods. In brief, cells were thawed, washed once in complete RPMI, incubated in 1 mL of complete RPMI with 2.5 mM MgCl₂, (Thermo Fisher, cat. #A195330G and 0.5 mg/mL 15% 10X PBS for use in FlowPath (Creative BioMolecules, cat. #L5002060) for 10 min at room temperature, and washed in 10 mL of complete RPMI. Cells were counted and 1 × 10⁶ or 2 × 10⁶ live cells were aliquoted for staining. Cells were washed with FACS buffer (HBSS without calcium or magnesium and supplemented with 2% BSA). Cells were resuspended in 100 µL of 1:500 Zombie Green fixable viability stain (BioLegend, cat. #423111) in PBS without calcium or magnesium and incubated for 15 min at room temperature, and then washed in the dark. Cells were washed in FACS buffer, resuspended in a master mix of fluorochrome-conjugated antibodies (Supplementary Fig. 4b) using the supplier-recommended 5 µL of each antibody per 1×10⁶ cells, and incubated for 30 min at 4°C in the dark. Cells were washed with FACS buffer, resuspended in 100 µL of FACS buffer, and kept on ice in the dark until sample analysis. Approximatley 1×10⁶ or all possible events were acquired for each sample. For single-color compensation controls, one drop of negative control and one drop of anti-mouse compensation beads (BD, cat. #552843) were incubated in 100 µL of FACS buffer and 5 µL of antibody for 15 min at room temperature in the dark and then kept on ice in the dark until analysis.

**Data analysis.** Bead normalized sample files were obtained from the Helios instrument using on-board software. FlowJo was used for cleaning up files, correcting for compensation, and setting compensation controls. Samples were run in parallel using a random seed, 1000 iterations, perplexity of 30, and theta of 0.5. Combined data were concatenated by group. In total, 5000 events were randomly subsampled from individual samples by group. In total, 5000 events were randomly subsampled from each concatenated file, and clustering was run on all concatenated files in parallel using a random seed, 1000 iterations, perplexity of 30, and theta of 0.5. GraphPad PRISM 7 was used to plot some figures and conduct some targeted statistical tests, and Microsoft Excel was used for some basic data maneuvering. MorphoSys was utilized for constructing heatmaps and Pearson correlation maps based on significant parameters selected using t-tests with 10,000 permutations and Benjamini and Hochberg (BH) FDR correction for multiple testing (documentation available at: https://www.morphoSys.com/morpheus/software/documentation.html). R was used for additional statistical analyses (see below). There were three patients in cohort 1 and one patient in cohort 2 excluded from analyses involving trafficking receptors because they were treated with an a4β7 antagonist.

**Statistics.** All p-values were derived from two-tailed unpaired T tests without adjustment for multiple testing unless otherwise noted. Most p-values were corrected for multiple testing as indicated using the BH FDR correction method in R (see above). When correcting for multiple testing manually, the p.adjust function with the BH correction method (also known as FDR) was used in R according to its documentation to account for the expected proportion of false discoveries among the rejected hypotheses. Correlation tests were based on Pearson correlations using R according to its documentation. For paired blood and tissue correlations, each biopsy location was paired with blood from the same individual, and Pearson correlations were calculated for each pair. The mean value

---

**ARTICLE**

NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-019-10387-7
for tissue was used if there was more than one tissue sample from a given location. P-values varied with the number of samples in each subgroup being compared, as p-adjust incorporates the number of comparisons being made. Correlated parameters were eliminated if driven by an outlier or less than three data points. For receiver-operating characteristic (ROC) analysis, generalized linear models (GLM) were constructed in R using the glm package according to its documentation using features significantly different between disease groups. GLMs were calculated for each individual parameter and then compared with a model including all significant parameters. For each ROC curve, the area under the curve and 95% confidence intervals were calculated using the ROCR package in R according to its documentation.

**Gene expression deconvolution.** Gene expression deconvolution analyses were conducted according to published methods using the immuneStates basis matrix for major cell lineages. Only major cell lineages were identifiable using this approach due to the availability of sorted cell gene expression data used in the basis matrix. In brief, we measured the mean expression of each gene in the 312-gene matrix for major cell lineages and performed deconvolution with support vector regression using the CIBERSORT algorithm. We then compared the resulting estimated cell-type frequencies between classes and subsequently calculated the effect sizes. We downloaded all data sets from Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo/) using the MetaIntegrator package from CRAN.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability.** The data sets generated during the current study are available from the corresponding author on reasonable request that does not include confidential patient information. The source data underlying Figs 1b–d, 2a–d, 3a–g, 4c, d, 5b, 6b, 7b, and Table 2a–c, as well as Supplementary Figs 3a, b, 4c, f, 5a–e, 6a–c, 7b–e and 8a, b are provided as a Source Data file.

**Code availability.** The computer code generated and used during this study is available from the corresponding author on reasonable request that does not include confidential patient information. Web-based tools utilized in this study are freely accessible for the foreseeable future (https://software.broadinstitute.org/morpheus/).

Received: 5 October 2018 Accepted: 29 April 2019

**References**

1. Molodecky, N. A. et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology 142, 46–54 (2012).
2. Abraham, C. & Cho, J. H. Inflammatory bowel disease. N. Engl. J. Med 361, 2066–2078 (2009).
3. Bilsho, J., Targan, S. R. & Snapper, S. B. Therapeutic Targets in Inflammatory Bowel Disease: Current and Future. Am. J. Gastroenterol. Suppl. 3, 27–37 (2016).
4. Ananthakrishnan, A. N. Epidemiology and risk factors for IBD. Nat. Rev. Gastroenterol. Hepatol. 12, 205–217 (2015).
5. Marsal, J. & Agace, W. W. Targeting T-cell migration in inflammatory bowel disease. J. Intern Med 272, 411–429 (2012).
6. Gasch, C. et al. A simple classification of Crohn’s disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998. Inflamm. Bowel Dis. 6, 8–15 (2000).
7. Satsangi, J., Silverberg, M. S., Vermeire, S. & Colombel, J. F. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. Gut 55, 749–753 (2006).
8. 10. Yantiss, R. K. & Odze, R. D. Diagnostic difficulties in inflammatory bowel disease pathology. Histopathology 48, 116–132 (2006).
9. McGovern, V. J. & Gaulston, S. J. Crohn’s disease of the colon. Gut 9, 164–176 (1968).
10. Islam, S. A. & Luster, A. D. T cell homing to epithelial barriers in allergic disease. Nat. Med 18, 705–715 (2012).
11. Habtezion, A., Nguyen, L. P., Hadeiba, H. & Butcher, E. C. Leukocyte Trafficking to the Small Intestine and Colon. Gastroenterology 150, 340–354 (2016).
12. Bendall, S. C., Nolan, G. P., Roederer, M. & Chattopadhyay, P. K. A deep profiler’s guide to cytometry. Trends Immunol. 33, 323–332 (2012).
13. Li, J. et al. Measuring gene expression across multiple human primary tissues. Bioinformatics 31, 1102–1108 (2015).
14. Song, J. et al. Enhancing the accuracy of gene expression deconvolution. Bioinformatics 32, 306–312 (2016).
15. Li, J. et al. Improved gene expression deconvolution analysis using a weighted linear model. Bioinformatics 32, 1287–1294 (2016).
16. McGovern, D. P., Kugathasan, S. & Cho, J. H. Genetics of inflammatory bowel diseases. Gastroenterology 149, 1163–1176 e1162 (2015).
17. Kim, S. V. et al. GPR15-mediated homing controls immune homeostasis in the large intestine mucosa. Science 340, 1456–1459 (2013).
18. Ananthakrishnan, A. N. Epidemiology and risk factors for IBD. Gut 63, 1263–1268 (2017).
19. Morrison, R. W. & De Jong, K. A. Measurement of population diversity. Front. Ecol. Evol. 31, 31–41 (2012).
20. Muroi, K. et al. CD56 expression in B-cell lymphoma. Leuk. Res 22, 201–202 (1998).
21. Vanacker, H. H., Capsomidis, A., Smits, E. L. & Van Tendeloo, V. F. CD56 in the immune system: more than a marker for cytotoxicity? Front Immunol. 8, 892 (2017).
22. Brand, S. Crohn’s disease: Th1, Th17 or both? The change of a paradigm: new immunological and genetic insights implicate Th17 cells in the pathogenesis of Crohn’s disease. Gut 58, 1152–1167 (2009).
23. Vallania, F. et al. Leveraging heterogeneity across multiple data sets increases accuracy of cell-mixture deconvolution and reduces biological and technical biases. Nat. Commun. 9, 4735 (2018).
24. Qiao, Y. Q. et al. LncRNA DQ986243 affects Treg related CREB and Foxp3 expression in Crohn’s disease. J. Biomed. Sci. 20, 87 (2013).
25. Wu, Y. C., Kipling, D. & Dunn-Walters, D. K. The relationship between CD27 negative and positive B cell populations in human peripheral blood. Front Immunol. 2, 81 (2011).
26. Fecteau, J. F., Cote, G. & Neron, S. A new memory CD27-IgG+ B cell population in peripheral blood expressing VH genes with low frequency of somatic mutation. J. Immunol. 177, 3728–3736 (2006).
27. Fleisher, S. J. et al. Increased frequency of a unique spleen tyrosine kinase bright memory B cell population in systemic lupus erythematosus. Arthritis Rheum. 66, 3424–3435 (2014).
28. Wei, C. et al. A new population of cells lacking expression of CD27 represents a notable component of the B cell memory compartment in systemic lupus erythematosus. J. Immunol. 178, 6624–6633 (2007).
29. Moir, S. et al. Evidence for HIV-associated B cell exhaustion in a dysfunctional memory B cell compartment in HIV-infected viremic individuals. J. Exp. Med. 205, 1797–1805 (2008).
30. Rojas, O. L., Narvaez, C. F., Greenberg, H. B., Angel, J. & Franco, M. A. Characterization of rotavirus specific B cells and their relation with serological memory. Virology 380, 234–242 (2008).
31. Mills, S. & Stamos, M. J. Colonic Crohn’s disease. Clin. Colon Rectal Surg. 20, 309–313 (2007).
32. Frizelle, F. A. & Burt, M. J. Review: the surgical management of ulcerative colitis. J. Gastroenterol. Hepatol. 12, 670–677 (1997).
33. Hidrick, T. L. & Friel, C. M. Colonic crohn disease. Clin. Colon Rectal Surg. 26, 84–89 (2013).
34. Rao, D. A. et al. Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis. Nature 542, 110–114 (2017).
35. Wang, X. et al. Cyclic AMP-responsive element-binding protein (CREB) is critical in autoimmunity by promoting Th17 but inhibiting Treg cell differentiation. Elife 25, 165–174 (2017).
36. Yew, J. F., De Souza, A. C. & Eaves, C. J. CD4+ T helper cell differentiation and inflammatory diseases. Cytokine 47, 149–156 (2009).
37. Galvez, J. Role of Th17 cells in the pathogenesis of human IBD. ISRN Inflamm. 2014, 928461 (2014).
38. Jäimes, M. C. et al. Maturation and trafficking markers on rotavirus-specific B cells during acute infection and convalescence in children. J. Virol. 78, 10967–10976 (2004).
39. Wang, W. et al. Identification of a novel chemokine (CCL28), which binds CCRI (GPR15). J. Biol. Chem. 275, 22313–22323 (2000).
40. Trivedi, P. J. et al. Intestinal CCL25 expression is increased in colitis and correlates with inflammatory activity. J. Autoimmun. 68, 98–104 (2016).
41. Bekker, P. et al. CCR9 antagonists in the treatment of ulcerative colitis. Mediat. Inflamm. 2015, 628340 (2015).
42. Yacyshyn, B. R. & Meddings, J. B. CD45RO expression on circulating CD19+B cells in Crohn's disease correlates with intestinal permeability. Gastroenterology 108, 132–137 (1995).
43. Yacyshyn, B. R. & Pilarski, L. M. Expression of CD45RO on circulating CD19+B cells in Crohn's disease. Gut 34, 1698–1704 (1993).
44. Nguyen, P. M., Putoczki, T. L. & Ernst, M. STAT3-activating cytokines: a therapeutic opportunity for inflammatory bowel disease? J. Interferon Cytokine Res 35, 340–350 (2015).
45. Pickert, G. et al. STAT3 links IL-22 signaling in intestinal epithelial cells to mucosal wound healing. J. Exp. Med 206, 1465–1472 (2009).
46. Reinsch, W. et al. An interleukin 13 monoclonal antibody, in active UC: efficacy and safety from a phase IIa randomised placebo-controlled, double-blind, placebo-controlled trial. Gut 64, 894–900 (2015).
47. Danese, S. et al. Tralokinumab for moderate-to-severe UC: a randomised, double-blind, placebo-controlled, phase IIa study. Gut 64, 243–249 (2015).
48. Hueber, W. et al. Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn’s disease: unexpected results of a randomised, double-blind placebo-controlled trial. Gut 61, 1693–1700 (2012).
49. Feagan, B. G. et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. N. Engl. J. Med 369, 699–710 (2013).
50. Sandborn, W. J. et al. Vedolizumab as induction and maintenance therapy for Crohn’s disease. N. Engl. J. Med 369, 711–721 (2013).
51. van Uden, V. et al. Mass cytometry of the human mucosal immune system identifies tissue- and disease-associated immune subsets. Immunity 44, 1227–1239 (2016).
52. Lees, C. W., Barrett, J. C., Parkes, M. & Satsangi, J. New IBD genetics: common pathways with other diseases. Gut 60, 1739–1753 (2011).
53. Kessel, A. et al. Human CD19(+)CD25(high) B regulatory cells suppress proliferation of CD4(+) T cells and enhance Foxp3 and CTLA-4 expression in T-regulatory cells. Autoimmun. Rev. 11, 670–677 (2012).
54. Vadazs, Z., Haj, T., Kessel, A. & Toubi, E. B-regulatory cells in autoimmunity and immune mediated inflammation. FEBS Lett. 587, 2074–2078 (2013).
55. Bing, X., Linlang, L. & Keyan, C. Decreased Breg/Th17 Ratio Improved the Prognosis of Patients with Ulcerative Colitis. Can. J. Gastroenterol. Hepatol. 2018, 5760849 (2018).
56. Jackson, S. M. et al. CD45RO enriches for activated, highly mutated human germinal center B cells. Blood 110, 3917–3925 (2007).
57. Fernandez, R. & Maeker, H. Cytokine-stimulated phosphorylation of PBMC Using CyTOF mass cytometry. Bio Protoc 5, e1496 (2015).
58. Samsuk, N., Good, Z., Spitzer, M. H., Davis, K. L. & Nolan, G. P. Automated expression profiling of phenotype space with single-cell data. Nat. Methods 13, 493–496 (2016).
59. Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. Nat. Methods 12, 453–457 (2015).
60. Haynes, W. A. et al. Empowering multi- cohort gene expression analysis to increase reproducibility. PLoS. Symp. Biocomput 22, 144–153 (2017).

Acknowledgements
We thank Y. Yang, Y. Wei, A. Fan, and A.R. Ji for technical assistance, S. Toppo and M. Adelus for clinical assistance with patient recruitment, N. Sigal and H. Maeker for their mass cytometry expertise, A. McMillan for expert biostatistics consultation and guidance, L. Stell for general statistical advice, and the patients and healthcare providers without whom this study would not be possible. We also thank the Stanford Human Immune Monitoring Center and the National Institutes of Health S10 Shared Instrumentation Grant (S10RR027431-01) for access to mass cytometers and expert technical services. This work was supported by the National Institutes of Health (R01 DK101119) and the Ann and Bill Swindells Charitable Trust as well as Leslie and Douglas Ballinger. S.J.S. Rubin thanks W.D. Rubin for invaluable discussions and inspiration, W.H. Robinson for mentorship and guidance, as well as the National Science Foundation Graduate Research Fellowship and the Stanford Graduate Fellowship for their generous support. This material is based in part upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE—1656518. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: Anti-human a4β7 integrin (Act-1) (cat#11718) from Dr. A.A. Ansari.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-10387-7.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Peer review information: Nature communications would like to thank anonymous reviewers for their contribution to the peer review of this work.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2019