ORIGINAL ARTICLE

T cell clonal expansions in ileal Crohn's disease are associated with smoking behaviour and postoperative recurrence

Matthieu Allez,[1,2] Claire Auzolle,[1] Marjolaine Ngollo,[2] Hugo Bottois,[2] Victor Chardiny,[2] Ana Maria Corraliza,[3] Azucena Salas,[3] Kevin Perez,[2] Carmen Stefanescu,[4] Stéphane Nancey,[5] Anthony Buisson,[6] Benjamin Pariente,[7] Mathurin Fumery,[8] Harry Sokol,[9] Xavier Tréton,[10] Nicolas Barnich,[11] Philippe Seksik,[12] Lionel Le Bourhis,[2] REMIND Study Group

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2018-317878).

For numbered affiliations see end of article.

Correspondence to
Professor Matthieu Allez, Gastroenterology, Hopital Saint Louis, Paris 75010, France; matthieu.allez@aphp.fr

Received 5 November 2018
Revised 21 January 2019
Accepted 22 January 2019

To cite: Allez M, Auzolle C, Ngollo M, *et al*. *Gut* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/gutjnl-2018-317878

ABSTRACT

T cell clonal expansions are present in the inflamed mucosa of patients with Crohn's disease (CD) and may be implicated in postoperative recurrence after ileocolonic resection.

Methods T cell receptor (TCR) analysis was performed in 57 patients included in a prospective multicentre cohort. Endoscopic recurrence was defined by a Rutgeerts score >i0. DNA and mRNA were extracted from biopsies collected from the surgical specimen and endoscopy, and analysed by high throughput sequencing and microarray, respectively.

Results TCR repertoire in the mucosa of patients with CD displayed diverse clonal expansions. Active smokers at time of surgery had a significantly increased proportion of clonal expansions as compared with non-smokers (25.9%vs17.9%, p=0.02). The percentage of high frequency clones in the surgical specimen was significantly higher in patients with recurrence and correlated with postoperative endoscopic recurrence (area under the curve (AUC) 0.69, 95% CI 0.54 to 0.83). All patients with clonality above 26.8% (18/57) had an endoscopic recurrence. These patients with a high clonality were more frequently smokers than patients with a low clonality (61% vs 23%, p=0.005). The persistence of a similar TCR repertoire at postoperative endoscopy was associated with smoking and disease recurrence. Patients with high clonality showed increased expression of genes associated with CD8 T cells and reduced expression of inflammation-related genes. Expanded clones were found predominantly in the CD8 T cell compartment.

Conclusion Clonal T cell expansions are implicated in postoperative endoscopic recurrence. CD patients with increased proportion of clonal T cell expansions in the ileal mucosa represent a subgroup associated with smoking and where pathogenesis appears as T cell driven.

Trial registration number NCT03458195.

Significance of this study

What is already known on this subject?

- Early postoperative endoscopic recurrence occurs in the majority of patients after surgery.
- Smoking behaviour is the most consistent risk factor for postoperative endoscopic recurrence.

What are the new findings?

- Smoking behaviour is associated with specific changes in the T cell repertoire characterised by reduced repertoire diversity and an increased proportion of T cell clonal expansions.
- Patients with an increased proportion of T cell clonal expansions at time of surgery have an increased risk postoperative endoscopic recurrence.
- Expanded T cell clones persist overtime and may play a major role in postoperative recurrence.

How might it impact on clinical practice in the foreseeable future?

- We identify a subset of patients, characterised by a higher proportion of smokers and a T cell-driven pathogenesis.
- This subset of patients could be potentially better treated with drugs targeting T cell-driven pathways.

INTRODUCTION

Crohn's disease (CD) is an inflammatory bowel disease (IBD) that is characterised by a chronic inflammatory process involving one or several segments of the gastrointestinal tract.[1] This complex disease results from the interaction of genetic heritable traits and environmental factors including the microbiota.[1 2] Genetic studies have highlighted possible defects in the integrity and function of the intestinal mucosal barrier as well as immune regulation, leading to an exacerbated immune response towards the intestinal microbiota.[3 4] Importantly, smoking is the most clearly defined environmental risk factor for the development and progression of CD, but the underlying mechanisms of this harmful effect are poorly understood.[5 6]

Despite improvements in the medical management of CD, more than two-thirds of patients need

intestinal resection, specifically patients with ileal and ileocolonic disease.[1] Furthermore, removal of the diseased segment of the intestine is not a cure, as the majority of patients will have a subsequent recurrence of the disease.[7 8] Postoperative endoscopic recurrence is most frequently observed at the anastomosis and/or in the neoterminal ileum.[7 8] Smoking behaviour is considered to be the most consistent risk factor for postoperative endoscopic recurrence.[8 9]

Mucosal inflammation in CD is associated with an infiltration of activated immune cells.[10] T cells play a dominant role in cell-mediated immunity, recognising and eliminating a vast spectrum of antigens and modulating the immune response. T cells accumulate in inflamed tissues, where their production of inflammatory cytokines and other effector functions contribute to mucosal damage.[11]

T cell antigen specificity is mediated by the T cell receptor (TCR), which is generated in the thymus following stochastic rearrangements of the DNA region encoding this extracellular receptor. Hypervariable complementarity-determining regions (CDRs) of the protein are generated de novo during this process. The sequence of the CDR3 region is a unique feature of each T cell, hence expansions following TCR-dependent antigen recognition spawn clones with identical CDR3 sequences.[12 13] The immune system encounters foreign antigens through microbial colonisation of mucosal surfaces or infections, inducing clonal expansions that shape the TCR repertoire.[12 13]

The antigen specificity of the T cells located in the intestinal mucosa in health or in IBD remains poorly understood. However, an alteration of the TCR repertoire in CD was observed 30 years ago by studies showing decreased TCR diversity and the presence of CD4 T cells expansions in patients with CD.[12 14 15] We described persistent clonal expansions in the inflamed mucosa of patients with CD that may be associated with disease recurrence.[16] The study of the TCR repertoire by next-generation sequencing enhanced our ability to track these clonal expansions.

We sequenced the TCR repertoire of a well-characterised prospective cohort of patients with CD followed from surgery to postoperative endoscopy 6 months later, aiming to identify predictors of early postoperative endoscopic recurrence.[8] Here we report next-generation sequencing analysis of the TCR repertoire from 171 mucosal samples from 57 patients with CD at different locations on the surgical specimen and on endoscopic biopsies performed 6 months later. We show a strong association between smoking behaviour and the T cell repertoire, characterised by increased proportion of clonal T cell expansions and reduced repertoire diversity in the ileal mucosa. We show that the presence of expanded clones in the surgical specimen is associated with an increased risk of endoscopic disease recurrence 6 months later. Furthermore, the persistence of the TCR repertoire over time was associated with disease recurrence. Transcriptional analyses of whole biopsies and TCR sequencing of sorted T cells from the mucosa revealed that the majority of these clonal expansions were in the CD8 T cell lineage.

MATERIALS AND METHODS

CD patients and controls

The present study was undertaken in parallel with a prospective multicentre study performed by the Recherche Maladies INflammatoires Digestives (REMIND) study group and aimed to identify predictors of early postoperative endoscopic recurrence.[8] Inclusion criteria were: age >18 years, ileal or ileocolonic CD and indication of CD-related intestinal surgery (ileocolonic resection). A postoperative treatment (no treatment/5-aminosalycilic acid, thiopurines or antitumour necrosis factor [TNF] agents) was proposed according to a pre-established algorithm, based on the following risk factors: current smoking, previous bowel resection, penetrating phenotype and active perianal disease. Six to 12 months after surgery, a colonoscopy was performed to assess the endoscopic recurrence according to the Rutgeerts score.

Table 1 Patients' characteristics

	n=57
Men, n (%)	32 (56)
Mean age (years, range)	37.4 (20–70)
Median disease duration (years, Q1–Q3)	7.9 (2–11)
Median delay of colonoscopy (months, Q1–Q3)	6.9 (6.1–8.0)
Active smoker at time of surgery, n (%)	20 (35)
Previous surgery, n (%)	11 (19)
Surgical indication, n (%)	
Stricturing disease	35 (61)
Penetrating disease	22 (39)
Disease location (Montreal classification), n (%)	
Ileal (L1)	37 (65)
Colonic (L2)	0
Ileocolonic (L3)	20 (35)
Anoperineal lesion	10 (17)
Granuloma	12 (21)
Extra-intestinal manifestations, n (%)	
Joint manifestations	7 (12)
Skin manifestations	2 (3)
Eye manifestations	1 (2)
Median Harvey-Bradshaw index (Q1–Q3)	5 (3–8)
Mean body mass index (kg/m^2, min–max)	21.3 (14.7–29)
Preoperative enteral nutrition	6 (10)
Preoperative parenteral nutrition	10 (17)
Previous exposure to anti-TNF	29 (51)
Anti-TNF therapy <3 months before surgery, n (%)	24 (42)
Adalimumab	15 (63)
Infliximab	9 (37)
Previous exposure to thiopurines, n (%)	35 (63)
Thiopurines <3 months before surgery, n (%)	18 (32)

TNF, tumour necrosis factor.

Biopsies were collected from the surgical specimen in the inflamed portion (M0) of the ileum and at the proximal margin of the surgical specimen (MI). All patients underwent endoscopy after surgery and biopsies were sampled from the neoterminal ileum (M6). The control group consisted of individuals (n=5) with normal ileocolonoscopy. Biopsies were collected in the ileum and stored at −80°C. Clinical characteristics are given in table 1. All subjects gave written informed consent.

TCR sequencing

DNA was extracted from frozen biopsies by TRizol-based technique according to the manufacturer's protocol. The quantity and purity of total isolated DNA were determined by measuring the absorbance at 260 nm and 280 nm using a NanoDrop 2000 Spectrophotometer.

Five hundred nanograms of DNA from biopsies were sent to Adaptive Biotechnology (Seattle, USA) for sequencing.[17] Shortly, the somatically rearranged CDR3 region was amplified from genomic DNA using a two-step, amplification bias-controlled multiplex PCR approach. Specifically, the first PCR consists of

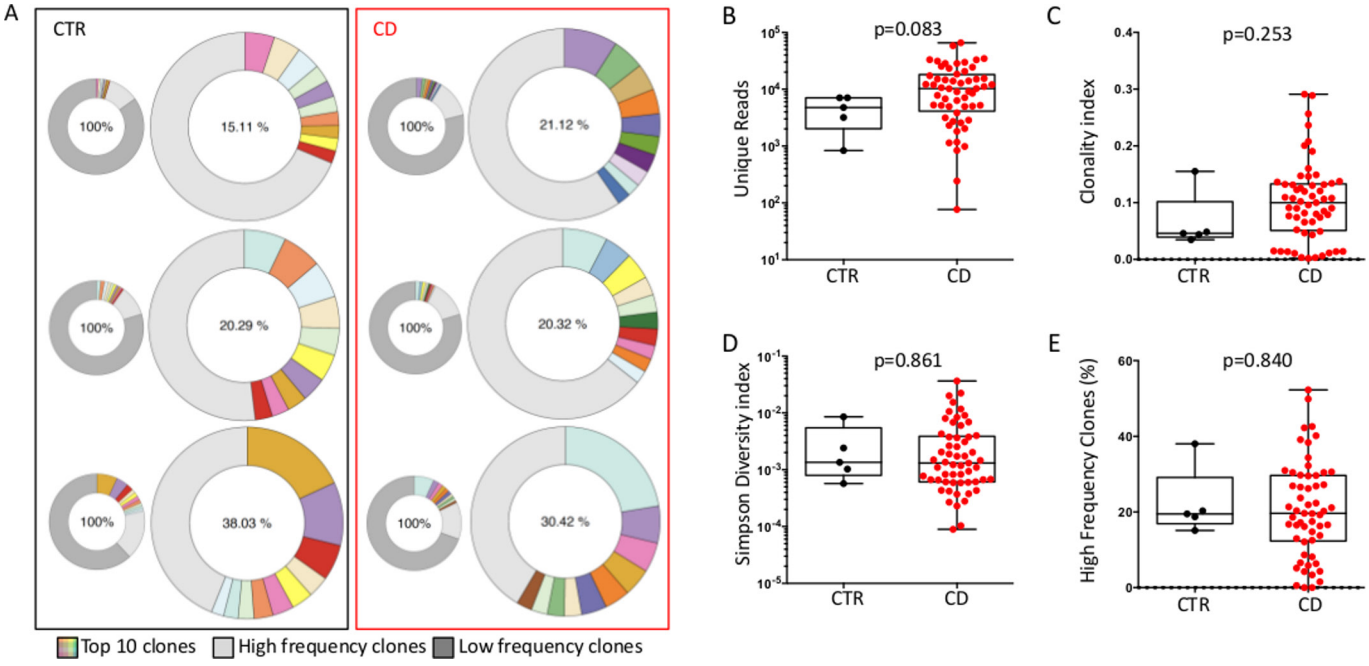

Figure 1 T cell repertoire analysis of ileal intestinal mucosa in Crohn's disease (CD). (A) Donut plots showing the TCRB repertoire analyses from sequencing data of three representative patients and controls. The whole repertoire is depicted by the small donut (100% of the repertoire). The dark grey areas are the fractions occupied by low frequency clones. The enlarged donut plots represent the fractions occupied by high frequency clones in each individual sample (percentage of high frequency clones is indicated at the centre of the donut). The top 10 high frequency clones present in the sample are shown in colour. (B) Number of unique reads detected by sequencing of the TCR repertoire in mucosal ileal biopsies from control (CTR; n=5) and from patients (CD; n=57). (C) Clonality index calculated from the sequencing data of mucosal ileal biopsies from CTR and patients with CD. (D) Comparison of the Simpson diversity index of repertoires from CTR and patients with CD. (E) Percentage represented by high frequency clones within the whole TCR repertoire of CTR and CD patients. No significant differences calculated by Mann-Whitney U test. TCR, T cell receptor.

forward and reverse amplification primers specific for every V and J gene segment and amplifies the hypervariable CDR3 of the immune receptor locus. The second PCR adds a proprietary barcode sequence and Illumina adapter sequences. CDR3 libraries were sequenced on an Illumina instrument. Pipelines provided by Adaptive Biotechnology were used for some analyses. Raw data were also extracted and R/Bioconductor statistical environment in R V.3.3.0 was used to analyse repertoire parameters.

Lymphocyte isolation and sorting of specific subsets

Lamina propria lymphocytes were isolated from surgical specimens as previously described.[18] Lymphocytes were obtained by separation on a Ficoll-hypaque density gradient and then resuspended in PBS for cell sorting. CD8 and CD4 T cells were sorted using the ARIA III (Becton Dickinson) with mean purity >99%.

Transcriptomic analysis by microarray

Total RNA from whole intestinal mucosa biopsies was extracted using TRizol reagent and purified using RNeasy micro kit (Qiagen), according to the manufacturer's instructions.

The quantity and purity of total isolated RNA were determined by measuring the absorbance at 260 nm and 280 nm using a NanoDrop 2000 Spectrophotometer. When the A260/A280 ratio was between 1.8 and 2.2, the isolated RNA was qualified to be pure and was used in subsequent experiments. Samples were then tested for quality with the Caliper LabChip GX High-Throughput Bioanalyzer (Life Sciences).

Microarray data were generated using the Affymetrix GeneChip Human Genome U133 Plus 2.0 arrays, which comprised of 54 675 probe sets. A total of 250 ng of mRNA was labelled using a 3' IVT PLUS kit (Affymetrix) following the manufacturer's instructions. Briefly, mRNA was reverse transcribed using a method that primes the poly(A) tail of mRNA. The mRNA was then amplified with T7 in vitro transcription technology. This method generates amplified and biotinylated complementary RNA. Fifteen micrograms of fragmented and labelled complementary RNA was hybridised overnight according to the manufacturer's instructions. Finally, the arrays were washed to the Affymetrix GeneChip Fluidics Station and scanned on an Affymetrix scanner. Intensity data were extracted using the Affymetrix extraction software.

The intensity data of microarrays were extracted and normalised using robust multiarray average method as implemented in the Affy R package. Batch analysis was performed using SVA R package. All analyses were performed using R/Bioconductor statistical environment.

CIBERSORT is an analytical tool that estimates the relative fraction of 22 immune cell types in a gene expression matrix, such that the sum of all fractions is equal to 1 for a given mixture sample.[19] These 22 cell fractions were calculated via an online calculator (https://cibersort.stanford.edu/). To characterise and to quantify each immune cell subtype, CIBERSORT uses gene expression signature consistent of 547 genes, the gene signature file LM22. The 22 immune cell types were quantified together for each sample and for all patients.

TCR analyses

A unique read is a unique nucleotide sequence generated by immunosequencing of a given sample. This term is synonymous with the term 'clonotype'. The inherent diversity of the immune repertoire generated by V(D)J recombination makes it

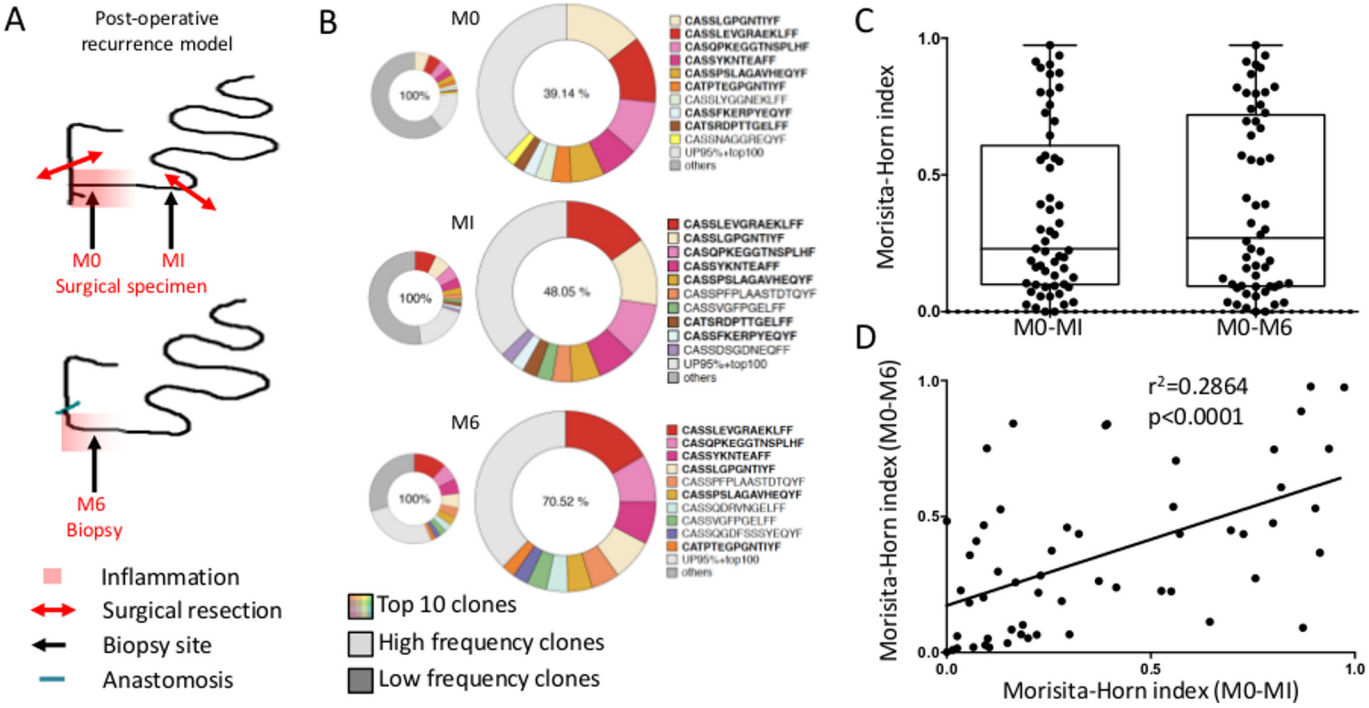

Figure 2 T cell clonal expansions persist over time in the ileum of patients with Crohn's disease (CD). (A) Postoperative recurrence model of ileal CD and location of biopsies sampled from the surgical specimen (M0 and MI) and during endoscopy 6 months later (M6). (B) Donut plots showing the T cell receptor beta chain (TCRB) repertoire analyses from sequencing data of biopsies sampled at time of surgery (M0 and MI) and during endoscopy 6 months later (M6) from one patient with highly persistent clones. The repertoire is depicted as in figure 1A. Clonotypes that are present at M0 and persistent at MI or M6 within the top 10 clones are shown in bold. (C) Analysis of TCR repertoire similarity for each patient by calculated Morisita-Horn index between repertoire at M0 and MI (M0–MI) as well as M0 and M6 (M0–M6). (D) Correlation between M0-MI and M0-M6 Morisita-Horn Indexes. Spearman correlation analysis showed significant correlation (r^2=0.2864; p<0.0001). TCR, T cell receptor.

highly improbable that the same CDR3 nucleotide sequence will be independently created twice, effectively making each CDR3 sequence a unique tag for any specific T or B cell and its clonal descendants.

Clonality is a measure equal to the inverse of the normalised Shannon entropy (a statistic from information theory) of all productive clones in a sample. Values for clonality range from 0 to 1. Values near 1 represent samples with one or a few predominant clones dominating the observed repertoire. Clonality values near 0 represent polyclonal samples. For the TCRB assay, the median clonality of an adult T-cell repertoire in blood is about 0.075. The Simpson diversity index measures the degree of concentration when individuals are classified into types; in the context of the TCR repertoire, it equals the probability of two sequences taken at random from the dataset to be identical.

Highly expanded clones correspond to the percentage taken within the whole repertoire by the hundred most expanded clones. To take into account the possibility of very low frequency T cell clones being present stochastically in the mucosa, we tested these hundred clones for their presence in a 95% interval of confidence. The resulting metric depicts the percentage occupied by clear clonal expansions in the total repertoire (high frequency clones (%)).

Morisita-Horn index is a statistical measure of dispersion of individuals in a population. It was used here to compare overlap among samples and how similar the TCR repertoires are between samples. Values near 1 represent two identical repertoires (independently of sample size and diversity and taking clonal size into account). Values near 0 represent dissimilar repertoires.

Statistical analysis

Quantitative variables were expressed as median and IQRs (Q1–Q3]) or mean and SD according to their distribution. Qualitative variables were expressed by frequency and percentage. The comparison of qualitative variables was realised by χ^2 or Fisher's exact tests. For quantitative variables, we used the Student's t-test or the Mann-Whitney U test when appropriate. We used receiver operating characteristic (ROC) curve analysis to evaluate the performance of our model and evaluate specificity and sensitivity. We determined a remarkable cut-off for which the specificity was at 100% (ie, all patients had a recurrence).

All statistical analyses were two tailed, and a p value of less than 0.05 was considered as statistically significant associations. All analyses were performed using R software, version 3.3.0.

RESULTS

Fifty-seven patients from the REMIND cohort were included in this study (table 1). Median disease duration was 7.9 years. Twenty (35%) patients were active smokers at time of surgery and 11 (19%) already had at least one previous resection. All patients had been operated for a complication, stricturing (61%) or penetrating (39%). Endoscopy was performed within the year after surgery in all patients with a median delay of colonoscopy of 6.9 months. Twenty-nine patients (51%) received a postoperative therapy (thiopurines: n=11; anti-TNF: n=18), including 10 of the 20 smokers at baseline.

Intestinal T cell repertoire displays a high variability of clonal expansions in patients with CD

Analysis of the TCR repertoire in the inflamed ileum (M0) showed important numbers of unique TCR sequences and low clonality indexes, with a trend towards increase number of unique reads and clonality in patients with CD compared with controls (figure 1A–C). The diversity of the TCR, measured with the Simpson index, was high variable among patients with no statistical difference compared with controls (figure 1D). We also analysed the distribution of the different clones in the whole repertoire by calculating the proportion occupied by high frequency clones. This proportion of the most abundant clones within the whole repertoire varied between patients. Indeed, clonal expansions represented as a mean 20.7% (SD=12.6%) of the whole repertoire in the cohort but up to 50% in some patients (figure 1E). Individual expanded clones could represent up to 10% of the repertoire. Importantly, none of the sequences found in high frequency clones were shared by patients in our cohort, demonstrating that there are no common expanded disease-specific TCR sequences (online supplementary figure 1). Taken together, these results show that the TCR repertoire present in the ileal mucosa displays clonal expansions that are highly variable between patients with CD.

Persistence of the TCR repertoire in patients with CD

We compared the TCR repertoire of the inflamed (M0) and macroscopically non-inflamed (MI) areas at the proximal margin of the surgical specimen (figure 2A,B). The TCR repertoire in these two locations was not significantly different when comparing the number of clonotypes, clonality or the percentage of high frequency clones (online supplementary figure 2A–C). The repertoire in these two locations could be very similar in some patients and very different in others. Indeed, the Morisita-Horn index, which quantifies how similar two repertoires are, fluctuated from close to zero (completely dissimilar) to close to one (completely identical) (figure 2C).

We next compared the T cell repertoire from the altered mucosa at the time of surgery to the one at time of postoperative endoscopy 6 months later (M6). Similarly to the comparison between M0 and MI, the Morisita-Horn index calculated between M0 and M6 varied greatly within the cohort (figure 2C). Importantly, we could see a correlation between Morisita-Horn indexes comparing the two different locations (M0-MI) and the stability of the repertoire over time (M0-M6) (Spearman, $r^2=0.2864$; $p<0.0001$) (figure 2D). This indicated that if the repertoire was similar between the inflamed part (M0) and the proximal resection margin (MI), it was also similar 6 months later (M6).

Increased proportion of clonal expansions in the ileal mucosa is associated with smoking

The variability in mucosal TCR repertoires of patients could be related to diverse disease characteristics, including natural history, environmental factors, disease duration and phenotype, as well as drug exposure. We assessed the mean proportion of high frequency clones in the whole repertoire according to baseline characteristics. Only active smoking at time of surgery was significantly associated with an increased proportion of clonal expansions as compared with non-smoking (25.9% vs 17.9%, p=0.02) (figure 3A). Furthermore, the Simpson index was significantly increased in active smokers, demonstrating a reduced diversity of the TCR repertoire among smokers (figure 3B). These results suggest that smoking behaviour have an impact on the T cell compartment in the intestinal mucosa of patients with CD.

We next tested if the stability of the TCR repertoire in the mucosa could be associated with smoking. The Morisita-Horn index comparing M0 and M6 was significantly higher in smokers as compared with non-smokers (p=0.004) (figure 3C), suggesting that smoking was also associated with persistence of the TCR repertoire over time.

Predictive value of clonal expansion at time of surgery on postoperative endoscopic recurrence

The proportion of high frequency clones at baseline was significantly higher in patients with recurrence at postoperative endoscopy (Rutgeerts score >i0) compared with patients with no recurrence (23.2% vs 13.8%, p=0.01) (figure 4A). Of note, patients with high frequency clones were mainly at risk of moderate endoscopic recurrence (i1 and i2) (online supplementary figure 3). Also, the Simpson index was significantly increased at baseline in patients with a postoperative recurrence, demonstrating that a lower diversity was associated with postoperative endoscopic recurrence (p=0.003) (figure 4B). These results show that increased proportion of high frequency clones in the intestinal mucosa at time of surgery is associated with an increased risk of disease recurrence.

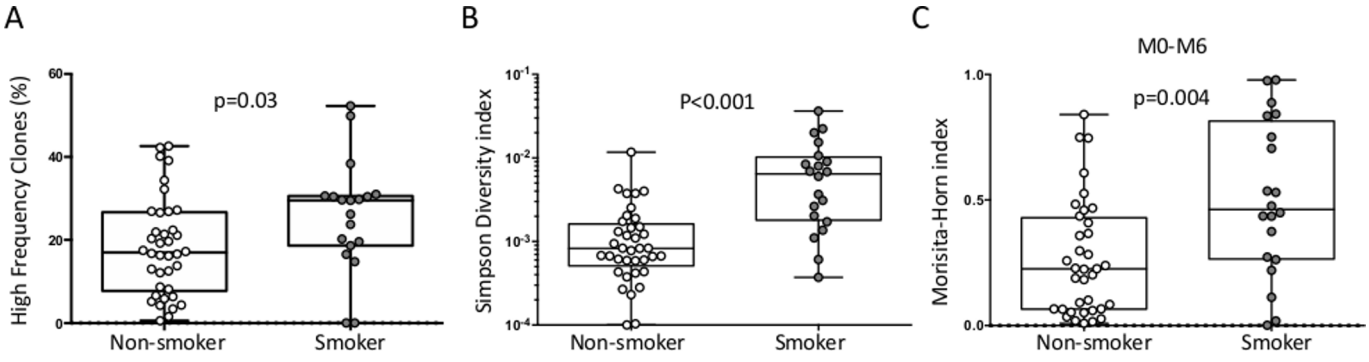

Figure 3 TCR repertoire diversity at time of surgery and persistence over time is correlated with smoking behaviour. (A) Comparison of the Simpson diversity index of TCR repertoires stratified according to smoking behaviour (Mann-Whitney U test). (B) Percentage represented by high frequency clones within the whole TCR repertoire (as in figure 1E) of mucosal ileal biopsies from patients at time of surgery stratified according to smoking behaviour as in figure 2A (Mann-Whitney test). (C) Comparison of Morisita-Horn indexes of repertoires at M0–M6 stratified according to smoking behaviour (Mann-Whitney U test). TCR, T cell receptor.

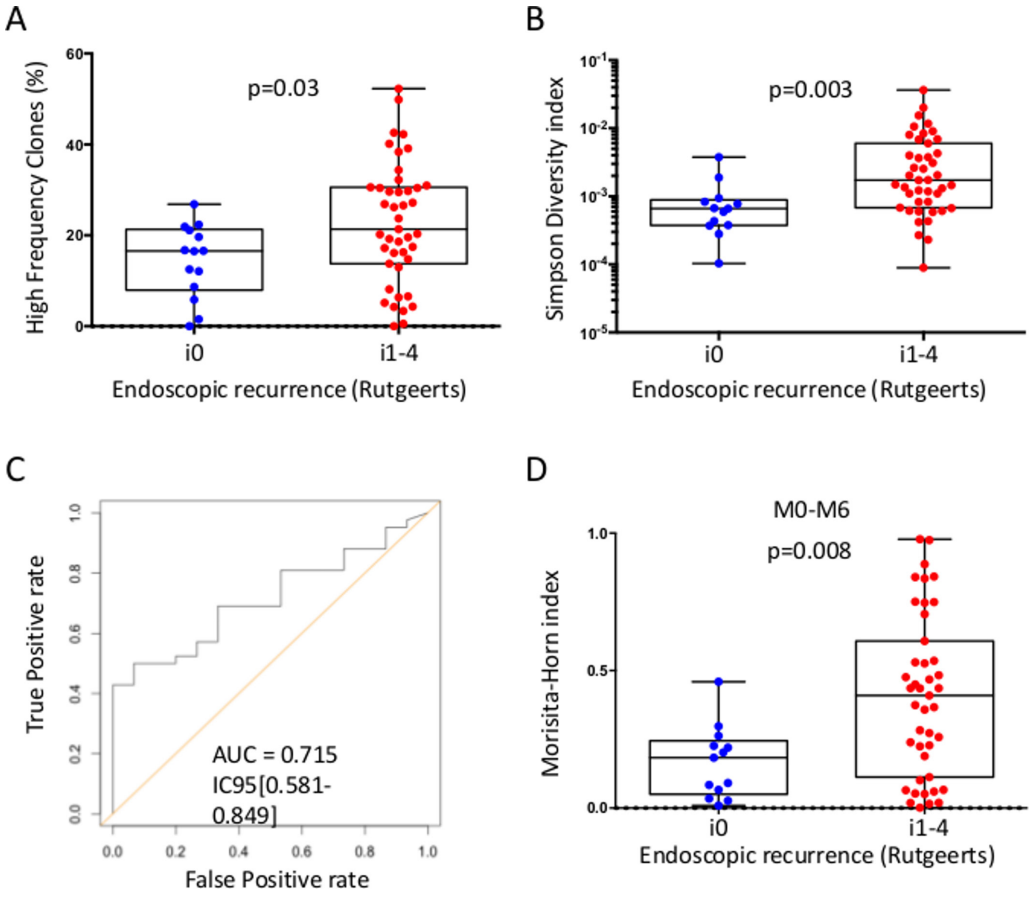

Figure 4 TCR repertoire diversity at time of surgery and persistence over time is correlated with endoscopic recurrence. (A) Comparison of the Simpson diversity index of TCR repertoires stratified according to the endoscopic recurrence Rutgeerts score (i0: no recurrence; i1-4: recurrence) determined 6 months after (Mann-Whitney test). (B) Percentage represented by high frequency clones within the whole TCR repertoire (as in figure 1E) of mucosal ileal biopsies from patients at time of surgery stratified according to recurrence as in figure 3A (Mann-Whitney test). (C) ROC curve for the prediction of postoperative recurrence according to TCR repertoire clonality (high frequency clones percentage). (D) Comparison of Morisita-Horn indexes of repertoires at M0–M6 stratified according to recurrence at M6 according to the endoscopic Rutgeerts score (i0: no recurrence; i1–4: recurrence) (Mann-Whitney U test). ROC, receiver operating characteristic; TCR, T cell receptor.

ROC curve analysis of the percentage of high frequency clones to predict endoscopic recurrence determined an AUC of 0.715 (95% CI 0.581 to 0.849) (figure 4C). After an internal validation with a bootstrap sample procedure of 300 iterations, the AUC was at 0.69 (95% CI 0.54 to 0.83). We selected 26.8% of high frequency clones as a remarkable cut-off for which sensitivity and specificity were of 43% and 100%, while negative and positive predictive values were of 38.5% and 100%, respectively. These results show that proportion of high frequency clones in the intestinal mucosa at time of surgery has a good predictive value on the risk of disease recurrence. Again, patients with high clonality (high frequency clones >26.8%) were more frequently smokers compared with those with a low clonality (61% vs 23%, p=0.05) (table 2).

The Morisita-Horn index comparing the repertoires at M0 and M6 was significantly higher in patients who had a recurrence (p=0.008) (figure 4D), indicating that the persistence of the T cell repertoire over time in CD was associated with disease recurrence.

High clonality is linked to CD8 T cells

We divided our cohort according to the percentage of high frequency clones, defining 'high clonality' and 'low clonality' as above or below 26.8%, respectively, and compared mucosal transcriptomic analysis. We could segregate them with 300 significantly regulated genes (figure 5A). In the low clonality subgroup, we observed an increased inflammatory signature indicated by

Table 2 Comparison of baseline characteristics in patients with high clonality and low clonality

	Low clonality	High clonality	
	n=39 (68%)	n=18 (32%)	P value
Smoking, n (%)	9 (23)	11 (61)	0.005
Men, n (%)	20 (51)	5 (28)	0.10
Previous resection, n (%)	6 (15)	5 (28)	0.30
Surgical indication, n (%)			0.30
Stricturing disease	22 (56)	13 (72)	
Penetration disease	17 (44)	5 (28)	
Disease phenotype, n (%)			0.30
B3	20 (51)	6 (33)	
B1 and B2	19 (49)	12 (67)	
Anti-TNF <3 months, n (%)	15 (38)	9 (50)	0.30
Thiopurines <3 months, n (%)	13 (33)	5 (28)	0.70
Extraintestinal manifestations, n (%)	7 (18)	2 (11)	0.70

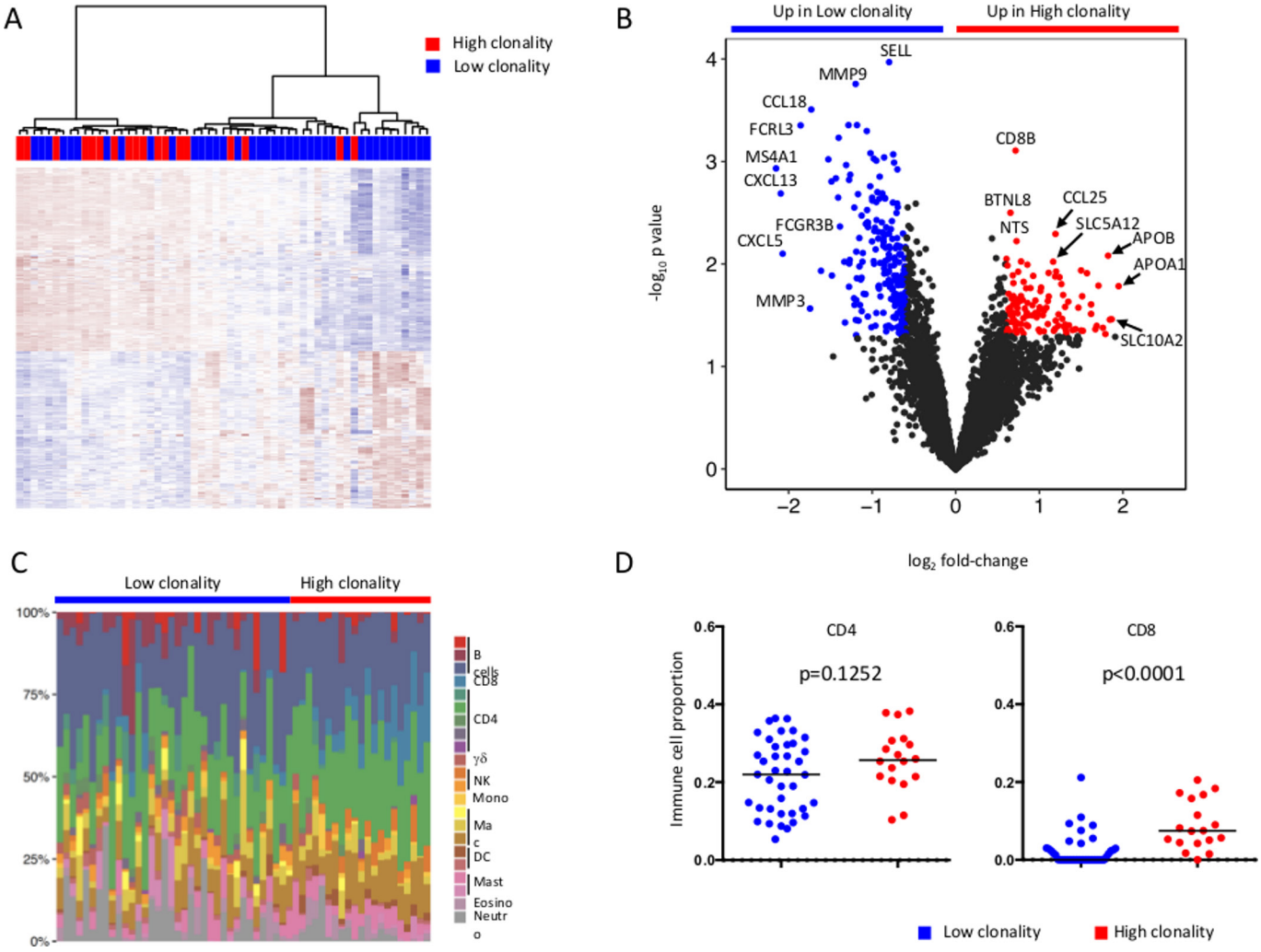

Figure 5 Transcriptomic analyses reveal a CD8 signature in high clonality patients. (A) Hierarchical clustering of differentially expressed genes observed in the inflamed tissue at time of surgery (M0) between patients with 'high clonality' (red) and 'low clonality' (blue). (B) Volcano plot of significantly differentially expressed genes upregulated in high versus low clonality patients. (C) CIBERSORT deconvolution analysis of the transcriptomic data at M0 showing the proportion of major immune cells inferred from gene expression in the mucosa of 57 patients sorted as high (red) or low clonality (blue). (D) Inferred CD4 and CD8 T cell proportions in patients with high (red) or low clonality (blue) (Mann-Whitney test).

increased expression of calprotectin (S100A8) and neutrophil chemoattractant (CXCL5) (figure 5B), associated with pathways involved in inflammatory responses (online supplementary figure 4). Interestingly, analysis of the genes that were significantly up regulated in the high clonality subgroup revealed the CD8-BETA gene as the most significantly differentially expressed between the two subgroups (figure 5B), suggesting an association between high clonality and cytotoxic CD8 T cell expansions.

Gene expression is influenced by cell distribution. Deconvolution analysis of the transcriptome showed different proportion of immune cell types in high versus low clonality patients (figure 5C). There was a significant increase in the inferred proportion of CD8 T cells in high compared with low clonality patients (figure 5D). In contrast, neutrophils and M1 macrophages were increased in low clonality patients (online supplementary figure 5). The data demonstrate that the disease could be driven by different cell types and pathways when comparing high and low clonality patients.

We also sorted T cells from the intestinal mucosa in a subgroup of patients and compared the TCR repertoire of the whole intestinal mucosa with the repertoire of sorted CD4 and CD8 T cells for each patient (figure 6A). Sorted CD8 T cells showed increase proportion of high frequency clones compared with CD4 T cells (figure 6B, online supplementary figure 6). The repertoires from whole biopsies were more similar to sorted CD8 T cell repertoire than CD4 T cell (figure 6C). These results indicate that most of the clonal expansions associated with postoperative disease recurrence in CD were within the CD8 T cell lineage.

DISCUSSION

T cells play a key role in the physiopathology of CD. Early studies highlighted the potential modification of the T cell repertoire in the context of CD.[12 14 15] We previously reported that clonal expansions might persist over time in the mucosa of patients with active uncontrolled CD.[16] In this study, we demonstrate for the first time that the presence and persistence of T cell clonal expansions are associated with smoking and postoperative endoscopic recurrence.

The T cell repertoire in the intestinal mucosa of patients with CD displays several thousand unique TCR sequences in a single biopsy, a large portion of which had low frequencies (under 0.01% of the total TCR). However, the most abundant clones could represent a significant proportion of the whole repertoire.

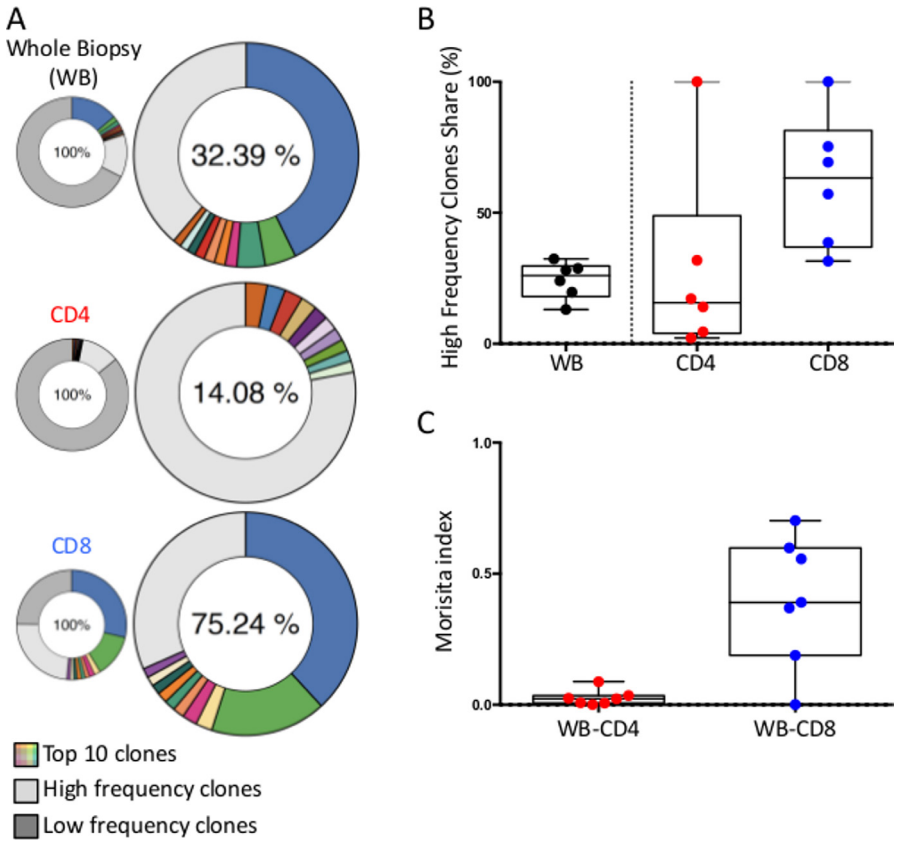

Figure 6 TCR repertoire in whole intestinal biopsies and sorted T cell population shows major expanded clones in CD8 T cell subset. (A) Donut plots showing the TCRB repertoire analyses from sequencing data of biopsies sampled at time of surgery (M0; whole biopsy) compared with sorted CD4 and CD8 T cell populations from the same surgical specimen of one representative patient. The repertoire is depicted as in figure 1B. (B) Percentage represented by high frequency clones (as in figure 1B) within the whole biopsy TCR repertoire (WB) or within the whole CD4 and CD8 repertoire. (C) Analyses of TCR repertoire similarities for each patient by calculated Morisita index between repertoire of the whole biopsy and the CD4 (WB-CD4) or the CD8 (WB-CD8) population. TCR, T cell receptor.

In this study, we show that increased proportion of T cell clonal expansions in the mucosa is associated with postoperative endoscopic recurrence of the disease. In fact, most of the patients with major clonal expansions at the time of surgery had a postoperative endoscopic recurrence, suggesting that they may represent a subgroup with a poor outcome.

In this cohort of patients with CD, with multiple paired samples from the surgical specimen and from postsurgical endoscopy at 6 months, we demonstrate that the T cell repertoire can be extremely stable along the intestinal mucosa and across different time points. This stability is associated with a poor outcome and suggests that the persistence of clonal expansion may be a causal factor of disease recurrence after surgery.

While some expanded clones were found in the CD4 T cell subset, we show that the majority of expanded clones in the intestinal mucosa are from the CD8 T cell lineage. These observations suggest that functions such as cell-mediated cytotoxicity could be involved in the physiopathology of CD, as suggested by our previous work on CD4 T cell expressing NKG2D.[18 20]

A major finding in our study is the strong association between smoking behaviour and clonal expansions in the T cell repertoire. Indeed, we demonstrate that smokers at time of surgery have a significantly decreased T cell repertoire diversity and an increased proportion of clonal expansions. Smoking is an independent risk factor for the development of CD and is associated with recurrence after surgery and a poor response to medical therapy.[6] Until now, the mechanisms that mediate these effects remain unclear. Cigarette smoke contains thousands of components, which can potentially alter the intestinal microbiota, epithelium, vasculature and immune system. Passive smoke exposure in mice leads to increased epithelial cell apoptosis, inflammation and subsequent T lymphocyte recruitment in the intestine.[21] Interestingly, T cell oligoclonal expansions are found within the lungs of patients with chronic obstructive pulmonary disease (COPD), and the number of CD8 T cells found in the lungs of patients with COPD correlates with disease severity. In a mouse model of COPD, it was shown that chronic cigarette smoke exposure causes oligoclonal expansions of lung CD8+ T cells.[22] It is interesting to note that the efficacy of thiopurines in preventing postoperative recurrence, drugs that are known to inhibit T cell proliferation, is mainly observed in smokers.[23 24] We also found a downregulation of neutrophil-associated genes in the group of patients with high clonal expansions. Interestingly, it has been shown that smokers with CD have a significant reduction in mucosal interleukin (IL)-8 levels, a cytokine involved in neutrophil recruitment.[25] The exact mechanism leading to clone expansion in the intestinal tissue of smokers remains to be elucidated. One hypothesis could be that some components from cigarette smoke induce the proliferation of antigen specific T cells, through direct effects on T cells and/or on antigen presentation.

Several studies have reported that patients could share some identical CDR3 sequences.[15 26–28] In our study, we could find such sequences shared by few patients but not the presence of previously reported specific CDR3 sequences. We also observed a very modest increase in the interindividual similarity index among patients with CD compared with non-IBD controls. Importantly, none of the common sequences were highly expanded in more than one patient. For this reason, we conclude that these identical CDR3 sequences shared by several patients do not play a key role in the development and persistence of the disease.

In conclusion, we demonstrate that smoking behaviour is associated with specific changes in the T cell repertoire characterised by reduced repertoire diversity and an increased proportion of T cell clonal expansions. The presence of these clonal expansions in the mucosa at time of surgery is associated with a high risk of postoperative endoscopic recurrence. These expanded T cell clones persist overtime and may play a major role in the disease progression. Altogether, our data may help to define a subset of patients, characterised by a higher proportion of smokers and a T cell-driven pathogenesis. This subset of patients could be potentially better treated with drugs targeting T cell driven pathways.

Author affiliations
[1]Department of Gastroenterology, Hopital Saint Louis, Paris, France
[2]INSERM U1160, Hôpital Saint-Louis, Paris, France
[3]Hospital Clinic Barcelona, IDIBAPS, CIBERehd, Barcelona, Spain
[4]Service de Gastroentérologie, MICI et Assistance Nutritive, Hôpital Beaujon, Clichy, France
[5]Department of Gastroenterology, Lyon Sud Hospital, Hospices Civils de Lyon, Pierre Benite, Lyon, France
[6]Gastroenterology Department, University Hospital Estaing, Clermont-Ferrand, France
[7]Department of Gastroenterology, Hopital Claude huriez, Lille, France
[8]Hepato-Gastroenterology Department, CHU Amiens, Amiens, France
[9]Department of Gastroenterology, Hopital Saint-Antoine, Paris, Île-de-France, France
[10]Service de Gastroentérologie, MICI et Assistance Nutritive, Hopital Beaujon, Clichy, France
[11]M2iSH, UMR Inserm U1071, USC INRA 2018, Université d'Auvergne, Clermont Ferrand, France
[12]Department of Gastroenterology, Hopital Saint-Antoine, Paris, Île-de-France, France

Correction notice This article has been corrected since it published Online First. The author affiliation details have been corrected.

Acknowledgements We thank our Institute technological platform for their help on transcriptomic and cytometry analyses. We are grateful to Dr Purnima Bhat for the English review of our manuscript.

Collaborators REMIND Study group (REMIND Study Group, France): Pierre Cattan; Mircea Chirica; Nicolas Munoz-Bongrand; Hélène Corte; Jean-Marc Gornet; Clotilde Baudry; Nelson Lourenco; Mariane Maillet; My-Linh Tran-Minh; Andrée Nisard; Laurent Beaugerie; Anne Bourrier; Isabelle Nion-Larmurier; Cecilia Landman; Sylvie Rajca; Elodie Quevrain; Loic Brot; Najim Chafai; Jeremie Lefevre; Emmanuel Tiret; Magali Svreck; Nadia Hoyau-Idrissi; Jean François Flejou; Nathalie Guedj; Yves Panis; Leon Magiorri; Mariane Feron; Yoram Bouhnik; Olivier Corcos; Dominique Cazals-Hatem; Xavier Dray; Ulrika Chaput; Philippe Marteau; Gilles Bommelaer; Marion Goutte; Emilie Vazeille; Michael Rodrigues; Pierre Sauvanet; Pierre Desreumaux; Maria Nachury; Coralie Sommeville; Jean-Louis Dupas; Franck Brazier; Denis Chatelain; Christophe Attencourt; Charles Sabbagh; Martine Leconte; Gilles Boschetti; Bernard Flourié; Yves François; Eddy Cotte; Anne-Laure Charlois; Peggy Falgon; Driffa Moussata; Marion Chauvenet; Sarah Boyer; Xavier Hebuterne; Jérome Filippi; Paul Hofmann; Madeleine Bezeault; Carole Margalef; and Patricia Detre.

Contributors MA and LLB conceived and designed the project, interpreted data and wrote the manuscript. MA, CA, KP and LLB analysed TCR sequencing data. AMC and AS provided analysing scripts for TCR sequencing data. MA and CA organised and analysed clinical parameters. MN, VC and KP performed transcriptomic analyses. LLB, HB and VC performed cell isolation experiments. CS, SN, AB, BP, MF, HS, XT, NB and PS provided essential materials and were involved in discussions. REMIND investigators provided IBD patient samples and ethical approval for the project.

Funding This study was supported by grant from the Helmsley Charitable Trust, the Association Francois Aupetit and MSD France laboratories.

Competing interests MA received honoraria from Abbvie, MSD, Janssen, Takeda, Pfizer, Novartis, Ferring, Tillots, Celgene and Genentech/Roche. SN received honoraria from MSD, Abbvie, Takeda, Janssen, HAC Pharma, Tillots, Ferring and Novartis. AB received honoraria from MSD, Abbvie, Ferring, Takeda, Vifor Pharma, Sanofi-Aventis, Hospira and Janssen. BP received honoraria from AbbVie, MSD, Takeda, Janssen, Bioagaran and Ferring. MF received honoraria from AbbVie, MSD, Takeda, Janssen, Pfizer, Ferring and Boehringer. HS received unrestricted study grants: Danone, Biocodex and Enterome; board membership, consultancy, or lecture fees: Carenity, Abbvie, Astellas, Danone, Ferring, Mayoly Spindler, MSD, Novartis, Roche, Tillots, Enterome, Maat, BiomX, Biose, Novartis and Takeda; cofunder of Nextbiotix. XT received honoraria from Abbvie, MSD, Takeda, Ferring, Norgine and Janssen. PS received honoraria from Takeda, MSD, Biocodex, Ferring and Abbvie, and non-financial support from Takeda. CA, MN, HB, VC, AMC, AS, CS and NB declare no competing interest.

Patient consent for publication Obtained.

Ethics approval The ethical committee of Hospital Saint-Louis approved this study (CPP#2009/17).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Torres J, Mehandru S, Colombel J-F, *et al*. Crohn's disease. *The Lancet* 2017;389:1741–55.
2 Dalal SR, Chang EB. The microbial basis of inflammatory bowel diseases. *J Clin Invest* 2014;124:4190–6.
3 McGovern DP, Kugathasan S, Cho JH. Genetics of Inflammatory Bowel Diseases. *Gastroenterology* 2015;149:1163–76.
4 Liu TC, Stappenbeck TS. Genetics and Pathogenesis of Inflammatory Bowel Disease. *Annu Rev Pathol* 2016;11:127–48.
5 Cosnes J, Carbonnel F, Beaugerie L, *et al*. Effects of cigarette smoking on the long-term course of Crohn's disease. *Gastroenterology* 1996;110:424–31.
6 Parkes GC, Whelan K, Lindsay JO. Smoking in inflammatory bowel disease: impact on disease course and insights into the aetiology of its effect. *J Crohns Colitis* 2014;8:717–25.
7 Rutgeerts P, Geboes K, Vantrappen G, *et al*. Natural history of recurrent Crohn's disease at the ileocolonic anastomosis after curative surgery. *Gut* 1984;25:665–72.
8 Auzolle C, Nancey S, Tran-Minh ML, *et al*. Male gender, active smoking and previous intestinal resection are risk factors for post-operative endoscopic recurrence in Crohn's disease: results from a prospective cohort study. *Aliment Pharmacol Ther* 2018;48:924–32.
9 Buisson A, Chevaux JB, Allen PB, *et al*. Review article: the natural history of postoperative Crohn's disease recurrence. *Aliment Pharmacol Ther* 2012;35:625–33.
10 Giuffrida P, Corazza GR, Di Sabatino A. Old and New Lymphocyte Players in Inflammatory Bowel Disease. *Dig Dis Sci* 2018;63:277–88.
11 Neurath MF. Cytokines in inflammatory bowel disease. *Nat Rev Immunol* 2014;14:329–42.
12 Probert CS, Saubermann LJ, Balk S, *et al*. Repertoire of the alpha beta T-cell receptor in the intestine. *Immunol Rev* 2007;215:215–25.
13 Matsuda JL, Gapin L, Sydora BC, *et al*. Systemic activation and antigen-driven oligoclonal expansion of T cells in a mouse model of colitis. *J Immunol* 2000;164:2797–806.
14 Probert CS, Chott A, Turner JR, *et al*. Persistent clonal expansions of peripheral blood CD4+ lymphocytes in chronic inflammatory bowel disease. *J Immunol* 1996;157:3183–91.
15 Doorenspleet ME, Westera L, Peters CP, *et al*. Profoundly Expanded T-cell Clones in the Inflamed and Uninflamed Intestine of Patients With Crohn's Disease. *J Crohns Colitis* 2017;11:831–9.
16 Camus M, Esses S, Pariente B, *et al*. Oligoclonal expansions of mucosal T cells in Crohn's disease predominate in NKG2D-expressing CD4 T cells. *Mucosal Immunol* 2014;7:325–34.
17 Robins HS, Campregher PV, Srivastava SK, *et al*. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* 2009;114:4099–107.
18 Pariente B, Mocan I, Camus M, *et al*. Activation of the receptor NKG2D leads to production of Th17 cytokines in CD4+ T cells of patients with Crohn's disease. *Gastroenterology* 2011;141:217–26.
19 Newman AM, Liu CL, Green MR, *et al*. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 2015;12:453–7.
20 Allez M, Tieng V, Nakazawa A, *et al*. CD4+NKG2D+ T cells in Crohn's disease mediate inflammatory and cytotoxic responses through MICA interactions. *Gastroenterology* 2007;132:2346–58.
21 Verschuere S, Bracke KR, Demoor T, *et al*. Cigarette smoking alters epithelial apoptosis and immune composition in murine GALT. *Lab Invest* 2011;91:1056–67.
22 Motz GT, Eppert BL, Sun G, *et al*. Persistence of lung CD8 T cell oligoclonal expansions upon smoking cessation in a mouse model of cigarette smoke-induced emphysema. *J Immunol* 2008;181:8036–43.

23 Mowat C, Arnott I, Cahill A, *et al*. Mercaptopurine versus placebo to prevent recurrence of Crohn's disease after surgical resection (TOPPIC): a multicentre, double-blind, randomised controlled trial. *Lancet Gastroenterol Hepatol* 2016;1:273–82.
24 Ben-Horin S, Goldstein I, Fudim E, *et al*. Early preservation of effector functions followed by eventual T cell memory depletion: a model for the delayed onset of the effect of thiopurines. *Gut* 2009;58:396–403.
25 Sher ME, Bank S, Greenberg R, *et al*. The influence of cigarette smoking on cytokine levels in patients with inflammatory bowel disease. *Inflamm Bowel Dis* 1999;5:73–8.
26 Chapman CG, Yamaguchi R, Tamura K, *et al*. Characterization of T-cell Receptor Repertoire in Inflamed Tissues of Patients with Crohn's Disease Through Deep Sequencing. *Inflamm Bowel Dis* 2016;22:1275–85.
27 May E, Lambert C, Holtmeier W, *et al*. Regional variation of the alphabeta T cell repertoire in the colon of healthy individuals and patients with Crohn's disease. *Hum Immunol* 2002;63:467–80.
28 Holtmeier W, Hennemann A, May E, *et al*. T cell receptor delta repertoire in inflamed and noninflamed colon of patients with IBD analyzed by CDR3 spectratyping. *Am J Physiol Gastrointest Liver Physiol* 2002;282:G1024–G1034.