



ARTICLE

Expanded TCR β CDR3 clonotypes distinguish Crohn's disease and ulcerative colitis patientsJ. Wu¹, A. H. Pendegraft², M. Byrne-Steele³, Q. Yang³, C. Wang³, W. Pan³, T. Lucious¹, T. Seay¹, X. Cui⁴, C. O. Elson¹, J. Han³ and P. J. Mannon¹

We aimed to determine whether the TCR repertoires of Crohn's disease (CD) patients contain highly prevalent disease-specific T-cell clonotypes reflective of the characteristic and highly shared aberrant serum antibody reactivity to gut commensal flagellin antigens. The CD4 TCR β CDR3 sequence repertoires from active CD ($n = 20$) and ulcerative colitis (UC) ($n = 10$) patients were significantly more diverse, and individual sequences over-represented, compared to healthy controls (HC) ($n = 97$). While a very small number of expanded public CDR3 sequences are highly shared between active CD and UC, the majority of significantly expanded TCR β CDR3 clonotypes are private to CD and UC patients with equivalent prevalence among IBD patients. Further defining TCR clonotypes by V β -CDR3 linkage showed significant differences in the TCR repertoires between UC and CD. Flagellin antigen exposure induced expansion of several TCR β CDR3 sequences in CD4 cells from a flagellin-seropositive subject including sequences highly shared by or relatively private to CD (and UC) patients. These data suggest that flagellin-reactivity contributes to the expansion of a small number of CD4 clonotypes but does not support flagellin antigens as predominantly driving CD4 cell proliferation in CD. Disease-specific expanded TCR β CDR3 clonotypes characterize CD and UC and the shared exposure to the gamut of gut microbial antigens.

Mucosal Immunology (2018) 11:1487–1495; <https://doi.org/10.1038/s41385-018-0046-z>

INTRODUCTION

Inflammatory bowel disease (IBD) results from excessive activation of the mucosal immune system by gut microbes, an outcome enabled by a dysregulated immune response or a defective epithelial barrier. Although risk of Crohn's disease is associated with major genomic loci linked to innate immune function, the activated adaptive immune response also contributes to the pathologic effects of the inflammation. The importance of the adaptive immune response in IBD is demonstrated by the requirement for microbial antigen-reactive T cells to induce experimental colitis in animal models of Crohn's disease.¹ In these murine models, T-cell-mediated damage occurs only in the presence of established gut microbiota.^{2,3} In Crohn's disease, gut microbes induce both aberrant B- and T-cell reactivity to microbial antigens. B-cell immune memory in IBD, demonstrated by seropositivity for certain microbial antigens, is not only characteristically shared by Crohn's disease patients but it even predicts the clinical course of disease.^{4,5} T cells from Crohn's disease patients can display inflammatory reactivity to flagellin and other microbial antigens.^{6,7} Since B-cell immunoglobulin class switching requires interaction with CD4 T cells reactive to epitopes of the same antigen, we sought to determine whether Crohn's disease patients also shared characteristic T-cell clonotypes that reflect this shared antibody reactivity to the same microbial antigens.^{8,9} The identification of highly shared disease-associated T-cell clonotypes that are expanded and activated in Crohn's disease by a small set of antigens could help assess their status as high-value targets for novel diagnostic and therapeutic strategies.

Antigen-specific T-cell reactivity is largely conferred by the highly variable complementarity-determining region 3 (CDR3) of the T-cell receptor (TCR), composed of α and β chains, that interacts with the antigen peptide-MHC complex. Sequence analysis of the TCR β chain CDR3 region shows that there are $0.5\text{--}3 \times 10^6$ unique TCRs per person and perhaps up to 20 million overall.^{10–12} While there are identical TCR β CDR3 sequences that are highly shared among a healthy population, so-called “public” sequences, the potential for numerous TCR α pairings, V and J region use, and the different HLA backgrounds among individuals challenge the assumption that these shared TCR β sequences will confer equivalent recognition of the same antigen.¹³ On the other hand, it is clear that highly shared TCRs characterized by the same V β usage or the same TCR β CDR3 sequence is observed after infections, antigen-driven autoimmune inflammation and even vaccination.^{14,15} Such observed TCR bias or oligoclonality has been proposed to be categorized as Type I (conserved V chain, *TRBV* or *TRAV* gene, use with non-conserved CDR3 sequences), Type II (conserved *TRBV* or *TRAV* gene use with conserved CDR3 sequences but random J chain, *TRBJ* or *TRAJ* gene, use), Type III (complete TCR β and/or TCR α V-D-J conservation) and Type IV (a modification of Type III where minor CDR3 sequences differences occur).^{15,16} Therefore, it could be hypothesized that the combination of excess exposure to common, perhaps “immunodominant”, gut microbial antigens, together with the potential for promiscuous antigen binding among HLA backgrounds, could result in disease-relevant (healthy control vs IBD, Crohn's disease vs ulcerative colitis) TCR bias or oligoclonality at

¹Division of Gastroenterology and Hepatology, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA; ²Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA; ³Repertoire, HudsonAlpha Institute, Huntsville, AL, USA and ⁴Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, USA
Correspondence: P. J. Mannon (pmannon@uabmc.edu)

Received: 29 November 2017 Revised: 26 April 2018 Accepted: 30 April 2018
Published online: 9 July 2018



differently defined levels of clonotype (e.g. shared TCR β CDR3 sequences only or shared V β -TCR β CDR3 pairing).¹²

In this study we compared the TCR β CDR3 sequences and V β linkage in independent groups of active and inactive Crohn's and ulcerative colitis patients compared to a healthy reference population. We found that IBD patients shared significantly diverse repertoires of expanded TCR β CDR3 sequences that distinguished them from healthy controls but which at this level of clonotype definition were similar between Crohn's disease and ulcerative colitis. Comparison of V β -CDR3 linkage, however, showed that Crohn's disease TCR repertoires could be significantly different from ulcerative colitis. Furthermore, flagellin antigens stimulated proliferation of a small number of public and relatively private clonotypes but did not identify a dominant clone or set of clones that was specific to Crohn's disease. These data support significant expansion of disease-related TCR β CDR3 sequence repertoires in Crohn's disease and ulcerative colitis that may reflect the excessive exposure to a broad variety of gut microbial antigens that accompany the gut inflammation in both IBDs.

MATERIALS AND METHODS

Patient enrollment

Patients with verified IBD (endoscopic, radiographic, and histologic criteria) signed an IRB-approved protocol consent explaining specimen collection procedures, study rationale, and risk. Active disease status was defined as a Harvey Bradshaw score ≥ 5 for Crohn's disease or a Simple Clinical Colitis Activity Index of ≥ 4 for ulcerative colitis and confirmed by endoscopy or imaging.

TCR β CDR3 sequencing and repertoire alignment

Peripheral blood mononuclear cells were isolated by standard Ficoll gradient separation from whole blood (50 mL) within 4 h of collection. Isolation of lamina propria mononuclear cells from endoscopic biopsies was done as previously described.¹⁷ After CD14-depletion, CD4 cells were isolated by magnetic CD4 Microbeads (Miltenyi Biotec GmbH) and processed for total RNA (RNeasy, Qiagen, Valencia, CA). RT-PCR reactions were conducted using a set of commercially available nested sequence-specific primers covering human V-beta genes (forward primers) and constant C-beta primers (reverse primers) (iRepertoire Inc., Huntsville, AL). An Illumina paired-end sequencing communal primer B was linked to each forward inside V-beta primer. Illumina paired-end sequencing communal primer A and a barcode sequence of 6 nucleotides were linked to the reverse inside C-beta primer. In brief, cDNA was reverse transcribed from 50 ng of total RNA sample using a mixture of forward V-beta and reverse C-beta primers and reagents from the OneStep RT-PCR kit (Qiagen). The first round of RT-PCR was performed at: 50 °C, 41 min; 95 °C, 15 min; 94 °C, 30 s, 60 °C, 5 min, 72 °C, 30 s, for ten cycles; 94 °C, 30 s, 72 °C, 3 min, for ten cycles; 72 °C, 5 min. After the first round of PCR, 2 μ L of PCR1 product was added to 48 μ L of PCR2 Multiplex Mix (Qiagen). The second round PCR was performed with a set of communal primers that complete the Illumina adaptor sequences at: 95 °C, 15 min; 94 °C, 30 s, 72 °C, 90 s, for 30 cycles; 72 °C, 5 min. After PCR2, 50 μ L of PCR2 product was run on a 2% agarose gel to assess amplification success. Amplicon bands were excised and gel purified using a Qiaquick Gel Purification kit (Qiagen) per manufacturer's instructions, and the final library was eluted in 37 μ L of nuclease-free water. The TCR beta libraries were pooled and sequenced with Illumina HiSeq 2500, v4 250 cycle kit using 125 paired-end reads through the Genome Services Laboratory (GSL) at HudsonAlpha Institute for Biotechnology (Huntsville, AL). The output of the TCR beta sequence covers within the third framework region through the CDR3 and the beginning of the constant region. The sequence information for all primers used for the library preparation can be found in US Patent Office (US9012148).

Raw data were analyzed using the previously described iRmap program.^{18,19} Briefly, sequence reads were de-multiplexed according to barcode sequences at the 5' end of reads from the TCR beta constant region. Reads were then trimmed according to their base qualities with a 2-base sliding window. If either quality value in this window was lower than 20, the sequence stretches from the 3' end were trimmed out from the original read. Trimmed pair-end reads were joined together through overlapping alignment with a modified Needleman–Wunsch algorithm. If paired forward and reverse reads in the overlapping region were not perfectly matched, both forward and reverse reads were thrown out without further consideration. The merged reads were mapped using a Smith–Waterman algorithm to germline V, D, J, and C reference sequences downloaded from the IMGT website.²⁰ To define the CDR3 region, the position of CDR3 boundaries of reference sequences from the IMGT database were migrated onto reads through mapping results, and the resulting CDR3 regions were extracted and translated into amino acid sequences.

CD4 cell proliferation studies

Isolated CD4 cells were cultured overnight in serum-free medium supplemented with IL-2 (20 μ g/mL). Autologous APCs were isolated from whole PBMCs by adherence to plastic culture plates after 45 min and then incubated with antigens overnight (CBir1 or A4 Fla2 flagellin (10 μ g/mL), tetanus toxoid (5 μ g/mL)), as well as no antigen (negative control) or anti-CD3/28 beads (cells:beads 5:1 ratio (Miltenyi Biotech GmbH), positive control). The APCs were washed twice and then incubated (1:10 ratio APCs:CD4 cells) with CD4 cells (10–50 \times 10⁶ cells/mL in 0.1% human AB serum in PBS) that had been loaded with CFSE (5 μ M DMSO, incubated in the dark for 8 min at 25 °C, quenched with 100 μ L 2% human AB serum in PBS, incubated for 10 min at 37 °C, washed and re-suspended in culture medium). After 4 days, cultures were processed for multi-color flow cytometry to assess proliferation (CFSE^{low}) or RNA extraction for CDR3 sequence analysis.

Statistical analysis

Calculations of TCR β CDR3 sequence diversity were determined across a range of $\alpha = 0$ to $\alpha = 4$ by a step size of 0.2 according to the methodological steps of the bioinformatics framework presented by Grieff et al.²¹ That is, for an individual with n total clonotypes observed, the α -parameterized diversity is ${}^{\alpha}D(f) = (\sum_{i=1}^n f_i^{\alpha})^{\frac{1}{\alpha}}$ in which f_i represents the frequency of species i . A one-way analysis of variance (ANOVA) adjusted for unequal sample sizes with the Welch correction was performed to assess the influence of patient group on diversity across the α levels of interest. To further specify significant pairwise differences, Tukey's honest significant difference (HSD) was considered for post-hoc analyses. Fisher's exact test was used to determine whether the frequencies of TCR β CDR3 clonotypes varied between the patient groups; no adjustment for multiple comparisons was applied due to exploratory nature. V β usage bias among the combination of all expanded TCR repertoires was examined with ANOVA using Tukey's HSD to make post-hoc pairwise comparisons between patient groups. The identification of usage differences about specific V β primers with CDR3 clonotypes was completed with Negative Binomial regression with multiple testing using false discovery rate. Differential abundance analysis via Negative Binomial regression normalized by the trimmed mean of M values algorithm was used to model count variability of V β selection within a given CDR3. All resulting probability values found to be less than a decision rule of 0.05 were noted as significant. Statistical and graphical analyses were computed with R v3.3.3 via the Bioconductor package edgeR dependent upon the high-performance computing resources of the University of Alabama at Birmingham Cheaha cluster. Additional packages used along with

Table 1. Demographic and clinical data for patient groups

	Active Crohn's (n = 20)	Inactive Crohn's (n = 19)		Active UC (n = 10)	Inactive UC (n = 8)	
Gender						
(M/F)	8/12	9/10		6/4	6/2	
Concomitant meds						
None	10 (50%)	1 (5%)		2 (20%)	0	
IS	3 (15%)	4 (21%)		2 (20%)	0	
Biologic	3 (15%)	13 (68%)		0	3 (38%)	
CS	5 (25%)	3 (16%)		3 (30%)	1 (13%)	
5ASA	0	5 (26%)		8 (80%)	7 (88%)	
Activity Score			p value			p value
Harvey–Bradshaw	10.4 ± 0.8	0.7 ± 0.2	<0.0001			
SCCAI				6.6 ± 0.7	0.75 ± 0.3	<0.0001
Labs						
CRP	19.7 ± 8.2	4.3 ± 0.9	0.09	13.3 ± 4.9	1.2 ± 0.5	0.06
WBC	9.2 ± 0.9	6.2 ± 0.5	0.01	7.9 ± 1.4	5.8 ± 0.6	0.24
Absolute lymphocyte count	1667 ± 176	1745 ± 203	0.81	1283 ± 244	1823 ± 189	0.16
Total CDR3 reads	3,908,590 ± 316,789	3,938,002 ± 477,300	0.96	4,109,487 ± 1,028,689	5,856,776 ± 1,484,569	0.32
Unique CDR3 seq	244,040 ± 18,946	197,920 ± 16,828	0.08	205,247 ± 18,476	212,669 ± 29,772	0.83

IS immunosuppressant drug, CS corticosteroid, 5ASA mesalamine drug, SCCAI Simple Clinical Colitis Activity Index, WBC white blood cell count, CRP C-reactive protein

the built-in functions included edgeR, ggplot2, gplots, HMP, and onewaytests.

RESULTS

The IBD TCR β CDR3 repertoire is significantly diverse

There was an average of $4,211,279 \pm 2,589,178$ (mean \pm SD) total reads per sample, of which $217,458 \pm 78,247$ were unique sequences in each patient's CD4 CDR3 repertoire. The vast majority of these unique sequences were private (i.e. found in only one individual 87% for HC, 94% for active CD, 97% for active UC) though there were a small number of sequences that were highly shared by patients within each group (active Crohn's, active UC and controls) (Suppl Fig. 1). The numbers of reads were not significantly affected by disease, disease activity, concomitant medications, or total lymphocyte counts though inactive Crohn's disease patients had lower numbers of unique CDR3 sequences compared to active Crohn's disease patients (Table 1 shows demographic and clinical parameters). Furthermore, contamination was also not a factor affecting results as the composition of PCR-generated nucleotide sequences that gave the same deduced amino acid CDR3 sequences were different among patients sharing the same CDR3 clonotype.

The primary analysis included only TCR β CDR3 sequences that were expanded at 100 copies or more, assuming this population would be enriched for clonotypes relevant to active disease. We compared the CDR3 repertoire among the groups using a statistic that measured diversity by accounting for both numbers of unique sequences as well as their abundance (RNA copy numbers). Figure 1 shows that both Crohn's disease groups and active UC had significantly different diversity compared to healthy controls, while the inactive UC diversity measure was not significantly different from controls. This significance persisted across a range of conditions (varying the alpha level) dominated either by excessive numbers of single copy sequences or highly expanded clonotypes supporting the robust character of the IBD CDR3 repertoire (the same significant differences in diversity persisted when all clonotypes regardless of copy number were analyzed (data not shown)).

The IBD repertoire of expanded TCR β CDR3 sequences is different from healthy controls but similar between active Crohn's disease and ulcerative colitis patients

We next determined whether there were T-cell receptor clonotypes that were specifically associated with Crohn's disease by testing for increased frequency of TCR β CDR3 sequences that were not only expanded (increased expression of individual CDR3 sequences) in active Crohn's disease but also highly shared among Crohn's disease patients. Figure 2 shows the frequency distribution of individual CDR3 sequences (expanded at ≥ 100 copies) that were shared among patients in the active Crohn's disease group compared with the identical CDR3 sequence shared within other comparison groups (healthy controls, active ulcerative colitis, and a separate group of inactive Crohn's disease). Compared to healthy controls, there were 2347 expanded CDR3 sequences (out of 77,572 total) that were found to be shared at significantly higher frequencies among 10–90% of active Crohn's disease patients (Fig. 2a, red bubbles). There were far fewer (11 sequences) that were expanded significantly more frequently in healthy controls compared to active Crohn's disease (Fig. 2a, purple bubbles). Comparing active Crohn's disease against active UC, there were very few CDR3 sequences (7 sequences) that were expanded in a "Crohn's specific" way (Fig. 2b, red bubbles). The CDR3 sequences that were significantly expanded in Crohn's disease were also expanded in active UC at similar frequencies (Fig. 2b, blue bubbles). Moreover, many of the expanded CDR3 sequences in active Crohn's disease were expanded at similar frequencies in an independent group of inactive Crohn's disease patients (Fig. 2c, blue bubbles). Whether the cell phenotype (e.g. effector, memory or regulatory) of the shared CDR3 clonotype in active vs remitted disease is the same is unknown.

Similar results were seen for shared CDR3 sequence frequencies in active UC compared to controls, active Crohn's disease and inactive UC. Compared to healthy controls, there were 797 CDR3 sequences (out of 55,359 total) that were found to be shared at significantly higher frequencies among 20–90% of active UC disease patients (Fig. 2d, red bubbles). There were more clonotypes that were significantly expanded in active UC (597) vs



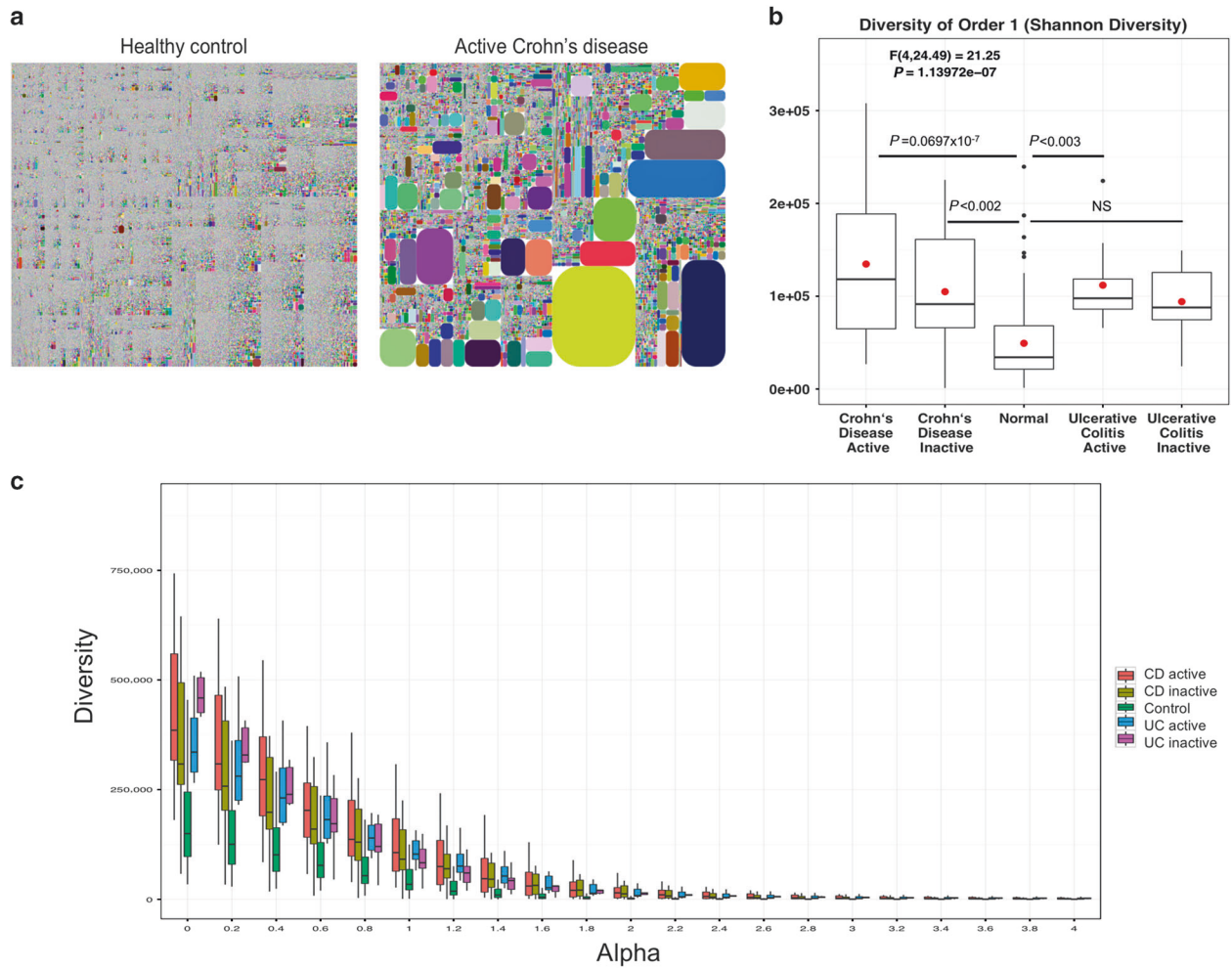


Fig. 1 TCR β CDR3 sequence repertoire diversity in active IBD patients and controls. **a** Tree maps of TCR β CDR3 sequences from a representative healthy control and an active Crohn's patient. Each spot represents a unique sequence; spot size correlates to sequence copy number. **b** Shannon diversity index ($\alpha = 1$) comparing the healthy control group to active Crohn's disease, inactive Crohn's disease, active ulcerative colitis, and inactive ulcerative colitis (p values from Tukey's post hoc test for multiple comparisons). Box and whisker plots are shown with the red dots representing the mean values. **c** Diversity index profile over a continuum of alpha parameters

active Crohn's (Fig. 2e, red bubbles) though most of these were relatively private, expanded in only 20–30% of the active UC patients. Similar to Crohn's disease, many of the expanded CDR3 sequences in active UC were expanded at similar frequencies in an independent group of inactive UC patients (Fig. 2f).

Significant differences in CDR3-V β linkage occurs among shared expanded CDR3 sequences in active Crohn's disease and ulcerative colitis patients

To look further into possible differences in the TCR repertoires between CD and UC patients, we first looked at V β usage by CD4 cell TCR β CDR3 sequences. Summary statistics of V β usage by the overall CDR3 repertoire was fairly similar between Crohn's and UC while significantly different from healthy controls (Table 2). However, when using a finer definition of clonotype, that is specific V β pairing with the same CDR3 sequence within groups of active Crohn's disease and active UC for instance, differences in oligoclonality for V β -CDR3 pairings are detected. For instance, testing for significant differences in V β usage by the top 20 most highly shared CDR3 sequences in the active CD group (also found at 40–90% prevalence in the active UC group, see Fig. 2b), 12 of the CDR3 sequences have significantly different V β usage in Crohn's disease compared to ulcerative colitis to varying degrees

(Suppl. Table 1). To illustrate this point, chord diagrams demonstrate the distribution of V β usage for a single CDR3 sequence (and same J β chain) among CD and UC patients (Suppl. Figure 2). The highly shared CDR3 sequence ASSPSTDTQY (found in Crohn's disease 90%, ulcerative colitis 90%, and healthy controls 64%) has no significant differences in V β usage in Crohn's disease vs ulcerative colitis and in fact share many of the V β -CDR3 linkages proportionately throughout the repertoire. In contrast, the highly shared CDR3 sequence ASSPQETQY (found in Crohn's disease 85%, ulcerative colitis 90%, and healthy controls 27%) has significantly different V β -CDR3 especially with TRBV6-3 and 7-2 where the ratio of usage is reversed (Suppl. Figure 2 and Suppl. Table 1).

To further investigate the prevalence of differences in oligoclonality based on V β -CDR3 pairing in Crohn's disease and ulcerative colitis, we tested for significant V β usage in all the CDR3 sequences shared by 15% or more of active Crohn's patients (total 1100) compared to active ulcerative colitis. Below the 15% level, numbers of patients per group (≤ 2) were so small as to preclude any chance of getting a significant result. Figure 3 shows the percentage of CDR3 sequences that have significantly different V β usage compared to ulcerative colitis at each level of their sharing among the active Crohn's disease group. The total number of sequences analyzed was 1100. So for instance, of the

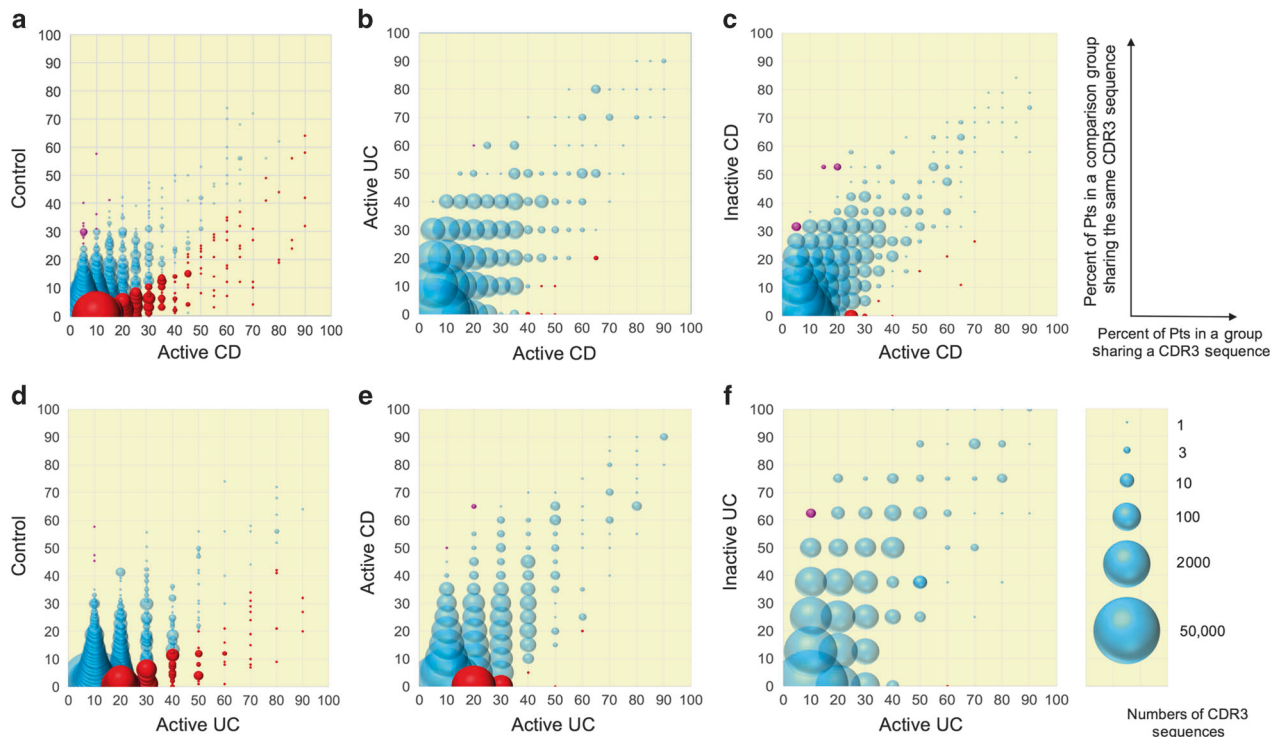


Fig. 2 Distribution of unique TCR β CDR3 sequences between subject groups. Histograms show the frequency of individual sequences present in a reference group (x-axis) and a comparison group (y-axis). A bubble represents a single or multiple CDR3 sequences that are found at the indicated frequencies in the two groups being compared (the legend shows estimates of numbers of sequences represented by bubble size using a logarithmic scale). The bubble color indicates whether the CDR3 sequences represented by the bubble have significantly different frequencies between the groups being compared: Blue bubble, no difference, Red bubble, $p < 0.05$ Fisher's exact test x-axis value vs y-axis value, purple bubble, $p < 0.05$ y-axis value vs x-axis value. **a** Active Crohn's disease, $n = 20$ (x-axis) vs healthy controls, $n = 97$ (y-axis). **b** Active CD vs active UC, $n = 10$. **c** Active CD vs inactive CD, $n = 19$. **d** Active UC, $n = 10$ vs healthy controls. **e** Active UC vs active CD. **f** Active UC vs inactive UC, $n = 8$

four expanded CDR3 sequences (Fig. 3, dotted line) that are shared by 90% of Crohn's disease patients, 25% (1 of 4) (Fig. 3 blue bar; also see Suppl. Table 1) had significantly different V β usage compared to the active ulcerative colitis group; of the 40 CDR3 sequences shared by 35% of Crohn's disease patients, 25% of them had significantly different V β usage. These data show two things, first that Crohn's disease and ulcerative colitis patients can share identical clonotypes based on V β -CDR3 pairings and that V β usage by CDR3 can be significantly different between the two disease groups.

The majority of TCR β CDR3 sequences expanded in lamina propria CD4 cells in active Crohn's disease are also expanded in the peripheral blood

Lastly, we investigated whether expanded clonotypes isolated from active Crohn's disease colonic tissue (endoscopic biopsies) were similarly represented in the peripheral blood compartment. We amplified the CDR3 repertoire from the LP T cells isolated from colonic biopsies and compared them to T cells isolated from paired peripheral blood (PB). Each patient ($n = 5$) had 5190 ± 257 (mean \pm SEM) unique CDR3 sequences that were expanded at >100 copies in the LP and/or PB T-cell populations (13,830 total unique sequences in this dataset overall) with $33 \pm 7\%$ sequences found in both PB and LP repertoires, $51 \pm 7\%$ only in PB and $16 \pm 0.7\%$ found in the LP only (576 private sequences seen in the LP of only one of the patients, 268 shared by 2 patients and 10 shared by 3 patients, Suppl. Figure 3). These data show that of all the CDR3 sequences expanded in the LP of active Crohn's patients, two thirds are found expanded in the patient's PB and one third expanded only in the LP.

Flagellin antigens expand IBD clonotypes in vitro

To test whether we could identify gut microbial antigen-specific CD4 cells belonging to the highly shared or "Crohn's-specific" expanded CDR3 sequences, we studied a patient with Crohn's colitis who was seropositive for anti-A4-Fla2 and anti-CBir1 antibodies. Incubation of isolated CD4 cells with autologous antigen-presenting cells induced proliferation of populations of CD4 cells after polyclonal T-cell stimulation (anti-CD2, CD3 and CD28 beads), common vaccine antigen tetanus toxoid and the cognate flagellin antigens that react with anti-CBir1 and A4-Fla2 antibodies (Fig. 4). We repeated the proliferation assay, collecting the cells after 4 days and sequencing the CD4 TCR β CDR3 repertoire to compare the effects of polyclonal T cell vs antigen-specific stimulation on expansion of CD4 clonotypes in order to identify candidate antigen reactive CRD3 sequences based on differential expansion. Table 3 shows the copy numbers of the top expanded sequences induced by A4 Fla2 or polyclonal T-cell stimulation vs no stimulation. Clonotype expansion was expressed as the fold-increase in sequence copy number from the flagellin antigen culture over the control culture and compared to proliferation induced by polyclonal T-cell receptor stimulation. Focusing on sequences that were greater than twofold amplified by A4 Fla2 exposure compared to polyclonal T-cell stimulation, we found two sequences expanded preferentially by A4 Fla2 exposure that were previously identified in our independent group of active Crohn's patients among the significantly expanded TCR β CDR3 sequences (>100 copies) compared to controls, ASSSYNEQF, and SASRNTEAF. The CDR3 sequence ASSSYNEQF is notable in that it is found in 100% of the Crohn's patients and 55% of the controls at any copy number level but is only significantly

Table 2. Statistical testing results of CDR3 sequence Vβ usage between the entire repertoires in control (HC), active Crohn's disease (CD), and active ulcerative colitis (UC) groups

Vβ	ANOVA	Tukey's HSD multiple comparison <i>p</i> values		
		HC-CD	UC-CD	UC-HC
hTRBV10.1	0.45	0.505	1	0.74
hTRBV10.2	0.73	0.76	1	0.9
hTRBV10.3	0.63	0.81	0.95	0.7
hTRBV11.1	0.93	0.93	1	0.98
hTRBV11.2	3.46E-12	1.18E-08	0.59	1.39E-07
hTRBV11.3	4.23E-03	0.01	1	0.08
hTRBV12.4	2.07E-11	1.50E-10	0.39	6.47E-04
hTRBV12.5	3.53E-03	2.52E-02	0.08	1
hTRBV13	0.53	0.53	0.67	1
hTRBV14	3.23E-04	3.58E-04	0.54	0.21
hTRBV16	1.91E-29	1.21E-14	0.11	2.14E-14
hTRBV18	9.18E-22	3.61E-14	0.64	4.89E-13
hTRBV19	7.63E-08	1.73E-05	0.7	5.58E-05
hTRBV2	0.33	0.3	0.84	0.9
hTRBV20.1	1.08E-13	3.73E-11	1	6.87E-07
hTRBV24.1	0.03	0.03	0.66	0.62
hTRBV25.1	0.76	0.89	0.97	0.807
hTRBV27	0.25	0.33	1	0.57
hTRBV28	0.36	0.71	0.82	0.4
hTRBV3.1	1.12E-27	9.21E-15	1	3.90E-14
hTRBV30	0.94	0.99	0.93	0.94
hTRBV4.1	2.27E-03	2.06E-03	0.53	0.4
hTRBV4.2	8.83E-06	4.82E-06	0.07	0.52
hTRBV4.3	0.75	0.87	0.74	0.87
hTRBV5.1	1.41E-07	1.06E-05	0.9	2.18E-04
hTRBV5.4	0.34	0.49	0.98	0.53
hTRBV5.5	5.60E-05	1.04E-03	0.9	4.58E-03
hTRBV5.6	6.20E-06	1.75E-04	0.92	1.69E-03
hTRBV5.8	0.92	0.91	0.98	1
hTRBV6.1	1.15E-28	8.55E-15	0.95	3.56E-14
hTRBV6.3	2.27E-11	3.54E-08	0.67	6.02E-07
hTRBV6.4	0.41	0.51	1	0.66
hTRBV6.5	9.29E-23	2.85E-14	1	6.46E-12
hTRBV6.6	1.17E-27	8.77E-15	0.97	4.74E-14
hTRBV6.8	3.71E-60	7.66E-15	0.96	7.66E-15
hTRBV6.9	3.67E-22	3.01E-14	0.94	3.82E-11
hTRBV7.2	0.01	0.05	0.76	0.03
hTRBV7.3	0.041	0.78	0.2	0.04
hTRBV7.4	0.82	0.87	1	0.9
hTRBV7.6	3.20E-11	2.04E-09	0.95	3.04E-05
hTRBV7.7	2.56E-21	2.76E-14	0.41	5.98E-09
hTRBV7.8	0.84	0.96	0.95	0.85
hTRBV7.9	0.08	0.11	0.97	0.43

expanded in active Crohn's disease 55% vs 7% of controls, $p < 0.000001$. In addition, there were other CDR3 sequences that met the level of >2-fold antigen-specific proliferation but were seemingly private to the patient (e.g. ASSLGANRGNTIY and ASSEPTGGSEAF). Extending this analysis of A4 Fla2-expanded CDR3 sequences to the dataset of all CDR3 sequences in active

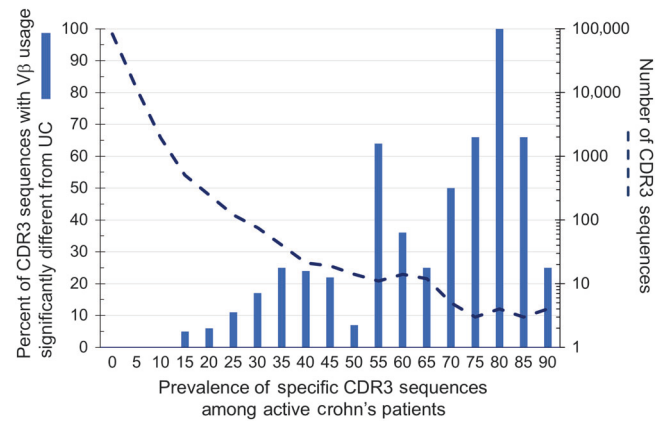


Fig. 3 Significantly different Vβ usage by CDR3 sequences highly shared by Crohn's disease patients compared to ulcerative colitis. The x-axis shows the percent of active Crohn's patients who share the same CDR3 sequences whose absolute numbers are represented by the dotted line (y-axis, right). The bars represent the percent of the CDR3 sequences at each level of sharing in the active Crohn's disease group that have significantly different Vβ usage compared to the same CDR3 sequence found in active UC patients

Crohn's disease and controls regardless of copy number, we actually identify more sequences as candidate A4 Fla2-reactive clonotypes that are shared significantly more frequently in active Crohn's disease vs healthy controls. For instance, the clonotypes SARTAVNTEAF, ASSRRMNTEAF, ASRFDEAF, and SARTASNQPQH are all significantly associated with the active Crohn's disease group but at a level <100 copy number which excluded them from the primary analysis

DISCUSSION

The main aim of this study was to determine whether Crohn's patients shared a T-cell receptor repertoire reflective of the shared B-cell reactivity to certain flagellin antigens. The data show that a relatively small number of expanded TCRβ CDR3 clonotypes are highly shared not only among Crohn's patients but also between Crohn's and UC patients. The surprising degree of sharing of expanded TCRβ CDR3 sequences among Crohn's and UC patients might have been anticipated since the gut inflammation in both diseases confers excessive exposure of mucosal immune cells to prevalent microbial antigens. TCR bias, the emergence of a TCR repertoire with conserved V regions and/or CDR3 sequences during an immune response, has been observed during chronic exposure to microbes as well as during autoimmune inflammation.^{18,22,23} While IBD is not considered an infectious disease or strictly an autoimmune inflammation, the fact that Crohn's disease patients develop shared serum antibody reactivity to a number of gut commensal microbial antigens, especially flagellins, implicates exposure to microbes that drives the ongoing inflammatory response. Why similar antibodies do not develop in UC patients is not clear. However, while not fully elucidated, the mechanisms for the susceptibility to colitis in each disease are not thought to be identical either. It is hypothesized that a dysregulated innate immune response in Crohn's disease drives an adaptive reactivity to immunodominant gut commensal bacterial antigens. On the other hand, the lack of aberrant B-cell reactivity to the same flagellins in UC could reflect the fundamental difference in mechanisms of susceptibility to colitis between the diseases as well as point to differences in the function of T and B cells in these diseases.²⁴ For instance, defective epithelial barrier function (as a mechanism of disease for ulcerative colitis) could allow an excessive exposure of microbial antigens to an overwhelmed

Table 3. Top TCR β CDR3 sequences expanded in vitro by flagellin antigen and polyclonal T-cell stimulation

	Fla2	CD2/28	No Ag	Fla2/no Ag	CD2/28/no Ag	ratio	% act CD > 100 copies	% HC > 100 copies	% act CD < 100 copies	% HC < 100 copies
ASSFDTEAF	1619	3184	1772	0.91	1.8	0.51	15% (3)	6% (6)	70% (14)	61% (58)
SARAASDTQY	1617	1575	510	3.17	3.09	1.03	0	0	15% (3)	5% (5)
SASRNTEAF	1038	366	111	9.35	3.30	2.84	15% (3)*	2% (2)	80% (16) [#]	32% (31)
SARTAVNTEAF	1249	275	244	5.11	1.13	4.54	0	0	15% (3)*	1% (1)
ASSRRMNTEAF	913	315	160	5.71	1.97	2.90	0	1%	55% (11)*	30 (29)%
ASSSYNEQF	709	98	100	7.09	0.98	7.23	55% (11) [#]	7% (7)	45% (9)	50% (50)
ASRFDEAF	696	42	108	6.44	0.39	16.57	0	1%	35% (7)**	10 (10)%
SARTASNQPQH	683	29	315	2.16	0.09	23.55	0	0	30% (6)**	6% (6)
ASSYGGVTEAF	609	17	52	11.71	0.33	35.82	0	0	5% (1)	2% (2)
ASSLSGRQPQH	560	748	75	7.46	9.97	0.75	0	0	20% (4)	20% (4)

* $p < 0.05$, ** $p < 0.01$, [#] $p < 0.005$

but otherwise intact innate immune response, resulting in lack of dysregulated antibody responses but driving an adaptive T-cell response to similar antigens.

Early studies into specific infectious or autoimmune causes for IBD had demonstrated T-cell receptor bias in the healthy human intestine. Data show that both intraepithelial lymphocyte and lamina propria T cells have preferential V β usage within an individual, but shared TCR V β bias has been more difficult to show between individuals.^{25–27} In IBD studies, most existing reports on TCR bias have relied on PCR- and immunohistochemistry-generated data to describe TCR V β frequencies in individual subjects at the peripheral blood and gut tissue level. No consistent V β usage pattern was robustly demonstrated as pathognomonic of IBD although V β 6 has been singled out for its predominant use among patients across studies and may differentiate intestinal T cells from peripheral T cells. Our data confirm a significant preference for V β 6.3 usage in the peripheral blood CD4 TCR repertoires of both active CD and UC patients, and significantly less usage of V β 6.1, 6.5, 6.6, 6.8 and 6.9, compared to the healthy control group.

Newer studies of TCR repertoire bias in IBD have used high throughput sequencing of DNA and RNA to define CDR3 amino acid sequence diversity and abundance. Consistent with our data, clonotypes defined by TCR β CDR3 sequence have been shown to be expanded in the lamina propria and are present in the peripheral blood as well and persist in tissue during inactive disease; however, the degree of sharing among patients and healthy controls had not been established.^{28–31} A recent report using NexGen sequencing to analyze TCR α and β CDR3 regions in cells from blood and surgical specimens of Crohn's patients identified several clonotypes from the tissue that were shared from 11 to 41% of the 17 patients, and at frequencies of 0.13 to 28.9% of total sequences; this is consistent with our findings although these expanded clonotypes are vastly larger (no individual clone in our dataset is larger than 0.1% of the total sequences) and they include CD4 and CD8 populations.²⁹ Similarly another report of TCR repertoire in the inflamed gut of Crohn's patients found clonal expansion of single clones at up to 58% (no analysis of sharing among the 19 patients studied), though this could reflect the low numbers of clonotypes examined (5000 per sample) and use of 454 sequencing technology.³² Moreover, in several reports, the predicted amino acid sequence of the largest clones' CDR3 regions from specific individuals could also be found in our dataset such as SARGDSNQPQH seen in one and ASSSGTSGYNEQF seen in three of our active CD patients (these reports do not show sharing of these regions among their small numbers of patients, but they were seen in 5–15% of our active Crohn's patient CDR3 sequence dataset).³³

Our focus on CD4 CDR3 clones characterized by high copy number was intended to identify T cells that were likely to be enriched for disease-relevant antigen-reactive clonotypes. Our data show that the degree of shared expanded CDR3 sequences within and between active CD and UC patients is higher compared to healthy controls. This is not unexpected in the setting of a chronic active intestinal inflammation where continuous exposure to microbial antigens can provide TCR stimulation for clonal expansion. It was recently shown that healthy humans have a highly heterogeneous memory T-cell reactivity for selected intestinal bacterial species, similar to IBD patients, but T-cell clonotype sharing, reactive or otherwise, among subjects was not reported (only three subjects were presented for analysis of clonotype sharing defined by V β usage of T cells reactive to bacterial lysates).³⁴ Interestingly, we found a startling similarity in the repertoire sharing of expanded clonotypes between independent groups of Crohn's patients who were studied during active or inactive disease states. This suggests that expanded disease-relevant clonotypes may be long-lived, perhaps maintained by continued antigen exposure, but whether they persist over time as the same phenotype (T_{eff} vs T_{reg} for example) remains to be tested. Clonotypes are also widely shared between inflamed and non-inflamed areas of the gut within the same subject.³² When we looked at sequences expanded in the lamina propria, in most of the patients these clonotypes were not expanded as often in the peripheral blood, but there were 25–43% of sequences similarly expanded in both compartments and 4.9–7.0% with expansion restricted to the gut tissue alone.

The long-term goal of this work is to determine whether the sharing of CDR3 sequence clonotypes or V β -CDR3 oligoclonality that we have shown in Crohn's disease and ulcerative colitis is an important primary feature of the dysregulated immune response or a secondary feature of the intestinal inflammatory damage and dysbiosis. Crohn's disease is unique in this regard as the characteristic B-cell antigen-reactive antibody responses to certain microbial antigens, flagellins among them, can serve as a link to T-cell immune memory given the requirement for interaction of antigen-specific T cells with B cells for isotype switching and antibody production. The epigenetic signature of B cells also overlaps with Crohn's disease finely mapped genetic risk single nucleotide polymorphisms significantly more than with UC.²⁴ Our data show that seropositivity for an immunodominant flagellin antigen can link to CDR3 sequences that are significantly expanded in Crohn's disease; when we activate peripheral CD4 T cells from an anti-A4 Fla2 antibody-positive Crohn's patient in vitro using polyclonal as well as A4 Fla2 antigen-specific stimulation, we significantly expanded a small number of antigen-reactive sequences, only two of which were



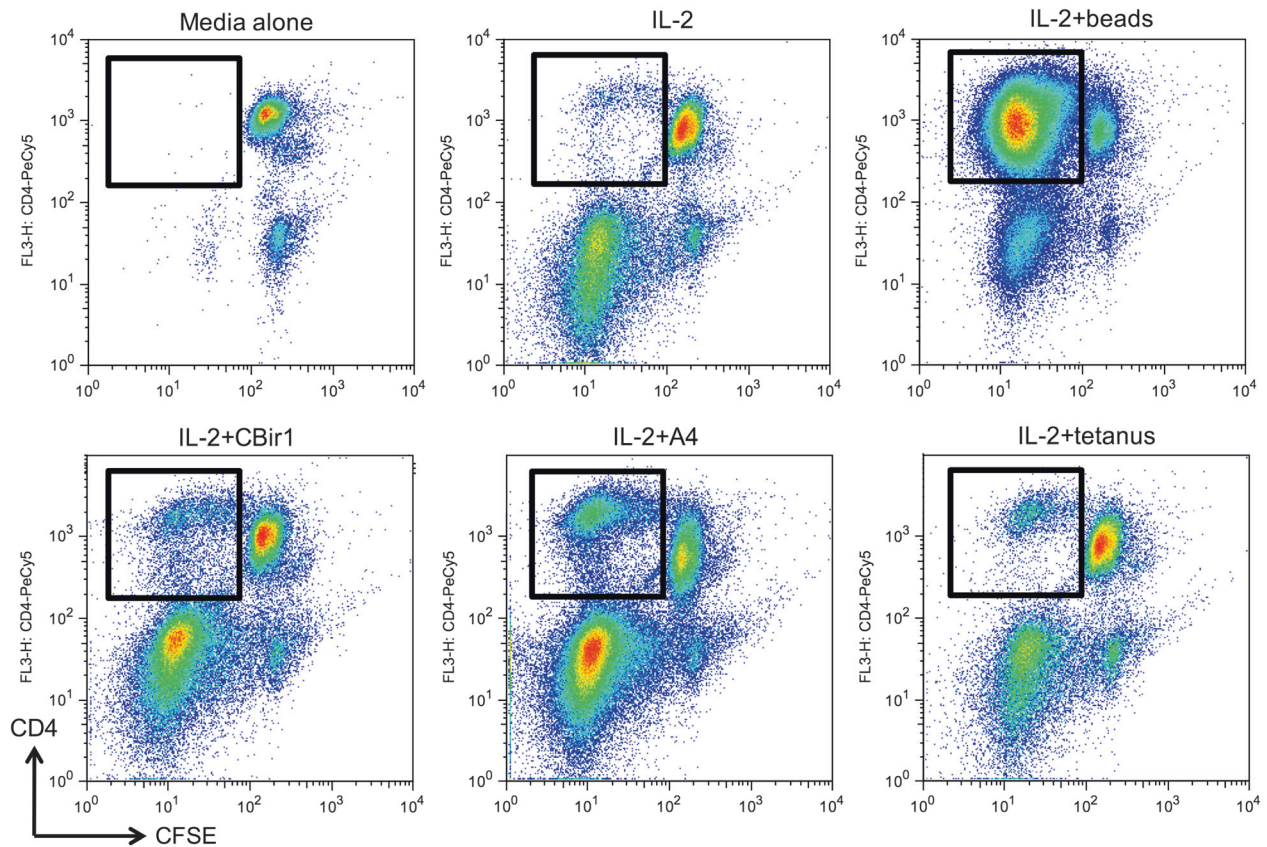


Fig. 4 Microbial antigen-specific CD4 T-cell reactivity detected by cell proliferation. CD4 T cells from a Crohn's disease patient seropositive for anti-CBir and anti-A4 Fla2 (flagellin in A4 bacterial strain) were loaded with CFSE and cultured for 4 days with IL-2 and autologous antigen-presenting cells (APCs) that were loaded with specific antigens. Controls were APCs alone with or without anti-CD2/3/28 beads

found in our set of expanded CDR3 sequences in active Crohn's disease but four more of which were highly shared in Crohn's disease but at lower levels of expression even in the setting of active disease. Additional studies are needed to validate the strength of the association of T-cell antigen-specific reactivity with seropositive status. Clonotype isolation for antigen-specificity testing will be needed to indisputably link TCR β CDR3 sequences to flagellin antigen peptide-MHC recognition. However, at present these data do not support the idea that responses to a very narrow group of flagellin antigens entirely explain IBD inflammation.

The T-cell repertoire in active Crohn's disease, both in the lamina propria and in the blood, are expected to contain clonotypes that expand after interaction with cognate antigens characteristic of the disease-causing agent or following epithelial barrier injury and unrestrained antigen exposure. This clonotype expansion influences the diversity of the TCR repertoire and possibly provides a "fingerprint" of the antigen-driven T-cell response. Moreover, the antigen-specific aberrant B-cell response may identify high value antigens central to isolating disease-relevant T cells. This would be particularly useful for comparing antigen-specific T cells (and especially identical CDR3 clonotypes) from Crohn's and UC patients. The sharing of TCR β CDR3 sequences between Crohn's and UC patients may also provide another clue as to how these diseases are similar but also different. V chain and CDR3 sequence linkage differences, for instance, are consistent with fine-mapping of the MHC region that shows most IBD risk-associated HLA alleles have a preferential role in either Crohn's disease or UC.³⁵ The details of TCR α and β CDR3 pairing, V-J usage, MHC specificity, antigen processing, thymic vs peripheral clonotype selection, and effects of T-cell phenotype

between Crohn's disease and controls and UC patients will provide an understanding of how such CDR3 sequence sharing can be observed in such different types of IBD. Furthermore, the differential contribution of shared vs private clonotypes to inflammatory responses in Crohn's disease, and whether they recognize a narrow group of antigen peptide-MHC complexes or a diverse antigen population will help determine their value as targets for innovative immunotherapies.

ACKNOWLEDGEMENTS

This work was supported by the Immunology, Autoimmunity and Transplantation Committee and the Program in Immunology at the University of Alabama at Birmingham.

AUTHOR CONTRIBUTIONS

JW: Specimen handling and library preparation; AHP: Statistical analysis and interpretation and manuscript evaluation; QY: Sequencing and data analysis; CW: Sequencing and data analysis; WP: Sequencing and data analysis; TL: Patient enrollment and data collection; TS: Patient enrollment and data collection; XC: Statistical analysis and interpretation and manuscript evaluation; COE: Development of the work and manuscript evaluation; JH: Development of the work and manuscript evaluation; PJM: Conception of the work, data analysis, drafting of manuscript.

ADDITIONAL INFORMATION

The online version of this article (<https://doi.org/10.1038/s41385-018-0046-z>) contains supplementary material, which is available to authorized users.

Competing interests: Jian Han is the founder of iRepertoire. The remaining authors declare no competing interests.

REFERENCES

1. Kiesler, P., Fuss, I. J. & Strober, W. Experimental models of inflammatory bowel diseases. *Cell Mol. Gastroenterol. Hepatol.* **1**, 154–170 (2015).
2. Aranda, R. et al. Analysis of intestinal lymphocytes in mouse colitis mediated by transfer of CD4⁺, CD45RB^{high} T cells to SCID recipients. *J. Immunol.* **158**, 3464–3473 (1997).
3. Sellon, R. K. et al. Resident enteric bacteria are necessary for development of spontaneous colitis and immune system activation in interleukin-10-deficient mice. *Infect. Immun.* **66**, 5224–5231 (1998).
4. Prideaux, L., De Cruz, P., Ng, S. C. & Kamm, M. A. Serological antibodies in inflammatory bowel disease: a systematic review. *Inflamm. Bowel Dis.* **18**, 1340–1355 (2012).
5. Kugathasan, S. et al. Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. *Lancet* **389**, 1710–1718 (2017).
6. Calderon-Gomez, E. et al. Commensal-specific CD4(+) cells from patients with Crohn's disease have a T-helper 17 inflammatory profile. *Gastroenterology* **151**, 489–500 e483 (2016).
7. Shen, C., Landers, C. J., Derkowski, C., Elson, C. O. & Targan, S. R. Enhanced CBir1-specific innate and adaptive immune responses in Crohn's disease. *Inflamm. Bowel Dis.* **14**, 1641–1651 (2008).
8. Jager, E. et al. Simultaneous humoral and cellular immune response against cancer-testis antigen NY-ESO-1: definition of human histocompatibility leukocyte antigen (HLA)-A2-binding peptide epitopes. *J. Exp. Med.* **187**, 265–270 (1998).
9. Zeng, G., Wang, X., Robbins, P. F., Rosenberg, S. A. & Wang, R. F. CD4(+) T cell recognition of MHC class II-restricted epitopes from NY-ESO-1 presented by a prevalent HLA DP4 allele: association with NY-ESO-1 antibody production. *Proc. Natl. Acad. Sci. USA* **98**, 3964–3969 (2001).
10. Epstein, M., Barencio, M., Klein, N., Hubank, M. & Callard, R. E. Revealing individual signatures of human T cell CDR3 sequence repertoires with Kidera Factors. *PLoS ONE* **9**, e86986 (2014).
11. Warren, R. L. et al. Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome Res.* **21**, 790–797 (2011).
12. Sewell, A. K. Why must T cells be cross-reactive? *Nat. Rev. Immunol.* **12**, 669–677 (2012).
13. Singh, N. K. et al. Emerging concepts in TCR specificity: rationalizing and (maybe) predicting outcomes. *J. Immunol.* **199**, 2203–2213 (2017).
14. Li, H., Ye, C., Ji, G. & Han, J. Determinants of public T cell responses. *Cell Res.* **22**, 33–42 (2012).
15. Miles, J. J., Douek, D. C. & Price, D. A. Bias in the alpha β T-cell repertoire: implications for disease pathogenesis and vaccination. *Immunol. Cell Biol.* **89**, 375–387 (2011).
16. Turner, S. J., Doherty, P. C., McCluskey, J. & Rossjohn, J. Structural determinants of T-cell receptor bias in immunity. *Nat. Rev. Immunol.* **6**, 883–894 (2006).
17. Fuss, I. J. et al. Both IL-12p70 and IL-23 are synthesized during active Crohn's disease and are down-regulated by treatment with anti-IL-12 p40 monoclonal antibody. *Inflamm. Bowel Dis.* **12**, 9–15 (2006).
18. Sui, W. et al. Composition and variation analysis of the TCR beta-chain CDR3 repertoire in systemic lupus erythematosus using high-throughput sequencing. *Mol. Immunol.* **67**(2 Pt B), 455–464 (2015).
19. Wang, C. et al. High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. *Proc. Natl. Acad. Sci. USA* **107**, 1518–1523 (2010).
20. Lefranc, M. P. et al. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev. Comp. Immunol.* **27**, 55–77 (2003).
21. Greiff, V. et al. A bioinformatic framework for immune repertoire diversity profiling enables detection of immunological status. *Genome Med.* **7**, 49 (2015).
22. Luo, W. et al. Limited T cell receptor repertoire diversity in tuberculosis patients correlates with clinical severity. *PLoS ONE* **7**, e48117 (2012).
23. Sing, G. K. et al. A molecular comparison of T lymphocyte populations infiltrating the liver and circulating in the blood of patients with chronic hepatitis B: evidence for antigen-driven selection of a public complementarity-determining region 3 (CDR3) motif. *Hepatology* **33**, 1288–1298 (2001).
24. Farh, K. K. et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518**, 337–343 (2015).
25. Balk, S. P. et al. Oligoclonal expansion and CD1 recognition by human intestinal intraepithelial lymphocytes. *Science* **253**, 1411–1415 (1991).
26. Blumberg, R. S., Yockey, C. E., Gross, G. G., Ebert, E. C. & Balk, S. P. Human intestinal intraepithelial lymphocytes are derived from a limited number of T cell clones that utilize multiple V beta T cell receptor genes. *J. Immunol.* **150**, 5144–5153 (1993).
27. Van Kerckhove, C. et al. Oligoclonality of human intestinal intraepithelial T cells. *J. Exp. Med.* **175**, 57–63 (1992).
28. Camus, M. et al. Oligoclonal expansions of mucosal T cells in Crohn's disease predominate in NKG2D-expressing CD4 T cells. *Mucosal Immunol.* **7**, 325–334 (2014).
29. Chapman, C. G. et al. Characterization of T-cell receptor repertoire in inflamed tissues of patients with Crohn's disease through deep sequencing. *Inflamm. Bowel Dis.* **22**, 1275–1285 (2016).
30. May, E. et al. Regional variation of the alpha β T cell repertoire in the colon of healthy individuals and patients with Crohn's disease. *Human Immunol.* **63**, 467–480 (2002).
31. Nakajima, A. et al. Specific clonal T cell accumulation in intestinal lesions of Crohn's disease. *J. Immunol.* **157**, 5683–5688 (1996).
32. Doorenspleet, M. E. et al. Profoundly expanded T-cell clones in the inflamed and uninfamed intestine of patients with Crohn's disease. *J. Crohns Colitis* **11**, 831–839 (2017).
33. Gulwani-Akolkar, B. et al. Selective expansion of specific T cell receptors in the inflamed colon of Crohn's disease. *J. Clin. Investig.* **98**, 1344–1354 (1996).
34. Hegazy, A. N. et al. Circulating and tissue-resident CD4⁺T cells with reactivity to intestinal microbiota are abundant in healthy individuals and function is altered during inflammation. *Gastroenterology* **153**, 1320–1337 e1316 (2017).
35. Goyette, P. et al. High-density mapping of the MHC identifies a shared role for HLA-DRB1*01:03 in inflammatory bowel diseases and heterozygous advantage in ulcerative colitis. *Nat. Genet.* **47**, 172–179 (2015).

