Investigating checkpoint inhibitor-induced gastrointestinal disease : comparisons with idiopathic ulcerative colitis and health

Tarun Gupta

Linacre College

University of Oxford

A Thesis submitted for the degree of

Doctor of Philosophy in Clinical Medicine

Michaelmas 2022

To Anuradha and Jayant Gupta

who dedicated themselves to giving me an education

Abstract

In the last decade, immunotherapy with anti-PD1 and anti-CTLA4 drugs has revolutionized the treatment of metastatic cancer. A significant proportion of patients given these checkpoint inhibitors develop off-target inflammation in the colon (checkpoint inhibitor-induced colitis, CC). Ulcerative colitis (UC) is a well-characterized multisystem disorder, propagated by a combination of barrier defects, immune dysfunction and a disordered microbiome, all on a background of genetic susceptibility. The underlying mechanisms of both checkpoint inhibitor-induced colitis and ulcerative colitis remain poorly understood, with a significant proportion of patients failing therapy. Using clinical data, next-generation single-cell RNA sequencing, unbiased spatial transcriptomics and organoid model systems, we elucidate clinically relevant patterns of cellular behaviour across blood, epithelium, stroma and immune populations. We describe and characterize novel disease-specific cell populations and cellular interactions in tissue 'microdomains' with functional relevance. Our work has implications for therapeusis as well as puts forward a new approach for identifying drug-bound cells in vivo.

Contents

Table of Contents

List of Tables

List of Figures

Supplementary Figures

Abbreviations

Ab	-	Antibody
Abs	-	Absorptive epithelial cells 1
Abs2	-	Absorptive epithelial cells 2
ADT	-	Antibody-derived Tags
B	-	B-cells
BAC	-	Bacterial artificial chromosome
BCR	-	B-cell receptor
BSA	-	Bovine Serum Albumin
Casp-3	-	Cleaved caspase 3
CC/ CC Colitis	-	Checkpoint inhibitor induced colitis
CC_I	-	Checkpoint inhibitor-induced colitis, inflamed
CC_NI	-	Checkpoint inhibitor-treated patients, no colitis
CD	-	Crohns' disease
CITE-seq	-	Cellular indexing of transcriptomes and epitopes
CM	-	Conditioned medium for organoids
CMV	-	Cytomegalovirus
CS10	-	Cryostor Medium 10 (see Materials)
CT-E	-	Crypt-top enterocytes
CTLA4	-	Cytotoxic T-Lymphocyte Associated protein 4
Cyc	-	Cycling cells
CyTOF	-	Mass Cytometry Time-of-Flight analysis
DAB	-	3,3′-Diaminobenzidine
DAPI	-	4',6 - diamidino-2-phenylindole
DEGs	-	Differentially expressed genes
DM	-	Differentiation medium for organoids
DMEM	-	Dulbecco's Modified Eagle Medium
DMSO	-	Dimethyl sulfoxide
DP	-	Double positive CD4+ CD8+ T cells
DSS	-	Dextran sulphate sodium
DTT	-	Dithiothreitol
Duo/Dual CC_I	-	Combination therapy with aCTLA4 (ipilimumab) and aPD1 (Nivolumab)
E-cad	-	E-cadherin
EDTA	-	Ethylenediaminetetraaceticacid
EEC	-	Enteroendocrine cells
End	-	Endothelial cells
Ent	-	Enterocytes
Exh	-	CD8+ HAVCR2+ exhausted T cells
F.Ef	-	FGFBP2+ effector T cells
FACS	-	Fluorescence activated cell sorting
FCS	-	Foetal calf serum
FFPE	-	Formalin-fixed paraffin embedded
FMO	-	Fluorescence minus one sample
FMT	-	Faecal Microbial Transplant
G-Ef1	-	Granzyme K+ T effector cells 1
G-Ef2	-	Granzyme K+ T effector cells 2
GEX	-	Gene expression

GO analysis	-	Gene ontology resource pathway analysis
Gob	-	Goblet cells
GZMK	-	Granzyme K
H&E	-	Haematoxylin and Eosin
HBSS	-	Hank's Balanced salt solution
HC	-	Healthy controls
HEPES	-	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hIL26-Tg	-	Humanized IL-26 producing Transgenic mice
HTO	-	Hash-tagged oligos
HTS	-	High throughput sampler for FACS sample acquisition
IBD	-	Inflammatory bowel disease
IELs	-	Intra-epithelial lymphocytes
IF	-	Immunofluorescence
IFNg	-	Interferon gamma
IFN-R	-	Interferon response B cells
IgA	-	Immunoglobulin A
IgG	-	Immunoglobulin G
IHC	-	Immunohistochemistry
IL-21	-	Interleukin 21
IL-23	-	Interleukin 23
IL-23R	-	Interleukin 23 Receptor
IL-26	-	Interleukin 26
IMC	-	Imaging mass cytometry
IRAE	-	Immunotherapy related adverse events
ISH	-	In-situ Hybridization
M1	-	M1 type Macrophages
M2	-	M2 type Macrophages
MACS	-	Magnetic activated cell sorting
MAIT	-	Mucosal-associated invariant T cells
MF	-	Myofibroblasts
Mono CC_I	-	Monotherapy aPD-1 treatment (Nivolumab or Pembrolizumab)
MYL	-	Myeloid cells
NK cells	-	Natural Killer cells
NSAIDs	-	Non-steroids anti-inflammatory drugs
NSCLC	-	Non-small cell lung cancer
Nv CD8	-	Naïve CD8 T cells
PBMC	-	Peripheral blood mononuclear cells
PBS	-	Phosphate buffered saline
PCA	-	Principal component analysis
PD-1	-	Programmed Death 1
Pen-Strep	-	Penicillin-Streptomycin
Per	-	Pericytes
PL	-	Plasma cells
qRT-PCR	-	Quantitative real-time polymerase chain reaction
RIN	-	RNA integrity number
R-L	-	Receptor-Ligand
RPMI	-	Roswell Park Memorial Institute (RPMI) 1640 medium
S1	-	Stromal 1
S2	-	Stromal 2

S3	-	Stromal 3
S4	-	Stromal 4
sc-RNAseq	-	Single cell RNA sequencing
SNPs	-	Single Nucleotide Polymorphisms
ST	-	Spatial transcriptomics
TA	-	Transit-amplifying epithelium
Tc17	-	CD8 T-helper type 17 cells
TCM	-	Central Memory T Cells
TCR	-	T-cell receptor
TF	-	Transcription factor
Tfh	-	T- follicular helper
Th1	-	CD4 T-helper type 1 cells
Th17	-	CD4 T-helper type 17 cells
Th2	-	CD4 T-helper type 2 cells
TNFa	-	Tumour Necrosis Factor alpha
Treg	-	Regulatory CD4+ T cells
TRM	-	Resident Memory T cells
t-SNE	-	t-distributed stochastic neighbour embedding
UC /UC colitis	-	Ulcerative colitis
UC_I	-	Ulcerative colitis, inflamed
UC_NI	-	Ulcerative colitis, matched non-inflamed biopsy areas
UMAP	-	Uniform manifold approximation and projection
UMI	-	Unique molecular identifier
WT	-	Wild-type (mice)
ZNF683+	-	ZNF683+ Tissue resident memory cells

Acknowledgements

Attributions

This work would not have been possible without a collaborative effort. I was supervised by Professor Alison Simmons and Dr Oliver Brain.

Chapter 1

I (T.G.) worked together with Vincent Ting Fung Cheung (V.T.F.C.) to extract patient data for the Oxford cohort from electronic records, splitting patients equally between us. Liverpool data was sent to me by Anna Olsson-Brown (A.O.B), with the study being conceived by Oliver Brain (O.B.). Histopathological analysis was performed by Eve Fryer (E.F.) and Elena Collantes (E.C.); endoscopy report analysis was performed by JAG-accredited endoscopists O.B. and V.T.F.C.

I organised and collated the data into final spreadsheets, conceived and performed the analysis, determined conclusions and generated all the figures, under the supervision of O.B. This work is published, I am joint first author[1].

Chapter 2

The study was conceived by Daniele Corridoni (D.C.) and Alison Simmons (A.S.). D.C. generated the original CD8 T cell cohort (sc-RNAseq experiment), and conceived and performed the hIL26-Tg mouse experiment in Japan together with our collaborators in Juntendo University.

I (T.G.) conceived, optimized and performed the CD8 T cell CITE-seq experiment, including library preparation. I also extracted RNA from mouse tissues, assessed RNA quality and sent these for library prep and sequencing. I conceived, performed and analysed all the IHC, ISH, FACS and organoid experiments and generated all the figures from this data.

All the single cell and bulk transcriptome data was analysed by Agne Antanaviciute (A.A.A.) in discussion with and guided by D.C., A.S. and me. A.A.A. also generated all the figures derived from this data type. This work is published, I am joint first author[2].

Chapter 3

I (T.G.) conceived this work, supervised by Alison Simmons (A.S.) and Oliver Brain (O.B.). O.B. conceived and developed the PRISE study, which I supported, as detailed in methods. Anna Aulicino (A.A.) generated the CD45+ data from ulcerative colitis and health, whereas I generated the CD45+ data from checkpoint inhibitor induced colitis. A.A., Ana Sousa Geros (A.G.) and I generated the epithelial and stromal/CD45 dataset together, including optimization and library preparation. I conceived the experiments for and generated the remainder of the data (CD3 sc-RNAseq data, IF staining and imaging data, organoid data, co-culture model development, FACS data and spatial transcriptomic data), as well as generating the figures deriving from that. I consented and collected samples from the patients included in the study (assisted by V.T.F.C. and Lulia Al-Hillawi). Information about duration of inflammation/immunotherapy was extracted from medical records by Nicole Cianci and me.

Technical and practical advice was given by Kaushal Parikh and Marta Jagielowicz (organoids), David Fawkner-Corbett (D.F.C., Visium 10X) and Chaitanya Vuppusetty (immunofluorescence). I was assisted in slide sectioning by Marta Jagielowicz, Esther Bridges and D.F.C. Quantification of image data using Visiopharm was carried out by Nasullah Khalid Alham (N.K.A.) directed by me. Histopathological analysis was carried out by Dr Eve Fryer (E.F.) in discussion with me.

All the single-cell transcriptomic data analysis was carried out by Agne Antanaviciute (A.A.A.) in conjunction with Chloe Hyun-Jung Lee (C.L.) and Rosana Babu (R.B.) in discussion and guided by me and A.S. Spatial transcriptomic data analysis was carried out by A.A.A in discussion with me. A.A.A. and I jointly conceived of the idea for detecting bound cells. This data is being prepared for publication.

Research support

I was supported in sample collection by the Oxford GI Biobank team, led by James Chivenga and the TIP Study nursing team led by Simona Fourie.

Flow cytometry advice was provided by the Weatherall Institute of Molecular Medicine (WIMM) Cytometry lab, headed by Paul Sopp. Imaging advice was provided by the WIMM Imaging facility, led by Chris Lagerholm and Jana Koth. Sc-RNAseq advice was provided by the WIMM single-cell facility, headed by Neil Ashley and Maria Greco. Sequencing advice was provided by Timothy Rostron at the WIMM sequencing facility.

This study would not have been possible without the support and efforts of patients attending Oxford University Hospitals towards donating time and samples for medical research.

I am grateful to both Professor Alison Simmons and Dr Oliver Brain for their invaluable advice, support and guidance.

Funding

This study was supported primarily by the Lee Placito Medical Fund fellowship. Other supporting organisations included the British Research Consortium precision medicine cluster (BRC), Celgene Inc., the Medical Research Council (MRC) and the Oxfordshire Health Services Research Committee (OHSRC).

Impact of Covid-19

Due to Covid-19 contact and policy changes, there was no access permitted to the WIMM lab facilities between March 2020 and July 2020. On return to work, restrictions remained in terms of number of hours allowed and spaces available for hands-on work until September 2021. This limited the type and volume of work possible during this 18 month interval of time.

The British Society of Gastroenterology issued guidelines on research at the start of the pandemic in March 2020 which did not permit the collection of samples for medical research, citing risk to staff and patients. This remained in place until August 2020. Limited sample collection was permitted for clinically unavoidable procedures between August and December 2020. Prospective recruitment to the PRISE study was therefore halted from March 2020 to March 2021. Taken together, this 12 month period significantly impacted longitudinal sample collection, a key feature of the PRISE project, as well as research sample numbers overall.

Introduction

The human colon in health

The human colon in health fulfils multiple excretory and absorptive roles all whilst maintaining a symbiotic barrier between the microbiome and self. Its structure comprises an outer serosa continuous with the peritoneum, which encapsulates two muscularis mucosa layers, in between which is sandwiched the myenteric neural plexus. This is followed by the submucosa consisting of a rich vascular, neural and lymphatic network supported by fibroblast cells. Lining the submucosa is the mucosa, comprising the lamina propria (consisting of multiple diverse stromal and immune cells, both scattered and organised in superstructures) and the innermost rapidly dividing epithelial layer. All these are surmounted by a secreted mucus layer, which abuts the lumen containing the microbiome. Acting synergistically, these layers accomplish the absorptive, secretory and peristaltic functions in the healthy colon. The microbiome and its metabolome are not passive agents, but are now thought to play key roles in maintaining the epithelial barrier whilst also having diverse systemic effects. Although the full extent and mechanism of these processes remains unclear, they have been implicated in contexts as disparate as cancer immunology[3], cardiovascular health[4,5] and neurological function[6], and are an area of intense research interest. Consequently, understanding the interplay of immune, stromal, epithelial and microbial factors that are responsible for maintaining homeostasis is key.

Ulcerative colitis

Inflammatory bowel diseases are complex multisystem disorders that are driven by the failure of this homeostasis, and are a product of barrier failure, dysregulated immune networks and microbiome, all on a background of genetic susceptibility[7–9]. They primarily affect the young, with considerable impact on morbidity and functional status within an otherwise healthy and economically productive demographic. Their incidence is increasing and has been conflated with sociological development in industrial societies. Combined with upto 40% of patients with moderate to severe disease being either primarily resistant or developing resistance to current therapy[10], it therefore poses a compelling and urgent argument for improved therapeusis and understanding of barrier function. Multiple manifestations (e.g. Crohns' disease, Ulcerative colitis, Collagenous colitis, amongst others) have been described, distinguished on the basis of histopathological and clinical criteria. The most severe fall into two broad patterns of disease first described in the late nineteenth and early twentieth century – ulcerative colitis and crohns' disease.

Ulcerative colitis is distinguished by disease that usually affects the distal colon (comprising the rectum, sigmoid and descending colon), with more severe cases extending more proximally to cause a pan-colitis. There is typically sparing of the small bowel and stomach, with inflammation limited to the mucosa (rather than penetrating the deeper layers of the bowel wall). The disease tends to follow a remitting-relapsing pattern, and can manifest with extra-intestinal inflammation affecting the skin, joints and eyes, all of which are thought to be driven by colonic disease, although the mechanism remains unclear[11,12].

Novel insights into health and disease

Foremost among novel technologies applied to understanding ulcerative colitis was droplet-based single-cell RNA sequencing. First described in 2015[13], this was a method that utilized microfluidic technology combined with primer-based barcoding to 'tag' and recover mRNA transcriptomic information from individual cells. When combined with advances in sequencing, it allowed high-throughput phenotyping of cells from disease and health. Adaptations such as including oligonucleotide-tagged antibodies for proteins and specific primers for T- and B-cell receptor sequences yielded proteomic and lineage maps.

Another methodology that has proved extremely valuable was the development of organoid mini-gut model systems in 2009[14]. Utilizing an optimized concentration of growth factors and extracellular matrix, it allowed adult crypt-derived stem cells to replicate indefinitely without requiring transformation, as well as allowing maturation when growth factors were withdrawn, essentially recreating the crypt *in vitro*. The approach recovered the majority of cell types within the native colon, and provided a novel approach to understanding cellular maturation as well as function[15].

Cytometry-time of flight (CyTOF) analysis proved an improvement on the number of proteins that could be simultaneously detected by traditional flow cytometry, utilizing heavy metal tagged antibodies together with single-cell mass spectrometry. This was used to characterize cells with more granularity than possible before[16].

Finally, more recently, imaging techniques such as imaging mass-cytometry (IMC, an adaptation of CyTOF) and high resolution spatial transcriptomics (ST) have emerged as the next iteration of more traditional localization techniques. These seek to phenotype cellular (and sub-cellular) interactions utilizing protein (IMC) and transcriptomic (ST) information, at high fidelity in a relatively unbiased fashion, and are at the cutting edge of providing novel insights into how groups of cells interact[17,18].

Taken together with more established techniques such as fluorescence-activated cytometry (FACS) and murine modelling of disease, considerable advancements have been made in understanding disease processes in health and ulcerative colitis.

The colonic epithelium and stroma in health

The colonic crypt comprises a tightly regulated *LGR5+* stem-cell niche at the base, progressing through transit-amplifying states in the lower crypt to a terminally differentiated states containing goblet, enteroendocrine, mature enterocyte and BEST4/OTOP2+ cells near the top[19–21].

The self-renewing stem cell niche (identified by *LGR5+*) is maintained through the synergistic actions of Wnt, Noggin and R-spondin growth factor gradients maintained by the stromal population[22]. It is relatively robust, but there are indications that it is perturbed by inflammation[23].

Introduction

Although most cell types (e.g. goblet cells, enteroendocrine cells) are preserved across murine and human colons and perform similar functions, there are small, but significant differences between murine and human colonic crypts (such as the absence of Paneth and Tuft cells in humans, and absence of BEST4/OTOP2 cells in mice) suggesting cells may perform overlapping and partially redundant functions, with a degree of plasticity which is revealed under certain physiological conditions. Each cell type can be identified by a characteristic transcriptomic signature, and with the possible exception of BEST4/OTOP2 cells, is thought to arise wholly from the stem cell compartment.[24,25]

Development into the diverse populations observed in the crypt occurs as these gradients decrease moving up the crypt, and some role in differentiation and nutrient supply may also be played by microbiome-dependent factors and metabolic products e.g. butyrate[26–28].

In close association with the epithelium, and responsible for maintaining the stem cell niche and supporting the epithelium, is the stroma[20,29]. This comprises diverse cell types including fibroblasts, myofibroblasts, glial cells, axons from neurons located distant from their site of action, lymphatic venules, vascular structures and a variety of immune cells that act in synergy to support and re-enforce the epithelium in ways that are still incompletely understood.

Fibroblasts comprise the predominant population in this compartment, expressing pan-fibroblast connective tissue markers such as *VIM* and *COL1A1*, distinguished from contractile myofibroblasts by lower expression of muscle-associated transcripts such as *MYH11*. Detailed characterization by scRNAseq sequencing revealed three-four subtypes, each with a distinct function[20,29].

Stromal 2/*RSPO3+* populations are thought to perform a key role in maintaining the epithelial stem cell niche. High *WNT* expression by this population, combined with R-spondin and noggin expression act to allow the stem cell niche to persist[14,30].

In contrast, Stromal 4 cells express high levels of chemokines CCL19 and CCL21, and are thought to form the nidus for lymphoid structure formation. CCL19 and CCL21 bind to CCR7 receptors expressed on a wide range of T and B cells, and are therefore thought to initiate and maintain lymphoid structure formation. By facilitating close interactions between T-follicular helper (Tfh) and maturing B cells, these improve the production of high affinity B cells and plasma cells through somatic hypermutation and affinity maturation of antibodies. Disruption of this process, accompanied by class switching from IgA to IgG, is thought to play a pathogenic role in IBD[31].

Other stromal populations are thought to have more supportive roles – e.g. stromal 1 in maintenance of the extracellular matrix, whilst stromal 3 and pericytes support vascular structures. Emerging evidence would suggest roles for neural cells, comprising neurons as well as supportive glial cells – e.g. in maintaining the epithelial stem cell niche[32]. However, these roles are as yet incompletely understood.

Epithelial and stromal dysfunction in ulcerative colitis.

Analysis of clinical outcomes provided some of the earliest evidence that epithelial regeneration (and mucosal healing) is a marker of long-term remission. Multiple lines of evidence now indicate that a disordered epithelial state, transcriptionally, epigenetically and functionally exists in inflamed ulcerative colitis, and may be responsible for perpetuating inflammation. Stromal dysfunction has also been observed, and is thought to assist with maintaining the inflammatory response whilst the epithelial barrier remains breached, but persistence of these inflammatory phenotypes is now recognised to play a role in treatment resistance[33].

Transcriptional scRNAseq of epithelial subtypes identified stereotyped changes driven by interferon and TNF across all cells. Interferon response genes such as *ISG15,* chemokines such as *CCL20,* siderophores such as *LCN2*, opsonins and immunomodulators such as *SAA1,* oxidases such as *DUOXA2* and HLA molecules such as *HLADRA* were generally upregulated, whereas other changes such as *MUC1* (upregulated, immature enterocyte cells), *SPON1* (upregulated, stem cells) and *WFDC2* (downregulated, goblets) were more population specific[19,20]. The overall effect of these appears to be a balance between increasing anti-microbial effect and disordered adaptations to the inflammatory milieu that allows persistence of inflammation. This equilibrium is clearly not maintained in UC, as we observe an increase in necrotic cell death, which in turn has been implicated in modulating the inflammatory response into a more disordered, self-perpetuating state[34]. The exact mechanisms by which the sum total of these processes results in transient inflammation that resolves (e.g. an acute infectious colitis), versus inflammation that persists (e.g. UC), remains unclear.

The development of high throughput sequencing technologies has also expanded our knowledge of gene-loci associations with disease. Multiple GWAS analyses have been carried out on IBD cohorts (across a spectrum of nationalities, with locoregional differences), that has identified several genes that increase the risk of developing UC. Many of these at-risk loci genes are highly expressed by multiple populations, particularly immune cells, but also by epithelial and stromal populations, supporting the role of intrinsic barrier defects in contributing to colitogenesis[19,24,29].

Multiple mouse models of colitis that are driven by a defect in the epithelial barrier have been developed. One of the commonest agents used is a Dextran Sodium Sulfate (DSS) model of colitis. DSS is an orally administered molecule that disrupts the tight junctions between colonic epithelial cells, inducing a colitis that has histological and transcriptomic hallmarks mimicking UC. Targeted knockout of genes involved in tight junction and connective tissue integrity are also sufficient to induce spontaneous colitis in mice.

Recent evidence suggests that an inflammatory environment may result in lasting effects on epithelium, likely through hitherto undescribed epigenetic changes. Studies that compared the transcriptome of inflamed epithelium to adjacent non-inflamed epithelium demonstrated that many of hallmarks of inflammation persisted/were also present in non-inflamed tissue, and were distinct to health[35]. Although not clear whether this represented an intrinsic predisposition towards inflammation in non-inflamed tissue, or 'burnt out' disease, this somewhat surprising finding suggested that transcriptional patterns can be imprinted.

Other data also supports the premise that inflammation has lasting effects after it has resolved. Colonoids (derived from stromal cells taken at endsocopy, retrodifferentiated into pluripotent stem cells, and then forward differentiated into epithelial-stromal co-clusters) or epithelium-derived organoids from patients with UC behave differently to those derived from healthy tissue. Despite being removed from an inflammatory environment, they continue to maintain the same transcriptional signatures of actively inflamed tissue, and also demonstrate functional impairments, such as a reduced replication ability and lower mucus production[36–38].

Taken together, these results suggest that the epithelium in UC is prone to perpetuating inflammation in the presence of physiological stressors. It remains challenging to understand whether epithelial dysfunction is the triggering event in UC, and differentiating 'normal' changes in inflammation from those that are due to an 'abnormal' response. This is due largely to the lack of an appropriate comparator.

Immune dysfunction in ulcerative colitis

Non-T cell populations

The adaptive and innate immune systems act in synergy to both maintain the barrier and repel any micro-organisms that breach the mucosal lining. Tolerogenic macrophages, T-regulatory cells, Tissue-resident memory cells and IgA-producing B and plasma cells act in concert to recognise and prevent attachment and invasion by normal commensal organisms whilst avoiding inflammation and tissue destruction.

Shortly after its clinical description as a distinct entity, ulcerative colitis (UC) was found to be amenable to treatment by steroids and aminosalycilates more than antibiotics[39]. The immunological changes in UC in nearly every subtype of cell have therefore been the subject of intense research investigation, through a variety of techniques, from immunohistological, cytometry, functional, through to most recently, single-cell technologies.

Professional antigen presenting cells, such as dendritic cells and macrophages display a distinct phenotype in inflammatory bowel disease. There is a considerable body of work that recognizes different subtypes of macrophages, including pro (M1-type) and anti inflammatory (M2-type) macrophages in mice, with putative analogous signatures identified in humans. In particular, M1-type pro-inflammatory macrophages, expressing TNFa, IL6 and IL12, are enriched in IBD, and form part of a pro-inflammatory module that has been demonstrated to confer a resistance to treatment.

Other cells of the innate immune pathway have also been implicated in the maintenance of the immune response in UC, although the evidence is less direct and has been less well characterized. Polymorphisms in *KIR* genes, which encode receptors on innate Natural Killer cells have been associated with an increased risk of developing UC[40]. They are thought to be part of an axis that results in an increased production of IL13 and possibly IL17F which contributes to epithelial damage in UC.

Other such less well characterized but implicated populations include neutrophils and B/plasma cells. Neutrophils are enriched in inflammation and correlate clinically with severity of colitis, but their contribution to colitis has been found to be pro- or anti-inflammatory, depending on the mouse/rat model of colitis. Similarly, class switching from an IgA to IgG phenotype has been observed in UC, which is thought to be pro-inflammatory as the constant region of IgG binds and activates Fc-receptor expressing antigen-presenting cells, driving an IL17 response through IL1B signalling[31], whereas the IgA molecule does not, effectively neutralizing the antigen and driving its loss in stool. Although usually part of the physiological mechanism for limiting microbial invasion during barrier disruption, an abnormally active/prolonged class switch could contribute to colitis.

T cell populations

Multiple T cell subsets have been implicated in the pathophysiology of UC, and much of the research work till date has focused on these populations, phenotyping their subsets, proportional changes in inflammation and functional behaviour using, amongst others, FACS, bulk TCR analysis, mouse models of disease and bulk transcriptomics.

Clear stereotyped changes have been described in UC inflammation – such as increases in the proportion of regulatory T cells (Tregs), T central memory cells (TCM), T follicular helper cells (Tfh), and decreases in resident T cell populations (TRMs)[41,42] across multiple studies using different methodologies.

Tregs are reduced in the peripheral blood in UC and increase in response to anti-TNFa therapy that is correlated with disease response, with multiple mouse models of disease[43] appearing to demonstrate an amelioration in colitis with an improvement in the function and number of peripheral Tregs. More recently, this has formed the basis of early adoptive autologous T cell transfer treatments in single patient studies[44] that show promise.

Th17 cells, and the IL17-IL23 axis has been implicated in the pathogenesis of IBD through transcriptomic, FACS and mouse model based analysis[45], with clinical trials demonstrating efficacy of anti-IL23R/anti-IL12R treatment (Ustekinumab) in UC[46] (UNIFI Trial).

In a seminal analysis, bulk profiling of both tissue and peripheral CD8 T cells in a cohort of UC (and Crohns') patients revealed a transcriptomic signature associated with a worse prognostic outcome. A prospective randomized trial (PROFILE) is currently underway to assess whether these patterns are modifiable and can be used to help tailor therapy for individual patients in IBD [47].

Other T cell subsets remain less well explored, but are also thought to contribute to IBD pathogenesis. T follicular helper cells (Tfh) and IL21 secretion (produced by Tfh and Th1 cells) are increased in UC[48–52], and potentially help perpetuate aberrant B cell responses, Th17 cell differentiation and Treg resistance, that all help drive disease.

Tissue resident memory (TRMs, generally CD103+) have a less clearly defined role in UC. Depending on the method of measurement employed in the study, they are relatively depleted in UC, but with certain subsets possibly increased in the lamina propria in inflammation[53]. Although evidence suggested that αEβ7/CD103+ cells were predominantly pathogenic in UC[54], drug therapy targeting cells bearing the receptor was mixed[55,56], suggesting that our understanding of which cells occupy a truly 'resident' niche, as well as their origin and fate, remains incomplete.

Finally, it was recognised that there was a considerable degree of trafficking of T cells between blood and tissue in ulcerative colitis which contributed towards inflammation. In addition to systemic inflammation markers (such as CRP and Albumin), as well as transcriptional signatures that were prognostic of response[57], therapeutic approaches which restricted trafficking of populations to the tissue by targeting the α4β7 receptor (vedolizumab) were also effective. However, as stated in the context of understanding T cell residency, our knowledge of T cell compartment dynamics in humans remains incomplete.

Immune and non-immune cell interactions driving ulcerative colitis

Increasingly, analysis of inflammation has recognised the heterogeneity of disease severity between patients and sought to sub-stratify based on those responding quickly to established anti-TNFa therapy versus those that prove more recalcitrant to treatment.

Many of these studies have identified nexi of immune and non-immune cell interactions rather than cell-type based associations – such as OSM-OSMR[58] or GIMATS module[59] that describe primary or acquired resistance to conventional treatment, used as a proxy for more severe disease. This demonstrates how many of the cytokine axes seen in IBD are driven by diverse cell types, highlighting the importance of understanding cellular interactions as a whole as well as changes within particular subsets.

This has led to trials such PROFILE (peripheral CD8 transcriptome, UC) and IBD-Response (microbiome, UC and Crohns) that are seeking to prospectively classify patients into those that would benefit from higher grade immunosuppressive therapy earlier in their course of treatment, versus those that are likely to respond to more moderate regimens.

Host-microbiome interactions driving ulcerative colitis

The colon cannot be understood as an entity distinct from the microbiome which exists within it[60]. Separate streams of evidence indicate that diet, socioeconomic circumstances and geography (all of which affect the risk of developing inflammatory bowel disease) are associated with certain species of archaea, bacteria and fungi, and they, with a certain metabolome. An individual's microbiome/metabolome is derived partially from the mother at birth, and is modified by exposure to their environment until the age of around four, when it stabilizes[61]. Multiple host and environmental factors are likely responsible for certain organism-favouring niches, given that an individual's microbiome appears to be 'tailored' to them[62,63]. Although studies are plagued by a lack of consistency, both in terms of experimental protocol as well as the sheer diversity of organisms that colonise the human gut across different biospheres, the microbiome appears to have marked effect on host health[60,61]. Particularly in the context of the colon, inflammatory bowel disease and maintenance of healthy bowel appears to be affected by the metabolome/microbiome[64–69], although the effect sizes are variable and sometimes paradoxical. The mechanism is also unclear, but existing research indicates could be as pleiotropic as affecting epithelial cell development[70] and modulating immune cell function[69]. Although there is considerable heterogeneity in the efficacy of the approach, fecal microbial transplant (FMT) appears to be an effective treatment option for UC[64,65,67].

Despite our developing understanding of immune cell perturbations in UC, as well as the development and characterization of various mouse models of colitis, the initiating event in ulcerative colitis, as well as the cognate antigen, remains unknown. Viral, fungal, bacterial and self-antigens[71] have all been proposed with varying degrees of evidence as support as candidates for driving the colitis. There are well documented changes that occur both in the mutational landscape of the self[71], as well as in the makeup and metabolome of the microbiome in inflammation, but whether these are a cause of inflammation, or simply changes in response to the altered milieu, is not clear.

Immune checkpoints and immunotherapy for cancer

Immune checkpoints are a critical facet of the immune system that are thought to have developed in order to allow co-existence with antigens we are chronically exposed to and cannot eliminate. They are a broad class of membrane receptors that are predominantly expressed on T cells (with associated ligands expressed on all cell types), and act as a negative feedback mechanism to induce 'exhaustion' within the immune system whenever it is activated by presence of antigen. These molecules serve to blunt the T cell response to a persistent antigen, and are either present constitutively or are upregulated concomitant with an inflammatory response. T cells bearing these receptors act generally in a more 'exhausted' or anergic manner to other T cells bearing the same TCR[72–74]. This family of extracellular receptors includes, but is not limited to, PD-1, CTLA-4, LAG3 and TIM3. As such, they play a critical role in post-thymic tolerance.

In humans, associations have been described between single nucleotide polymorphisms (SNPs) in checkpoint pathway proteins and autoimmune diseases including, but not limited to systemic lupus erythematosus, psoriasis, rheumatoid arthritis, type 1 diabetes, allergy and progressive multiple sclerosis, these appear to be dependent on the ethnicity of the population analysed[75–78]. Multiple knockout mouse models also support the notion that checkpoint pathways play a role in physiological tolerance, not only in autoimmune disease[79], but also in the context of feto-materal tolerance in the placenta[80].

Although investigated in the context of chronic infection, the biggest therapeutic impact of these signalling pathways was discovered in cancer immunotherapy with drugs that blocked the ligand or receptor of checkpoint pathways, loosely categorized as "checkpoint inhibitors". Utilizing immune activation within a cancer microenvironment had been a potential strategy for the treatment of cancer since the discovery of 'Coley's toxin'[81], but no consistent strategy had been effective.

Introduction

In a series of seminal trials, blockade of PD-1 and CTLA-4 signalling by checkpoint inhibitors that durably bound to PD1 (Nivolumab, Pembrolizumab) and CTLA4 (Ipilimumab) transformed progression-free survival in malignant melanoma[82,83]. Since then, their use has been expanded to non-small cell lung cancer[84], lymphoma [85], renal [86] and prostate cancer (NEPTUNES trial), and trials are ongoing in a variety of other malignancies[87], with some reviews projecting a role in treatment in upto 50% of all cancer types in time[88,89].

PD-1 and CTLA-4 pathway blockade, either singly (PD-1 blockade, monotherapy) or concomitantly (PD-1 and CTLA-4 blockade, combination therapy), has been the mainstay of therapeutic approaches till now. Although clinical trials of all possible permutations have not been possible, generally combination therapy results in higher proportion of response and duration of remission, as PD-1 and CTLA-4 occupy non-redundant roles in the induction of anergy[90,91]. Following their success, blockade of other pathways is under active development, with LAG-3 inhibitors entering clinical use in early 2022[92].

There is a marked heterogeneity in response across patients, with approximately 15% of patients responding to the therapy long term[88,93,94]. Multiple factors have been identified as increasing the likelihood of response, such as high mutational burden due to mismatch repair defects[95] as well as the relative proportion of stem-cell like T cell subsets[96], amongst others, and this is an area of active research.

The exact mechanism through which these drugs act within the tumour remains controversial[97–100], but current opinion credits expansions of pre-existing, ‘proto-exhausted’ CD8 populations within the T cell compartment in tumour tissue[101]. However, this is unlikely to represent a complete picture, as these proteins are also expressed in B cells and antigen presenting cells[102].

The role of checkpoint pathways has also been explored in the context of chronic viral infection[103]. There are no large trials, and they are currently only to be considered in the therapeutic context of cancer and co-infection[104]

However, perhaps somewhat predictably, checkpoint inhibitor therapy causes undesirable auto-inflammatory disease in numerous tissue sites thought to be unrelated to viral exposure or sites of metastatic cancer deposits. These are most commonly barrier areas, such as the skin, colon and lung[104,105]. The type of autoimmune reaction and the tissue affected is also surprisingly protean, as although barrier tissues are most commonly affected, organs such as the liver, adrenal glands[106] and brain[107] are not spared. The culprit mechanism also appears to be variable, with some patients developing a T-cell mediated destruction[108], whereas others develop an antibody-mediated disease[109,110].

Within the gastro-intestinal system, these reactions range from mild to severe, requiring a range of responses from cessation of immunotherapy through to requiring sustained immunosuppression to prevent florid complications such as colonic perforation, liver failure and necrotizing pancreatitis. The mechanism for the induction and maintenance of these auto-inflammatory reactions remains unclear.

Checkpoint inhibitor induced colitis

Checkpoint inhibitor induced colitis (CC colitis) was recognized early as an immunotherapy-related adverse event (IRAE) and is described in the first large clinical trials[111]. It was categorized by oncologists according to clinical criteria of severity and treated with steroids, sometimes with prolonged courses with considerable associated morbidity.

It became apparent early in clinical trials that anti-CTLA-4 therapy, and in particular combination therapy with anti-PD1 and anti-CTLA4 therapy carried a high risk of inducing colitis[82,111,112].

Prior to 2017, at the inception of this study, relatively little was known about the pathophysiology and cellular response involved in CC colitis.

In a case series of 27 patients with CC colitis[113], the authors described a predominantly distal colitis (i.e. affecting descending colon, sigmoid and rectum) with clinical features of ulceration.

Introduction

Another case series of 9 patients with CC colitis[114](in the context of CTLA-4 blockade) demonstrated left sided colitis, with neutrophils in variable numbers and plasma cell infiltration in several patients. In a further case series of 8 patients (treated with anti-PD1 therapy)[115], the authors defined 2 predominant histological manifestations of colitis in those with diarrhoea. 62% developed a neutrophilic infiltrate, whereas the remainder had a predominantly lymphocytic infiltrate.

Another study[116] had assessed 25 patients with anti-PD1 and 5 patients with anti-CTLA4 induced colitis, and in both situations, described a predominantly T lymphocyte driven infiltrate. They performed the first comparison of the colitis induced by the two distinct immunotherapy regimens and described that CD4 lymphocytes predominated in aCTLA4 driven CC colitis (60% vs 40% CD8s), whereas CD8s were the predominant lymphocyte in anti-PD-1 colitis.

There were no papers describing stool or blood clinical markers that correlated with CC colitis (unlike UC), however it was noted that *CD177* RNA and related *CEACAM1* (two neutrophil activation markers) were markedly elevated in blood early after immunotherapy in those patients who went on to develop IRAE colitis. There was no baseline difference between the two groups with regard to these markers[117]. Those who had prior autoimmune disease were prone to developing a flare[118], but numbers were too small to ascertain if this was particularly true for colitis.

It had been shown that the presence of a microbiome was necessary for a good response to checkpoint inhibitors[119] and particular microbiome signatures could even predict response to immunotherapy[120–122], although the exact organisms responsible varied between studies in different populations[123]. The mechanism for this remained unclear, although it was shown that patients on checkpoint inhibitor therapy tend to develop antibodies to a variety of gut microbiome antigens[124], but there was no correlation between the presence of antibodies and the grade of IRAEs. However, microbiome studies demonstrated that patients with a high carriage of *Bacteroides* species were less likely to develop colitis[125]. Taken together, it suggested that the local microflora was immunogenic in CC therapy and might be a key driver in the development of CC colitis. In line with this, trials of FMT therapy demonstrated complete efficacy in a small case series of two individuals refractory to other treatment[126].

In line with the key role suspected of the local immune response, contemporaneous research with our own analysis went on to show, using sc-RNAseq and FACS analysis, that tissue resident (TRM) cells were significantly upregulated and expanded in CC colitis, along with an activation and enrichment of MAIT cells[127,128]. Clinical studies showed efficacy of anti-TNFa (infliximab) and vedolizumab[129,130] and anti-JAK/STAT therapy[128,131], in keeping with the patterns of response seen with UC. However in both diseases a significant number of patients presented with a non-response to therapy, representing a significant unmet clinical need[132].

Rationale for comparison of UC, CC colitis and health

Defining aberrant inflammation

Induction of inflammation is a critical part of homeostasis. Understanding why inflammation may be excessive (requiring resection of the inflamed segment), lead to aberrant fistulation (such as in the case of Crohns' disease) or excessively persistent (lasting months or years) is key to understanding the pathophysiology of IBD.

Studies in humans find it technically and ethically difficult to serially sample patients through the course of an inflammatory episode. Characterizing the response at the initiation, median phase and resolution of inflammation however is key to understanding how inflammation resolves.

Multiple studies have sampled patients early in the time course of disease, but their focus has been on correlating inflammatory patterns[133], methylation[134] or microbiome[135] with clinical outcomes or treatment efficacy[136]. Although these analyses, alongside efforts such as PROFILE (see above) further the goal of personalised medicine, they provide more limited insight into why inflammation in UC does not spontaneously resolve.

Such studies are clearly more feasible in experimental models of colitis in mice, and DSS-induced colitis has been the seminal model investigated so far. In this context, single-cell transcriptomic analysis[137], metabolome/microbiome[138] and high-resolution transcriptomics driven cellular interaction analysis[139] have all been successfully performed. These elegant analyses have identified novel, potentially critical fibroblast roles in the resolution of inflammation, either through the upregulation of certain genes such as *Serpina3n* and inhibition of tissue elastases[140] or development of certain micro-domains of inflammation-associated fibroblasts that favour repair[139]. Although these studies in mice draw comparisons with the inflammatory process in IBD, time course analysis of the inflammation and restitution process in human tissue samples is important because insights from mouse models do not always translate well to human disease.

The utility of the right control

Analysis of inflammatory bowel disease has focused on characterizing the inflammatory response and understanding the genetic, microbiological and environmental factors associated with a risk of developing aberrant inflammation. The comparator in these analyses has either been healthy individuals and tissue or an alternative form of IBD (such as comparisons between UC and Crohns').

From a clinical standpoint, comparisons with health have been very fruitful, yielding insights into the key pathways involved in inflammation, all of which have been critical to identifying new therapeutic agents which are now standard-of-care for the treatment of IBD. Currently, approaches centre on treating increases in cytokines such as TNF alpha (infliximab/adalimumab/golimumab), increased cell trafficking of effector T cells into the gut (vedolizumab), downstream activation of the JAK/STAT pathway (tofacitinib/filgotinib/upadacitinib) and upregulation of the Th17 CD4 T cell population via the IL12/23 pathway (ustekinumab). Such analysis continues to inform future drug development strategies, e.g. targeting epithelial cell tight-junction dysfunction[141] or macrophage modulation[142].

Key though such analysis has been, our understanding of the contribution of these pathways towards the aberrant nature of inflammation in IBD remains incomplete – multiple approaches that were expected to have efficacy e.g. blockade of interferon gamma[143–145 14656]inhibiting IL17[146], or blocking the of exit of pro-regulatory cells from tissue[56], or even drugs targeting the same pathway, but using a different mechanism e.g. TNF-alpha receptor inhibition[147] were ineffective or worsened inflammation. Moreover, all current treatment immunosuppresses the patient, rather than returning them to a healthy 'mean'. Patients remain at risk of developing flares in the future. No treatment offers a 'cure'.

Genome-wide association study (GWAS) analysis has also identified over 160 risk-increasing loci for inflammatory bowel disease. Although many are shared between Crohns' disease and Ulcerative colitis – e.g. pathways dealing with intracellular pathogen response/Th17, and act in a congruent way for both disease subtypes, some are disease specific (e.g. *NOS2,IFNGR2* in CD; *TNFRSF14, NFKB1* in UC) whilst some exert opposite effects (e.g. *PTPN22* and *NOD2* are protective for UC, but increase the risk of CD)[148,149]. Understanding how these changes at gene level affect the development of disease is confounded by the effects of the environment and microbiome[150]. The goal remains to translate genetic information into a prognostic tool, as monogenic forms of the disease have led us to understand that what we classify as 'UC' or 'Crohns' may actually represent a whole family of related diseases[151]. The ultimate target, to enable personalised medicine beyond a 'one-size-fits-all' strategy (through prospectively informing therapeutic choice, monitoring method, assessing cancer risk etc) remains out of reach for now.

Checkpoint inhibitor-induced autoimmune effects, by comparison, have only recently been characterized to the same depth as historical analysis of IBD. No drugs have been developed specifically for its treatment, as drugs already in use for the treatment of IBD were employed as immunotherapy colitis increased in prevalence[152]. There is, intriguingly, some evidence that inhibiting epigenomic changes at the induction of immunotherapy (histone de-acetylation in particular) may reduce the development of disease[153], but the mechanisms by which it might be doing so remain obscure. GWAS analysis has identified loci near the *IL7R* and *IL22RA* genes as increasing the risk of developing autoimmune side effects, possibly related to the development of certain CD8 T cell and B cell responses on exposure to immunotherapy[154,155]. The effect of the microbiome on checkpoint-inhibitor induced colitis is thought to be significant (see above), but currently remains unexplored, and plagued by issues similar to those analyses that have been used to characterize UC to date.

What utility then, of a comparison between UC, CC and health? As highlighted above, multiple lines of enquiry have shown a similarity between UC and CC colitis. Although one is clearly idiopathic and the other induced, both tended to affect the distal colon (as summarised above), with a lack of inflammation penetrating through the mucosal wall. Bamias et al showed that in the context of CC colitis, IFN gamma, IL17, IL10 and FoxP3 mRNA transcription were all higher compared to healthy controls[114], in the same range as IBD biopsies. Both UC and CC colitis had similar histopathological features, appeared to respond to similar drugs and were at higher risk of occurrence with NSAIDs[129]. Finally, once induced, both ulcerative and CC colitis were self-sustaining and chronic inflammatory processes in humans. There are more similarities between UC and CC than there are between UC and Crohns' disease.

By comparing ulcerative colitis, checkpoint inhibitor induced colitis and health, we sought to understand the pathophysiology of both in more detail, differentiating between changes seen consequent to inflammation, as part of a general response, and those that may be particular to either disease, ideally yielding mechanistic insights that had been difficult to distinguish before. Although we attempted to include self-resolving infectious colitis and endoscopically milder microscopic colitis as comparators, difficulty in collecting these samples, exacerbated by the Covid-19 pandemic, rendered this cohort incomplete and insufficient for analysis.

Although our analysis highlights facets of the inflammatory process that are unique to both UC and CC, and therefore potentially mechanistically relevant, there is currently an unmet need for comparison with acute inflammation that resolves spontaneously (e.g. infectious colitis), and ideally, an analysis of the natural history of inflammation in humans, and how it breaks down. Unbiased transcriptomic or spatial technologies, coupled with novel cell lineage tracing methodologies[156], may offer insight into what deficiencies or aberrant pathways drive persistence of inflammation, which is key to understanding IBD.

Aims of thesis

The results of our study of Checkpoint and Ulcerative colitis have been summarized in this thesis as follows:

(1) **Chapter 1**: An analysis of ~1000 patients given checkpoint inhibitors to identify clinical patterns of CC colitis.

Novel findings:

a. Assessment for risk factors for development of CC colitis
b. Risk stratification of patients to better identify those likely to develop severe disease, and highlight differences from UC
c. An analysis of the efficacy of the current treatments of CC colitis and impact on survival

(2) **Chapter 2**: A novel in-depth unbiased analysis of critical tissue CD8 T cell populations in idiopathic UC

Novel findings:

a. Identification of novel populations of CD8 T cells in health and disease and understanding tissue clonal dynamics of exhaustion and activation
b. An examination of interactions with diverse non-immune populations in tissue, identifying a novel IL26-epithelial interaction that may ameliorate colitis severity

(3) **Chapter 3**: Characterization of all immune and non-immune population behaviour in CC, UC and HC utilizing both unbiased single-cell as well as spatial transcriptomic information

Novel findings:

a. Identification of common and unique patterns of inflammation in both UC and CC
b. Characterization of novel 'micro-domains' of inflammation and repair in UC and CC with potential therapeutic implications utilizing novel unbiased spatial transcriptomics
c. Identification of populations in CC that may assist with predicting those who will develop the disease, suggesting a possible mechanism of development of the disease.

These results will be discussed in detail in each chapter, preceded by a discussion of the methods employed.

Methods

Chapter 1

Ethical approval

Given this was a retrospective anonymised analysis of data, no formal consent was sought from patients, however, the overall data collection and analysis was performed as part of preliminary work for the PRedicting Immunotherapy Side Effects (PRISE) Study (London-Surrey Research Ethics Committee: REC18/LO/0412) and the 'A Mechanistic Investigation into Drug and Chemical Induced Hypersensitivity Reactions (HYST)' study (REC12/NW/0525), both of which received NHS-Research Ethics Committee approval. Data was handled and stored in accordance with Caldicott principles of confidentiality and Good Clinical Practice.

Data collection

All adult patients who received checkpoint inhibitor therapy between January 2012 and October 2018 at the Royal Liverpool University Hospitals NHS Trust (headed by A.O.B) and Oxford University Hospitals NHS Foundation Trust (headed by O.B.) for cancer were identified using oncology drug prescribing electronic patient records (Aria® system at Oxford, Meditech system at Liverpool).

This included patients undergoing Ipilimumab (anti-CTLA4), Nivolumab and Pembrolizumab (anti-PD1) or Ipilimumab and Nivolumab (Combination) therapy for a variety of metastatic cancers (predominantly metastatic melanoma, non-small cell lung cancer, prostate and renal cancer).

Using electronic patient records (Cerner Millennium®), two medical professionals (V.T.F.C and T.G.) scanned through letters and hospital records to extract all relevant patient demographic information from the Oxford dataset, including but not limited to patient characteristics (e.g. date of birth, gender, smoking status, type of cancer), disease characteristics (e.g. development of any immune related adverse events, timing of immunotherapy, severity), treatment (e.g. requirement for infliximab, timing and number of doses) and outcomes (e.g. mortality, resolution of colitis), populating these into an excel spreadsheet, ensuring data protection guidelines were followed. Once collected, each patient could only be identified through a unique 7-9 digit identifier, preventing duplication in analysis, whilst minimizing bias.

Liverpool data (stored on a Meditech database) were collected locally and forwarded on in an anonymized fashion to T.G. If further data was required, a request was submitted to A.O.B, data extracted from the database and the data returned to T.G. using the same identifiers.

Patient characteristics including survival

Data such as date of birth, gender and date of death are recorded nationally and updated automatically to the electronic patient record. The type of immunotherapy and cancer was clearly documented locally on local systems given implications on clinical care. Smoking data was not collected in a proforma and relied on the assessing clinician or health provider to document during the multiple consultations a patient had during the course of their treatment.

Determining onset and severity of checkpoint inhibitor-induced colitis

Patients undergoing immune checkpoint therapy for cancer are extensively briefed about the possibility of developing immune reactions and the need to report this rapidly to oncology departments for early treatment. In addition, patients having these drugs are regularly reviewed in oncology clinic after induction with the drugs: anti-PD-1 inhibitor therapy recipients are reviewed at 8 weeks after starting then 3-monthly; those having dual therapy are reviewed at 4 weeks (i.e. after cycle 1) then at 6-weekly intervals. Severe immunotherapy reactions mandate a cessation of immunotherapy/change of regimen. Taken together, the aim is to assess how patients are coping with therapy and effectiveness, but it ensures that the development of immune related adverse events was well and clearly documented in electronic records. Given the nature of the treatment, patients did not discontinue/were not lost to follow-up.

Methods

The diagnosis of checkpoint inhibitor-induced colitis (CC colitis) was based on the decision of the treating oncologist or gastroenterologist. The definition was typically clinical, with immunotherapy induced diarrhoea being defined as per CTCAE version 5.0[11] (Table 1):

Table 1: CTCAE grading for diarrhoea

Term	**Definition**
Grade 1 Diarrhoea	Increase of < 4 stools per day over baseline
Grade 2 Diarrhoea	Increase of 4-6 stools per day
Grade 3 Diarrhoea	Increase of > 6 stools per day or incontinence or hospitalization
Grade 4 Diarrhoea	Increase of >10 bloody stools per day or life-threatening consequences
Grade 5 Diarrhoea	Death

Alternative diagnoses (like infection or use of non-steroidal anti-inflammatory drugs (NSAIDs)) were excluded by the treating clinician. The timing of last dose of immunotherapy was recorded as a date on the record, with the time of onset of diarrhoea as reported by the patient and recorded by the assessing clinical staff.

Biochemical markers of severity

Available data for available patients was extracted from the electronic patient record database by T.G., with the relevant time period identified by the timing of colitis. The parameters to be measured (Haemoglobin, C-reactive protein, Albumin) were pre-determined as these are known to change in idiopathic ulcerative colitis and are used as markers of severity in the Truelove-Witts (Hb, CRP)[157] and Ho indexes (Albumin)[158].

Determining endoscopic severity of disease

As oncology guidelines currently do not require an endoscopy or blood tests prior to initiation of treatment[159], not all patients received an assessment prior to initiation of treatment.

Therefore we conducted a blinded analysis of the subset of patients for whom endoscopic data were available (n = 40, 30%), by two experienced endoscopists (O.B. and V.T.F.C) both credited by the Joint Advisory Group on Gastrointestinal Endoscopy (JAG). These assessors independently graded the endoscopy images using the UCEIS and Mayo severity scores, with good agreement (UCEIS - Kappa = 0.51, SE = 0.09; Mayo - Kappa = 0.54, SE = 0.09; Landis and Koch criteria). The results were collected and analysed by a blinded researcher (T.G.), who highlighted any disputes back to the assessors. Then V.T.F.C and O.B. re-evaluated the disputed records to reach a consensus which was communicated back to T.G., then used for analysis. Endoscopists were blinded as to patient outcomes and treatment.

Determining histopathology in disease

Given not all patients underwent endoscopy, we analysed histological slides from those that were available (n=45). Two expert GI pathologists (E.F. and E.C.) were given patient details and separately scored the slides on the presence of ulceration, acute inflammatory cells infiltrate and chronic inflammatory infiltrate, which was then used to calculate the Nancy index of severity. The histopathologists (blinded to clinical outcome) then assigned an overall histological pattern to each patient: focal active colitis, lymphocytic colitis, collagenous colitis, ulcerative colitis and drug-induced/infectious. All the histological specimens were acquired at the index scope of diagnosing IRAE colitis and prior to infliximab use. The collected data was forwarded to T.G. for analysis..

Missing data

Given the retrospective nature of the analysis, there were multiple metrics for which we did not have data. Where data is missing, this is indicated and quantified in the relevant analysis.

Data analysis

Data analysis for all conclusions presented in the thesis as well as all figure generation was performed by T.G. using Graphpad PRISM™ (Ver 8.1, **Materials**), with the relevant tests and metrics annotated as per the figure legends. Continuous data were presented with mean (with standard error of mean) or median with interquartile range, as indicated in the figure. A p value < 0.05 was considered significant. Non-continuous data were presented as patient numbers and percentages. Differences between groups were determined using T-tests, ANOVA or linear regression analyses, as specified in individual figures.

Chapter 2

Ethical approval

All human samples (colonic biopsies and blood) were collected from patients randomly attending endoscopy at the Oxford University Hospitals NHS foundation trust as part of their routine clinical care. They were collected by GI Biobank staff or T.G. after acquiring informed consent from patients under the aegis of multiple NHS Research Ethics Committee (NHS REC)- approved studies - 18/WM/0237, GI 16/YH/0247 and IBD 09/H1204/30, all in accordance with the principles of Helsinki. Patient demographic and treatment data is summarized and is available (**Appendix D).**

Sample collection and storage

Sample handling All samples whilst being processed, unless specified otherwise, were kept on ice in specified medium and handled in sterilized tissue culture hoods. All washing steps were carried out at 4°C at 400g for 5 minutes in a balanced centrifuge. All incubators were maintained at 37°C with 5% CO_2 and 21% oxygen.

Colonic samples: 4 biopsy pairs were collected in biopsy medium (DMEM, 10% FCS, 100U-100ug/ml Pen-strep, 1X HEPES, 1X Pyruvate, **Materials**) on ice. Whilst collection was occurring, samples were stored on ice for a maximum of 2 hours, following which they were spun down at washed in ice cold PBS, and then suspended in 0.7mls of CS10 (**Materials**), cooling down at 1 degree per minute to -80degrees C (using a Mr Frosty system) in barcoded vials. The following day, they were transferred to liquid nitrogen for long term storage. Other members in the group had performed prior analysis to check that freezing samples in this way did not affect their single-cell transcriptional profiles as compared to fresh samples (unpublished data).

Methods

Blood samples: 10-20mls of blood were collected in standard EDTA tubes to prevent coagulation. Following storage at room temperature for a maximum of 2 hours, the sample underwent density gradient centrifugation separation to enrich for PBMCs. Briefly, the sample was diluted in a 1:1 ratio by volume with room temperature sterile PBS, mixed, following which it was layered carefully on top of 12-15mls of Lymphoprep in a 50ml Falcon tube, avoiding mixing of the blood and lymphoprep. The falcon was spun at 800g x 20 minutes at room temperature, with decelerations off. Following this the PBMCs formed a clear layer, bordered by serum on top and lymphoprep underneath. This was aspirated using a pasteur, dispensing into blood medium (RPMI 1640, 10% FCS, 1X Pen-strep, 1X HEPES). This was washed twice with 20mls of RPMI, followed by processing.

Formalin-fixed paraffin embedded (FFPE) samples: Colonic biopsies retrieved as above were fixed in 10% neutral buffered formalin for 48 hours before being placed in 70% ethanol, this was followed by a standard dehydration and wax embedding protocol, performed by automated processors. Blocks were stored at room temperature before 4uM sections were cut upto 1 month prior to staining.

Determining disease state The diagnosis of health or ulcerative colitis, as well as severity by UCEIS/Nancy index was confirmed by endoscopic and histopathological assessment of the clinical samples collected at the same time as the research samples, as well as with recourse to historical patient care records, accessed as per Caldicott principles of data management.

Biopsy dissociation

Biopsies transferred from liquid nitrogen to the laboratory on dry ice were defrosted into biopsy medium (DMEM, 10% FCS, 1X Pen-strep, 1X HEPES, 1X Pyruvate, **Materials**) at 37°C. Following washing twice in the same medium, they were chopped into fine pieces using a scalpel and suspended in 1ml of DMEM supplemented with Type 2 Collagenase and DNAse 1 (**Materials**) at a final concentration of 1mg/ml and 50ug/ml respectively in a flat 24 well plate. This digest solution was incubated in an incubator at 37 for 60 minutes, with the chopped pieces undergoing mechanical dissociation using a Stemcell 16 gauge needle at 20 minutes, 40 minutes and 1 hour intervals. Following this, the solution was passed through a 70uM filter to remove debris, washing the filter with biopsy medium. The final solution was spun down and resuspended in 200ul of biopsy medium, and the cells count and viability measured using a Countess II system.

FACS staining

All FACS staining was done in a round bottom 96 well plate, with washes with relevant medium at 800g x 1minute at 4dC unless stated otherwise. Single colour controls were created on beads (IL26 experiments) or cells (single-cell sorting experiments).

IL26 quantification PBMCs from a healthy donor and HDLM-2 cell line cells at passage 3 were isolated as described and 1 million cells at >95% viability of each were placed in different wells of a 96 well plate. Additional PBMC and HDLM-2 cells from the same source were used as FMO controls. The cells were washed x 2 in 200uL of PBS, followed by staining in 1:200 dilution of Zombie Violet for 12 minutes at room temperature in the dark. Cells were washed with 200uL cell staining buffer x 2, then resuspended in 50uL of staining buffer containing 1uL Trustain FcX block, 0.5uL CD45-FITC, 2uL CD3-APC (**Materials**) for 30 minutes at 4°C in the dark. Following this, cells were washed x 3 with 200uL staining buffer, then incubated with 100uL of fix-perm solution for 15minutes at 4dC in the dark in preparation for intracellular staining. Cells were washed x 3 with 200uL perm-wash solution, then resuspended in 50uL of perm-wash solution with a titration series of IL26-PE (2.5uL , 5uL, 10uL, 15uL, 17.5uL/1 million cells; recommended concentration 10uL/1 million cells) for 30 minutes at 4°C in the dark. Cells were then washed x 3 with 200uL of perm-wash solution, then resuspended in 200uL of staining buffer, before being run on an BD LSR 2 flow cytometer.

IL26 receptor quantification The known receptor for IL26 is a heterodimer of IL20RA and IL20RB[160]. It is known to be expressed on cell line SW-480[160,161], which was used as a positive control. Biopsies from healthy and ulcerative colitis patients were dissociated into a single cell suspension as described above. 1 million cells from each sample (>75% viability) were placed in a well of a 96 round bottom plate as described above, and washed x 3 in 200uL of staining buffer. Matched samples were taken for FMO controls (FMO IL10RB and FMO IL20RA). The cells were then suspended in 50uL staining buffer containing 1ul Trustain FcX Block + 0.5uL CD45-FITC + 1ul EPCAM-BV785 + 5ul AF647-IL10RB + 15uL PE-IL20RA (**Materials**) for 30 minutes at 4°C in the dark. Cells were then washed x 3 with 200uL of staining buffer and 1:1000 DAPI was added just prior to acquisition on an LSR 2 flow cytometer.

CITE-seq antibody testing We wished to confirm that the digestion protocol would not cleave epitopes for the hashing or key CITE-seq antibodies (e.g. CD103, PD1, CD45RO), as the former would render the single-cell runs useless (cells recovered without an antibody tag could not be attributed to an individual sample in a run, making comparisons impossible). We also wished to determine a safe level of titration for the hashing antibody (anti-B2 microglobulin). In order to do so, we isolated single cells from biopsies from a healthy donor, using the digestion protocol described above. The sample was split into multiple 2 million cells/ well in a 96 well plate and washed twice with 200ul staining buffer. Samples were then incubated in 100ul of an antibody mix containing Fc block, 1.66uL CD3-BV711, 1.66uL CD8-APC-R700 and 0.5ug of each tested CITE-seq antibody clone (conjugated to PE) and either 0.5ug or 0.25ug of anti-B2 microglobulin antibody clone (conjugated to PE) at 4°C for 30 minutes in the dark. We ensured that the clone used for FACS staining was the same as used for the CITE-seq staining. Cells were then washed with staining buffer, stained with 1:1000 DAPI and acquired on an LSR 2 cytometer. We established that 99.9% of all CD3+CD8+ cells bound the hashing antibody clone at both 0.25ug and 0.5ug/well concentrations, and substantial proportions of positive cells were detected by each CITE seq antibody, as a result of which we proceeded with this protocol.

Single-cell RNA sequencing sorting Single-cell suspension of biopsy cells from healthy controls and patients with ulcerative colitis were prepared from biopsy samples as described above. Cells were counted and resuspended in a 96 well plate, aiming for 2million cells per sample per well. Each well was washed twice in 200uL of staining buffer before resuspended in 50uL of staining buffer and addition of 5uL of Trustain FcX (**Materials**), followed by incubation for 10minutes at 4°C. During this interval, a 'master mix' of antibodies was prepared, containing 1.66uL CD3-BV711, 1.66uL CD8-APC-

R700 and 1uL (0.5ug) of *each* CITE-seq antibody (14 in total, full list in **Materials**) in 50ul of staining buffer per well. This master mix was then split into five fractions, and the appropriate hashing antibody added to each fraction (titrated to 0.5g i.e. 1uL per 2 million cells). Each fraction (containing FACS antibodies for CD3, CD8, all 14 CITE-seq antibodies and one hashing CITE-seq antibody in 50uL) was then added to each 50uL sample treated with FcX blocker and incubated with the appropriate sample well for 30minutes at 4°C in the dark. Cells were then washed x 3 with 200ul of staining buffer and 1:1000 DAPI was added just prior to sorting using a FACS Aria IIIu sorter. CD3+CD8+ live cells were sorted directly into eppendorfs containing 50uL of sorting buffer (PBS and 0.04% BSA) and kept on ice, all samples being run to dryness. Once all five samples had been sorted, cells were spun down, and each sorted CD8 T cell sample resuspended in sorting buffer in an appropriate volume to reach a concentration of 1x10^6 cells per uL. 8uL of each sample was taken and pooled to create a final volume of 40uL (40,000 cells), 20uL (20,000 cells) of which was added to the master mix and loaded onto the 5’ Chromium 10X platform.

10X single-cell RNA sequencing

We enriched CD3+ CD8+ T cells from healthy controls and inflamed ulcerative colitis biopsies through dissociation into a single cell suspension and FACS sorting as described above. Cells from five different donors across disease and inflammation were pooled, and then 11,000-20,000 were loaded per run onto the Chromium 10X platform. Library generation for CITE–seq and hashed samples was performed using 10x Chromium Single Cell 5’ V(D)J Reagent Kits with feature barcoding technology (user guide, no. CG000186). Resulting Gene expression (GEX), T-cell receptor (TCR) and CITE-seq protein and hashing antibody (ADT-HTO) libraries were sequenced on an Illumina Novaseq 6000 S4 platform to achieve an average of 50,000 reads per cell for the GEX libraries, 5,000 reads per cell for the ADT-HTO libraries, and 2,000 reads per cell for the TCR libraries. Total sequencing was based on estimated cell recoveries, with additional sequencing performed if the library had not been sequenced to saturation on bioinformatic.

The original cohort without CITE-seq (run by D.C.) comprised 3 ulcerative colitis and 3 healthy control samples, processed and stained as described above, but loaded without pooling on the 5' Chromium platform and sequenced on an Illumina HiSeq4000 platform to achieve an average of 50,000 reads per cell.

Quantitative real-time PCR

We performed quantitative real-time (qRT-PCR) on human colonic biopsy (4 pairs) and mouse whole colon tissue samples. UC non-inflamed samples were drawn from the same patients, from paired proximal biopsy areas that were uninflamed by endoscopic and histopathological analysis.

For both human and mouse tissues, samples were physically homogenized (with 100 mg of 1.4-mm ceramic beads, run at 4,000 r.p.m.), and total RNA was isolated using an RNAeasy Miniprep kit. cDNA was synthesized using the high-capacity RNA-to-cDNA kit (**Materials**), standardized to 5ng/ul and qRT-PCR performed on 1ul (5ng) of cDNA using TaqMan gene expression assays on the QuantStudio 7-Flex system. Reference genes used in data analysis are highlighted in the appropriate figure legend.

Bulk RNA sequencing

RNA from hTg-IL26 and WT mouse tissue was extracted as described for qRT-PCR. We assessed the RIN (RNA integrity number) quality of samples using an Agilent Tapestation kit as per the protocol, with high-quality samples with RIN scores >8.0 being converted to strand-specific cDNA libraries using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (no. 7420), with an insert size of 250–300 base pairs. Library prep of identified high quality samples was done by a commercial company (Novogene). Samples were pooled and sequenced to a depth of 20 million reads per sample on a Novaseq 6000 S4 platform.

Mouse experiments

Mouse experiments were performed by our collaborators in Juntendo university in Japan, and are reported fully in our manuscript[2].

Methods

Background Briefly, mice do not express IL26, but their cells express its cognate heterodimer receptor IL10RB/IL20RA, and appear to retain the ability to signal through this[160,162]. The Aune lab developed a transgenic mouse on a C57BL/6J (Black 6) background that expressed human IL26 (utilizing a bacterial artificial chromosome vector carrying IL26 and interferon gamma) and confirmed that under steady state conditions, this hIL-26Tg mouse expressed IL26 (and minimal interferon gamma) in the small intestine and colon[163]. We elected to use this model to investigate the effect of IL26 on colitis.

Ethics Mice were housed under standard conditions in the animal facility at the School of Medicine, Juntendo University, Tokyo, Japan. Animal experiments were conducted following protocols approved by the Animal Care and Use Committees at Juntendo University.

Experimental setup hIL-26Tg mice and B6 Wild-type littermates were housed together in micro-isolator cages under germ-free conditions with free access to germ-free food and water. They were exposed to a 12h light/dark cycle at 24°C +/- 2°C. Both male and female B6 and hIL-26Tg mice at 10–14 weeks of age were selected for the DSS-induced colitis model. For the hIL-26Tg mice, we devised an additional arm to the study where the mice were injected intraperitoneally with either control isotype or IL-26 neutralizing antibody. The neutralizing properties of this anti-IL26 antibody had been confirmed by our collaborators in Juntendo university and was established in literature. The dose and schedule was determined in discussion with these researchers who had extensive experience of this mouse model and utilizing the antibody for this purpose[164,165]. The neutralizing antibody (or its isotype control) was injected on days 0 and 3 of colitis induction.

DSS colitis DSS colitis[166] was induced through introduction of 2.5% DSS into the drinking water of both WT (n = 4 control mice, n = 6 DSS-treated mice) and hIL26-Tg mice (n = 4 control mice, n = 6 DSS-treated mice + isotype control injection, n = 5 DSS treatment + anti-IL-26 antibody injection mice). Mice were sacrificed on Day 6 and colons extracted for analysis.

Analysis of mouse colonic tissue H&E analysis of mouse tissue was organised by D.C. and performed by an accredited mouse histopathologist. Tissue preserved in RNA later was transferred to our lab in

Oxford, where M.J and T.G. extracted RNA, performed the initial qRT-PCR, and then sent tissue for library preparation and bulk RNA analysis after determining RNA quality. Bioinformatic analysis was carried out by A.A.A.

In-situ hybridization (ISH) for IL26

RNA-scope assays for IL26 We utilized the RNA-scope 2.5 HD Brown assay (**Materials**) developed by ACD Bio-techne to look for the presence of IL26 in tissue sections of ulcerative colitis with raised UCEIS scores (UCEIS 3-7). As control, we looked for *PPIB*, expressed by stromal fibroblasts in colon in the same sections of UC[29].

Methodology and optimization Briefly, the test slides were heated for 2 hours at 60°C, following which they were deparaffinized in Xylene and 100% ethanol, then air dried. Slides were incubated with proprietary hydrogen peroxide for 10 minutes, then washed in distilled water it. They were then subjected to incubation with proprietary antigen retrieval buffer for 30 minutes at 100°C. Following a distilled water wash and air dry, slides were incubated with a proprietary protease for an interval of 10-30 minutes (higher incubation times destroy cellular architecture but improve signal from RNA). T.G. determined that the optimum protease digestion time was 27 minutes for these samples. Following a wash, slides were incubated with proprietary anti-RNA probes (*IL26* for test, *PPIB* for controls) at 40°C for 2 hours. They were then washed in proprietary buffer, followed by 6 signal amplification steps using proprietary amplification reagents for between 15-30 minutes at 25-40°C (as per protocol). Finally, the slide was incubated with freshly reconstituted DAB (3,3'-Diaminobenzidine) for 10 minutes to develop signal (brown), then counterstained with Haematoxylin for 45 seconds. Slides were dehydrated through an ethanol and xylene series before being mounted with xylene compatible micromount mounting medium and visualized using a microscope (**Materials**).

Immunohistochemistry (IHC) for IL26

Antigen retrieval and staining 4uM FFPE sections were placed onto slides prior to staining from both healthy (control) and inflamed UC tissue (test).

Methods

Sections were initially deparaffinized - 100% Histoclear series (2 x 5 minutes each) then rehydrated – ethanol series from 100%-70% (4 minutes each), before being suspended in distilled water (4 minutes). Antigen retrieval was then performed with either pH 6 (10mM Citrate buffer) or pH 9 (10mM Tris/1.3 mM EDTA buffer) for 30 minutes at 96°C. Following cooling and washing in PBS at room temperature, slides were incubated with 30% hydrogen peroxide in PBS for 30 minutes at room temperature in the dark to block non-specific peroxidase activity. Slides were then blocked with 2.5% goat serum for one hour at room temperature. They were then incubated with primary antibody (as indicated below) diluted in 1% BSA in PBS for 60 minutes at room temperature, followed by washes in PBS-0.1% tween (PBS-T). Slides were then incubated with HRP-conjugated goat anti-mouse secondary antibody for 30 minutes, followed by another 3 washes with PBS-T. Slides were then incubated with freshly constituted DAB for 5 minutes, followed by a wash in distilled water. Sections were then counterstained with haematoxylin for 45 seconds, followed by a wash in distilled water. Slides were then dehydrated through an alcohol followed by histoclear, then covered with Micromount medium and coverslip applied, prior to imaging.

Antibodies tested with IL26: pH 6 & pH 9 retrieval, AK155, mouse anti-human, concentration of 1:50-1:500; pH 6 & pH 9, Juntendo anti-IL26 9/4, mouse anti-human, concentration of 1:2500-1:120,00; pH 6 & pH 9 Juntendo anti-IL26 10/A, concentration of 1:250-1:75,000 (**Materials**).

Cell culture experiments with IL26

Deriving Mo-DC from human PBMCs Human blood cones underwent lymphoprep separation (with additional dilution) as described for blood lymphoprep separation above. Following this, CD14 cells (monocytes) were separated out using CD14+ MACS separation.

Briefly, PBMCs derived fresh from blood cones were counted and viability checked. Depending on cell counts, they were resuspended in ice cold MACS buffer (0.5% BSA, 2 mM EDTA in PBS) to a concentration of 10^5 cells/uL. Working quickly on ice, cells were incubated with CD14+ microbeads (20uL per 10^7cells), mixed and left on a rotator in the cold room for 15 minutes. 1ml of MACS buffer was added per 10^7cells and cells centrifuged (300gx10min) to remove excess. The supernatant was resuspended in 500uL of ice cold MACS buffer and passed through a pre-wetted LS column within the magnetic MACS separator followed by a flush and washes (as per protocol), with the flow through (containing CD14- cells) being discarded/stored as necessary. The LS column was then removed from the MACS separator and flushed through with MACS buffer to collect the CD14+ fraction.

The monocyte CD14+ fraction thus isolated was then resuspended in RPMI medium (described above) at 1 million cells/ml with supplemented IL-4 and GMCSF (40ng/ml for both) for a period of 5 days, with refreshment of the cytokines on day 3. On day 5, dendritic cells were harvested[167].

Co-culture with IL26 Dendritic cells as derived above were plated at 2.5 million cells/well in a 48 well plate. They were then cultured with recombinant IL-26 for either 4hrs or 6hrs at a concentration of 100ng/ml or 1000ng/ml. We observed no clear differences in a pilot between these conditions, so a timepoint of 4 hours and 100ng/ml stimulation was chosen for the experiment. Following incubation, RNA was extracted using column-based separation and enrichment (Qiagen RNA mini kit) as per the published protocol.

Organoid co-cultures with IL26

Organoid establishment and propagation Human colonic organoids were generated using established published protocols[15].

Briefly, frozen biopsies were defrosted into warmed biopsy medium (as constituted above). These were chopped up into fine fragments to improve surface area for epithelial recovery, and placed in warmed chelation medium (96% HBSS, 1% Pen/Strep, 1% HEPES (1M stock), 1% 5mM EDTA (0.5M stock), 2mM DTT, 1%FCS)[19]. Epithelial crypts were chelated off from the lamina propria at 37dC with gentle agitation for 1 hour using a shaker. The isolated crypts were spun down (400g x 5min @ 4◦C) and resuspended in cold DMEM F12/BSA medium (1% BSA in DMEM F12) on ice, aiming for 1 million cells/ml. Thawed Matrigel at 4◦C was added in a 1:1 ratio to the suspension, mixed and 50ul rapidly dispensed into each well of pre-warmed flat bottom 24 well plate at 37◦C. The pellet was allowed to set for 10 minutes, following which 500uL of conditioning medium (CM) supplemented with 1:1000 Y-inhibitor (for the first passage) was added to each well. Conditioning medium was freshly made up as per the original published protocol for propagation of colonic organoids. Medium was changed every other day until organoids required passaging, when the Matrigel was dissolved using cold dissociation medium, spun down 400g x 5min at 4°C, refreshed with fresh Matrigel/DMEM F12-BSA of double the volume and re-plated as before.

Co-culture experiments All samples were passaged thrice, and experiments begun once the spherules were actively increasing in size on 5 days after passaging. The 4 arms were managed as follows: Medium and "Acute stimulation" arms – CM changed every other day. "Chronic stimulation" and "Chronic stimulation plus IL26" arms – TNFa (1ng/ml) and Interferon gamma (1ng/ml) was added to both types of medium; for the IL26 arm, 100ng/ml IL26 was added on the indicated day and refreshed with every medium change every other day. This concentration of reagents was chosen with recourse to literature[160,168,169] and our measurement of the concentration in tissue[2]. This was continued for 5 days (i.e. 2 medium changes), after which all wells were changed to differentiation medium (DM), which was made up as per published protocols[15]. All wells were incubated with DM for 3 days to allow development of the non-stem cell epithelial compartment. The TNFa, IFNg and IL26 refreshment was carried on as during the CM incubation. The "Acute stimulation" arm was incubated with TNFa and IFNg for 24 hours prior to harvesting and RNA extraction using the Qiagen RNEasy Mini kit. The organoids were imaged at 10X prior to harvesting.

Bioinformatics analysis

Bioinformatic analysis of single-cell RNA seq and bulk RNA sequencing data was carried out by A.A.A. with discussion with D.C and T.G.. All methods used were devised by A.A.A. and are described in the published manuscript *Single-cell atlas of colonic CD8 T cells in ulcerative colitis*[2].

Methods

Single-cell RNA datasets from the original (Accession: GSE134649) and validation (Accession: GSE148837) datasets were deposited online. The mouse bulk RNA-seq data was also deposited online (Accession: GSE148505) and all are freely available.

Chapter 3

Ethical approval

All human samples (colonic biopsies and blood) were collected at Oxford University Hospitals after obtaining full informed consent under the umbrella of multiple NHS-REC approved studies. Ad-hoc samples were collected at random from patients attending for routine clinical care under the following studies - GI Biobank: 16/YH/0247, IBD Biobank: 09/H1204/30, TIP: 18/WM/0237, as well as systematically under the PRISE Study: 18/LO/0412. Patient demographic and treatment data are summarized in **Appendix D**.

PRISE study

T.G. helped set up and manage the PRISE study (Predicting Immunotherapy Side Effects, conceived by O.B.) at Oxford University Hospitals. The study was conceptualised after a review of the literature by O.B. suggesting a lack of characterization of CC colitis at its inception, accompanied by an increasing number of cases seen in clinical practice. The trial protocol, patient information sheet and consent forms were developed in partnership with the Oxford University Hospitals R&D department, and with the support of the oncology service. T.G. assisted with securing ethical approval with the NHS REC committee, setting up a Redcap database to securely collect and store patient information, consenting patients and document management, as well as managing minor and substantial amendments. PRISE aimed to collect blood and colonic tissue samples from patients receiving immunotherapy at 6-8 weeks after initiation of treatment, with optional endoscopy *prior* to receiving immunotherapy, as well as samples at the time of any suspected colitis. Patients commencing mono (anti-PD1) and dual (anti-PD1 and anti-CTLA4) immunotherapy were identified at random in clinic, agreeing to screening endoscopy prior to initiation of the study.

The study was received well in public engagement events, and patients were happy to contribute samples and time, both at the time of their attendance at oncology outpatients, and also when approached for ad-hoc sampling. Although not formally calculated, the incidence of patients declining ad-hoc sampling for research was low. The PRISE study was mentioned in publications resulting from the project, as well as in scientific talks and posters presented locally and internationally, with acknowledgement of oncology support. Regular patient engagement events by the research nurse staff, Oxford University sponsorship, as well as involvement of the oncologists early in the study design and recruitment process, were key factors in fostering patient trust and recruitment to the study.

Categorization of samples

Patients were counselled by oncology prior to immunotherapy initiation of the possibility and variety of immune-related adverse events, and for the need for urgent treatment. As all patients continued to receive immunotherapy locally, it is unlikely that episodes of colitis were not reported.

Because of close links established by O.B. and V.T.F.C, the oncology service were provided with a clear pathway to flag up patients with colitis, particularly those individuals with more severe disease requiring infliximab. In addition to highlighting patients suitable for research, this process improved clinical care by providing a rapid access pathway for advice, investigation and treatment. However, for patients with mild disease, it depended on the clinical judgement of the oncologist or gastroenterologist whether to diagnose and treat the patient as having colitis. Therefore, colitis that resolved with loperamide was potentially less likely to come to the attention of research staff, and consequently, less likely to be sampled for analysis. That being said, our goal was to draw comparisons with UC, which required more clear-cut, CC disease of comparable severity, which we would have captured with this strategy without any readily apparent systematic sampling bias.

Nevertheless, in order to capture milder cases, understand the time course of disease, whilst also eliminating any unforeseen confounding factors, PRISE proposed routine endoscopy for all patients 8 weeks after initiating immunotherapy. Ancillary benefits included likelihood of earlier detection, more robust follow-up, and therefore benefits to patient care overall.

All samples, whether collected under biobank ethics, or PRISE, were screened prior to inclusion in experiments. As inflammation was not always apparent at endoscopy, samples were preserved as detailed below and a blinded, clinically experienced, GI-specific histopathologist opinion (E.F.) obtained prior to categorization of a sample as 'inflamed' or 'non-inflamed'.

Sample handling and storage

Fresh colonic, blood and FFPE samples were handled and stored as described in **Chapter 2 methods**. Disease state was confirmed through histopathological review of clinical samples stored from the same site and time as research samples.

Fresh frozen samples for spatial transcriptomics Three different methods were trialled for determining the optimal process for storage of tissue samples for spatial transcriptomics : placing samples in a 30% sucrose in PBS solution, samples washed in PBS placed in isopentane subsequently cooled with liquid nitrogen and samples washed in PBS placed in an OCT-filled (**Materials**) cryomould, cooled in an isopentane bath at -80°C. We compared the RNA quality (Agilent RNA Tapestation kit, **Materials**) and tissue architecture preservation of 10uM sections of tissue samples across the three methods.

For n = 2 samples, each split three ways, we determined that the RNA quality and tissue architecture were both inferior for the sucrose preservation medium (Mean RIN less by 1 point), but were equivalent for isopentane and OCT (data not shown). Given that biopsy samples were easier to orient and place in OCT, we opted for this method for storage.

Biopsy samples were therefore washed in PBS and then frozen as above, followed by long term storage in sealed containers at -80°C. Surgical tissue in excess of histopathological requirements was dissected in the lab to isolate the mucosa layer, then frozen in OCT as specified above.

Methods

Biopsy dissociation and fraction enrichment

From prior work in the lab, it was clear that running unenriched 'whole biopsy' samples on a 10X platform was an inefficient method for sequencing cell types as diverse as epithelial, immune and stromal cells at an adequate depth (given significant differences in the RNA content and diversity between epithelial cells and the others). Cell type enrichment was therefore required.

However, we also wished to extract these diverse cell types with minimal adverse effects on viability from the same biopsy samples in order to reliably impute cellular interactions using single-cell RNA transcriptomics.

Finally, given that we wanted to hash samples in order to render the exercise affordable and reduce inter-run variation, we needed to ensure that no digestion method was harsh enough to cleave hashing antibody or key CITE-seq antibody binding epitopes.

Determining digestion protocols compatible with antibody binding We trialled a number of different digestion protocols on PBMCs, comparing cell yield and viability, as well as cleavage of key antibody epitopes – including Type 2 collagenase digestion (**Chapter 2 methods**), Type A collagenase digestion, Liberase digestion, Mouse Umbilical cord kit digestion and Mouse Lamina propria kit digestion, with undigested cells as controls. We determined that Collagenase A, Liberase and Umbilical cord digestion cleaved key epitopes (CD4 and CD8) that would be utilized for analysis and therefore were not suitable. Type 2 collagenase and Lamina propria digestion kits were equivalent for minimal cleavage of antibody binding epitopes, but whereas cellular yield was higher with Lamina propria digestion, but the Type 2 collagenase digest was marginally better for epithelial survival.

Determining optimal method for epithelial, stromal and CD45 fraction enrichment We trialled three different methods – MACS bead-based enrichment, FACS sorting and a hybrid crypt enrichment and tissue dissociation strategy, each detailed below

Methods

MACS bead-based enrichment Colonic biopsies were mechanically digested using Type 2 collagenase (1mg/ml) in biopsy medium (specified in **Chapter 2, Materials**) for one hour into a single cell suspension. Each sample was resuspended in 200uL ice cold MACS buffer (0.5% BSA, 2 mM EDTA in PBS) to a concentration of ~ 1-2x10^4 cells/uL. Working quickly on ice, cells were incubated with 64uL EPCAM+ microbeads (16 uL per 10^6cells), mixed and left on a rotator in the cold room for 30 minutes. 1ml of MACS buffer was added per 10^6cells and cells centrifuged (300gx10min) to remove excess. The supernatant was resuspended in 500uL of ice cold MACS buffer and passed through a pre-wetted LS column within the magnetic MACS separator followed by a flush and washes (as per protocol), with the flow through (containing EPCAM- cells) spun down and resuspended in biopsy medium as the Lamina propria/CD45 fraction. The LS column was then removed from the MACS separator and flushed through with MACS buffer to collect the EPCAM+ epithelial fraction. Each fraction was then stained with hashing antibody as described below for 10X single-cell RNA sequencing, pooled with like fractions and run through the protocol to generate hashed pooled libraries. Analysis of this dataset however, revealed that cellular hashing had failed, and samples could not be demultiplexed. We anecdotally verified that this was a particular issue experienced by other groups with MACS based cell sorting and cellular hashing, for reasons that remained unclear. We therefore could not take this approach forward.

FACS- based enrichment Colonic biopsies were mechanically dissociated into a single cell suspension using Type 2 collagenase (1mg/ml) in biopsy medium over one hour. They were then stained for flow-cytometry based cell sorting as described below for CD45+ cells, and then sorted into epithelial (Live, EPCAM+), CD45+ (Live, CD45+) and stromal (Live, EPCAM-CD45-) fractions. These were counted, pooled, and run as described below for 10X single-cell RNA sequencing to generate hashed pooled libraries. Analysis of this dataset however revealed that epithelial and stromal cell recovery was extremely poor after FACS sorting, which was driven by the rapid acceleration and deceleration cells undergo during this process during single-cell droplet generation and collection in suspension.

Methods

Hybrid crypt enrichment and tissue dissociation Given the limitations of the two methods described above, we adopted a hybrid epithelial crypt and lamina propria digestion protocol. Epithelial crypts were enriched for utilizing the crypt chelation protocol specified for organoid generation (**Methods, Chapter 2**) using HBSS solution supplemented with EDTA and DTT (**Materials**). The supernatant containing epithelial crypts was digested into a single cell suspension by suspending the cells in TrypLE™ Express (**Methods**) for 30 minutes at 37˚C. Dead cells and debris were removed through filtering through 100uM and 40uM meshes to create a single-cell suspension in biopsy medium, which was stained as specified below for 10X single-cell RNA sequencing

The remainder from the crypt chelation protocol was subjected to digestion by the lamina propria kit (as per protocol, 2.35mls Buffer L + 100uL reconstituted Enzyme D + 50uL reconstituted Enzyme R + 12.5ul reconstituted Enzyme A, **Materials**) for 1 hour at 37˚C, followed by filtering through 100uM and 40uM filters to remove dead cells and debris, then suspended in ice cold biopsy medium, then stained for 10X as specified below.

CD45+ and CD3+ fraction enrichment: The entire biopsy was digested as per the Lamina propria kit protocol (described above) for 1 hour at 37˚C, followed by filtering through 100uM and 40uM filters to remove dead cells and debris, then suspended in ice cold biopsy medium, then stained for 10X as specified below.

PBMC handling and isolation

PBMCs were separated from whole blood using density gradient centrifugation employing Lymphoprep, followed by storage in liquid nitrogen as previously described (**Chapter 2 methods**). For experiments, samples were thawed into cold blood medium (RPMI, **Chapter 2 methods**), and stained as described below before loading onto the 10X platform.

Cell staining and FACS enrichment of relevant fractions for sc-RNAseq

Cell staining 0.5 million cells (Epithelial, Lamina propria, PBMC or full biopsies) were suspended in 50uL of biolegend cell staining buffer (**Materials**) then blocked with 5uL of Trustain FcX (**Materials**) for 10 minutes at 4°C. Following this, they were incubated with an additional 50uL of antibody mix for 30 minutes at 4°C. The cells were then washed and then subjected to FACS sorting or pooling as per the relevant protocol. The antibody mix comprised (per 0.5 million cells of sample): 0.75ul (0.37ug) of Totalseq-C CITE-seq/hashing antibody (anti-CD4, anti-CD8, anti-PD1, anti-CD45RO, anti-NKp46, anti-CCR6, anti-LAG3, anti-TIM3, anti-CTLA4, anti-CD137, anti-CD11b, anti-CD11c, anti-CD103, Hashing antibodies TotalseqC #1-#5, **Materials)**, 0.5uL of anti-CD45 (APC, 5B1, **Materials**), 2uL anti-CD236 (FITC, HEA-125, **Materials**), 1uL Anti-CD3 (PE-Dazzle594, UCHT1, **Materials**) as required for individual experiments.

FACS enrichment Cells were stained as described above, washed to remove excess unbound antibody and sorted using a BD FACSAria III Cell Sorter and/or BD Fusion Cell Sorter (for five samples, both sorters were run in parallel to minimize time each sample was kept following sorting). The accuracy of sorting was confirmed using beads and cells. Cells were gated based on size using Forward and Side scatter, followed by identification of singlets using FSC-H and FSC-A. After gating on live cells, CD3+/CD45+ cells were sorted, aiming to collect 100,000 cells into eppendorfs containing 50uL pure FCS. Cells were spun down at 600g x 2 minutes at 4°C in a 96 well round-bottom plate and resuspended in 100uL sorting medium (2% BSA, 0.01% Tween in PBS) at 4°C. Following a count, cells were pooled in a 1:1 ratio across all samples, spun down and resuspended to a concentration of 10^6 cells/ml, followed by immediate loading onto the 10X scRNA platform.

Single-cell RNA sequencing

Given the timeline over which these samples were collected and the iterative nature of the project as progressively acquired datasets were analysed, two separate 10X chemistries came to be utilized across the entirety of the experiment.

Methods

Tissue CD45+ cells In the first iteration, live CD45+ cells were sorted from HC, UC_I, CC_I and CC_NI as described above, following which 10,000 cells from each sample were loaded into the 10X scRNA platform without pooling, recovering approximately 4,000 cells per run. GEX Libraries were prepared using 3’ 10x Genomics Library Kits (10x Genomics, CG000183, Rev A).

Epithelial/Lamina Propria/matched PBMC cells: In the second iteration, epithelial and lamina propria (CD45/stromal) cells from up five different samples (HC, UC_I, UC_NI, CC_I and CC_NI) were pooled at 1:1 ratio. Each pool was loaded on a channel of the 10X Chromium single-cell platform, one for the epithelial fraction and the other for the lamina propria fraction. A superloaded input of 30,000 single cells per pool was added to each channel with a recovery rate of approximately 10,000 cells per channel. Libraries were prepared using 5’ 10x Genomics Library Kits (10X Genomics, CG000186,Rev A). For matched blood samples, cells from four different samples (HC, UC_I, CC_I and CC_NI) were pooled at 1:1 ratio. For each pool 30,000 cells were added to each channel, recovering approximately 8,000-10,000 cells. Gene expression (GEX), T-cell receptor (TCR) and Antibody-derived Tag (ADT) Libraries were prepared using 10x Genomics Library Kits (10X Genomics, CG000186,Rev A) for each of the lamina propria and PBMC pools.

CD3+ cells: In the third iteration, CD3+ cells from tissue and blood were sorted as described above. Cells from three samples were counted, pooled at a 1:1 ratio, and 20,000 cells from each pool were loaded into the 10X scRNA platform, recovering approximately 10,000 cells per run. GEX, ADT and TCR libraries were prepared using 5’ 10x Genomics Library Kits (10x Genomics, CG000208,Rev F).

Sequencing of libraries was carried out on an Illumina NovaSeq 6000 platform with 1% PhiX by a commercial agency (Novogene) given the sequencing requirements of the relevant libraries and cost considerations.

Spatial Transcriptomics

The 10X Visium platform was used for extracting matched H&E and spatial transcriptomic information as per published protocols (specified below).

Methods

Briefly, fresh frozen samples as described above were initially screened to identify blocks of high RNA quality. In order to do so, RNA was extracted from 5 x 10uM sections taken from blocks in OCT placed directly into Buffer RLT at 4°C which was followed by column-based RNA extraction (RNEasy plus Micro Kit, **Materials**). RNA quality and quantity were assessed using a high sensitivity RNA ScreenTape assay in a 4200 TapeStation (per protocol, **Methods**). All the samples retained a RNA Integrity Number (RIN) > 8.5.

Following on from this, we determined the optimum tissue permeabilization time for mucosal gut biopsy samples using the Visium Tissue Optimization kit and protocol as published (CG000238 Rev D; Imaged on a Leica DMI8 inverted microscope). Briefly, 10uM thick serial sections from an inflamed biopsy sample (Nancy Score 4) on a tissue optimization slide were incubated with permeabilization enzyme from a range of 3-24 minutes as per the described protocol The timepoint which yielded an optimum trade-off for RNA recovery across epithelium and stroma for colonic tissue was determined as being 12 minutes, which was used for the Visium ST protocol below.

We then proceeded to perform 10X Visium Spatial Transcriptomics on 10uM thick OCT-embedded biopsies and resections derived from HC, UC_I and CC_I patients as per the published protocol (Visium 10X, CG000239,Rev D). Briefly, sections were placed onto slides and processed to yield H&E sections (imaged at 10X on a Zeiss Axioscan z.1 slidescanner). These sections were then processed with a permeabilization time of 12 minutes to yield spatial transcriptomic libraries at a resolution of 55uM per spot (Visium 10X, CG000239,Rev D). Libraries were sequenced locally on an Illumina NextSeq platform.

Immunofluorescence and imaging of sections

Deparaffinized 4uM slices from biopsies were treated with pH 9 heat-mediated antigen retrieval in Tris-EDTA buffer for 30 minutes, then were blocked for 1.5 hours at room temperature with 2.5% goat serum (**Materials**). Sections were then incubated with optimised dilutions of primary antibodies (CD103 - Ab 238010, 1 in 350 dilution; E-cadherin - 24E10, 1 in 400 dilution; FOXP3 - HPA 045943, 1 in

4000 dilution; FABP1 - HPA028275, 1 in 800 dilution; iNOS2 - MAB9502, 1 in 1600 dilution; CD163 - Ed-Hu1, 1 in 200; Neutrophil elastase - ELA2, 1 in 800 dilution; Cleaved caspase-3 - ASP-175, 1 in 100 dilution; E-cadherin - 4A2, 1 in 50 dilution; **Materials**) in combinations of mouse and rabbit anti-human antibodies in PBS supplemented with 1% BSA overnight. Following extensive washing with PBS-T, slides were incubated with appropriate secondary goat antibodies (1 in 500 dilution) conjugated to 488 or 647 fluorochromes for 1 hour at Room Temperature (RT) in 1% BSA/TBS + 1ug/ml DAPI. slides were incubated with Vectashield Truview (**Materials**) for 4 minutes to reduce autofluorescence, followed by a 5 minute 1ug/ml DAPI incubation, and covering with Vectashield Immunofluorescence preservation medium. Each step was separated by multiple washes in TBS/TBS-Tween.

Slides were imaged at 20X on a Zeiss Axioscan z.1 slidescanner within 24 hours of staining, being kept at 4°C until acquisition. T.G. developed a standardized exposure and acquisition protocol that was used across all sections. A control 'secondary antibody only' section was used to correct for autofluorescence prior to image export at >75% original size and resolution TIF format used for image analysis.

Detecting nivolumab-bound cells using Flow cytometry (FACS)

The Davis group developed the non-competitive PD-1 binding clone as part of an independent research stream. Previously published work had demonstrated that the EH12.2H7 clone (utilized for CITE-seq) bound within the PD1-PDL1/L2 interaction site. Given that Nivolumab and Pembrolizumab also interfere with this binding, by induction (and on the basis of previous work) Nivolumab and Pembrolizumab-bound cells would be unable to bind the CITE-seq clone[170–172]. The Davis group had developed their clone to bind to PD-1 distinct to the binding site of either Nivolumab or Pembrolizumab, and provided us with a crystal structure, the schematic of which is presented in **Chapter 3**.

We validated the binding property of the Davis Noncomp PD-1 clone by comparing its binding versus Nivolumab binding to PD1. We first conjugated Nivolumab to APC and the Non-comp PD1 antibody to

AF488 using a commercial kit (ab201807-300ug, **Materials**). We then assessed the binding of these antibodies to a Jurkat T-cell line engineered to over-express PD1, confirming that both these antibodies had equivalent binding affinity to cells expressing PD1.

In order to validate Nivolumab interfered with the binding of the CITE-seq EH12.2H7 clone, but not Noncomp-PD1 in biopsies, we utilized FACS. Biopsies from healthy donors were digested as previously described using the lamina propria kit protocol (**Materials**) into a single-cell suspension. These were then stimulated using CD3/CD28 dynabeads (**Materials**) at a concentration of 10uL per 1 x10^6 cells for 24 hours in order to increase the expression of PD-1 molecules on T cells. Each biopsy was then split into two, resuspended in ice cold staining buffer and incubated with either Noncomp-PD1 (AF647, 10ug/ml) and CITE-seq clone EH12.2H7 (FITC, 10ug/ml) or the same two antibodies with added unlabelled Nivolumab (10ug/ml) with staining for CD3 across all samples for 30minutes at 4°C. The same samples were incubated with isotype control antibodies for both. Samples were then acquired on an LSR II flow cytometer.

In order to detect cells in biopsy samples that were bound to Nivolumab and Pembrolizumab given to the patient in vivo, we utilized FACS, but with a more extensive staining panel. Biopsies from patients with CC_I colitis were digested into a single cell suspension using the lamina propria digestion kit protocol described above. These cells were then stained for live dead (Zombie Aqua 1:100, **Materials**) discrimination in PBS for 12 minutes at room temperature, followed by suspension in ice cold staining buffer. They were then incubated with antibodies for Trustain FcX, CD3, CD8, PD1 (EH12.2H7, 10ug/ml), CCR6, CD103, CXCR5 and Noncomp-PD1 (10ug/ml), along with isotype controls for FITC (PD1) and AF647 (Noncomp-PD1) for 30 minutes at 4°C in the dark (clones specified in **Materials**). Following two washes in staining buffer, cells were lightly fixed with FACS Cytofix buffer (**Materials**) for 12 minutes on ice, before washing and resuspension in FACS Cytoperm buffer (**Materials**) containing antibodies against Ki67 for 30 minutes in the dark on ice. Following removal of excess antibodies with Cytoperm washes, cells were resuspended in cell staining buffer and acquired on a LSR X20 flow cytometer.

Timecourse analysis of incubation with Nivolumab (FACS)

We wished to check that incubation with Nivolumab itself was not sufficient to induce cellular replication or apoptosis. Briefly, we incubated freshly isolated PBMCs (as specified above) from n = 5 healthy volunteers. We incubated these with CD3/CD28 dynabeads (**Materials**) at a concentration of 2.5uL per 1x10^6 cells for 7 days, sampling at baseline, 3 days, 5 days and at the final 7 day timepoint with and without the addition of Nivolumab (**Materials**) at a concentration of 4ug/ml. In a separate experiment, we assessed the effect of Nivolumab on healthy-volunteer derived PBMCs at a range of concentration from 0.4ug/ml to 200ug/ml at a single 7 day timepoint (with and without the presence of CD3/CD28 dynabeads). Dynabeads were utilized for these experiments in order to induce the expression of PD1 in order to bring out any effects of Nivolumab, simulating inflammation. For both these series of experiments, the readout was assessed for Ki67, CD3, CD8 and Live-Dead staining (clones specified in **Materials**) by FACS.

Organoid co-cultures with TNF and Interferon

Organoid establishment and propagation Human colonic organoids were generated using established published protocols as described above for IL26 co-culture.

Briefly, frozen biopsies were defrosted into warmed biopsy medium (as constituted above). These were chopped up into fine fragments to improve surface area for epithelial recovery, and placed in warmed chelation medium (96% HBSS, 1% Pen/Strep, 1% HEPES (1M stock), 1% 5mM EDTA (0.5M stock), 2mM DTT, 1%FCS). Epithelial crypts were chelated off from the lamina propria at 37dC with gentle agitation for 1 hour using a shaker. The isolated crypts were spun down (400g x 5min @ 4°C) and resuspended in cold DMEM F12/BSA medium (1% BSA in DMEM F12) on ice, aiming for 1 million cells/ml. Thawed Matrigel at 4°C was added in a 1:1 ratio to the suspension, mixed and 50ul rapidly dispensed into each well of pre-warmed flat bottom 24 well plate at 37°C. The pellet was allowed to set for 10 minutes, following which 500uL of conditioning medium (CM) supplemented with 1:1000 Y-inhibitor (for the first passage) was added to each well. Conditioning medium was freshly made up as per the original published protocol for propagation of colonic organoids. Medium was changed every other day until organoids required passaging, when the Matrigel was dissolved using cold dissociation medium, spun down 400g x 5min at 4°C, refreshed with fresh Matrigel/DMEM F12-BSA of double the volume and re-plated as before.

Co-culture experiments All samples were passaged once, and experiments begun once the spherules were actively increasing in size 3 days after passaging. The 3 arms were managed as follows: the “Medium only” arm was incubated with conditioned medium (CM), changed every other day for 7 days, followed by differentiation medium (DM) for 3 days, followed by harvest and RNA extraction. The “Acute stimulation” arm received medium changes as for the “Medium only” arm, but for the last 24 hours of the differentiation medium step, TNFa (10ng/ml) and IFNg (10ng/ml) were added. For the “Chronic stimulation” arm, TNFa (10ng/ml) and Interferon gamma (10ng/ml) were added to medium from day 2 onwards, refreshed with every 48 hourly medium change (8 days total). This concentration of reagents was chosen with recourse to literature and our estimation of the concentration that could be expected in tissue. As before, all wells were incubated with DM for 3 days to allow development of the non-stem cell epithelial compartment. RNA extraction was carried out using column-based RNA enrichment using the Qiagen RNEasy Mini kit. The organoids were imaged at 10X with an Sartorius Incucyte scanner at 24 hour-intervals.

Once extracted, the concentration of RNA was determined with a nanodrop system, and cDNA conversion and RT-PCR done as per the quantitative real-time PCR protocol described in **Chapter 2**.

Epithelial-stromal-T cell co-culture organoid model system

Given the paucity of available model systems to explore the effect of checkpoint inhibitor action on immune cells, we wished to develop a model system within which this could be understood.

Methods

Experiments with tissue biopsies and PBMCs (fresh and frozen) incubated with a variety of Nivolumab concentrations (0.7ug/ml-70ug/ml) both with and without stimulation of the TCR (via CD3/CD28 dynabeads) demonstrated no additional activation (as assessed by TNF) at either the RNA or protein level in the presence of Nivolumab.

Other groups had employed an air-liquid-interface model system to replicate some of the effects of Nivolumab action in the context of malignancy[173,174]. We trialled multiple variations of air-liquid interface (ALI) and standard Matrigel 'dome' conditions in combination with epithelial seeding or biopsy fragmentation, also comparing outcomes when using a typical organoid[15] versus a modified medium[175] for propagation, for both fresh and frozen healthy colonic samples.

We assessed the suitability of each method by assessing cellular survival using FACS and a high throughput sampler (HTS) system. Briefly, every sample was divided into a number of equal fractions by weight/number of biopsy fragments. One fraction was considered an original 'starting' fraction, with the remainder 'result' fractions cultured in the variety of permutations described above. The HTS system is capable of aspirating a given volume of solution for analysis by FACS. After digesting the original and various result fractions using the lamina propria protocol described above (followed by staining for EPCAM, CD45, CD3, CD8 and CD4), sampling an equivalent proportion of each allowed for an accurate measurement of total cell number by cell type. We chose a timepoint of 7 days for measurement, given this was sufficient for inducing Nivolumab-induced activation within this model system[176]

We were able to ascertain that previously frozen samples were unsuitable for such analysis, with extensive stromal and immune cell death within the first 48 hours after thawing. Epithelial propagation and survival was however minimal in the context of simple biopsy fragmentation as practiced by the Kuo lab[174], perhaps unsurprising given their experiments were performed in the context of malignant cells. However, survival of stromal and T cells was reduced if the sample was completely dissociated into a single cell suspension, leading us to conclude that maintenance of physiological cell-cell contact was important for these cell types.

We also found reduced propagation of epithelial cells in an air-liquid-interface (ALI) system as compared to standard Matrigel domes. There was no significant difference in epithelial survival between standard and modified media, so we opted for a standard medium for consistency in methodology. We added IL-2 to improve the survival of T cells[176].

Taken in summary, we opted for a hybrid system whereby a sample was chopped into fine fragments utilizing a scalpel (capable of passing freely through a P1000 pipette tip), then subjected to crypt chelation as previously described. The chelated fraction was then combined back with the fragments prior to suspension in standard Matrigel domes and cultured in the presence of standard medium supplemented with IL-2. This allowed epithelial fragments to propagate in Matrigel whilst maintaining close proximity with stromal cells, which in turn maintained some physiological positional relationships with each other and immune cells. As assessed by FACS and HTS sampling, we were able to validate that all three broad compartments (epithelial, stromal and immune) survived upto the 7 day timepoint from freshly cultured biopsy samples.

Once we had optimized cell survival, we then wished to ascertain whether cells still behaved in a physiological manner in this model system. To this end, we investigated the effects of T cell activation on stromal and epithelial counterparts using a phytohaemagglutinin (PHA)-driven model[177]. We opted for this given its molecular size and likelihood of penetrating a Matrigel matrix (as opposed to antibody-driven activation, penetration of which was likely to be limited by Matrigel). The concentrations of PHA were chosen with recourse to literature, and output measured RNA transcripts as per our single-cell transcriptomic data. Nivolumab concentrations were chosen with recourse to primary phase I data[178]

This system is by no means definitive as the impact of the pandemic meant difficulties in acquiring fresh colonic biopsy samples in sufficient numbers to fully optimize the seeding density of this co-culture system, which we postulate is partially responsible for survival and the variability in results. We would also like to explore longer incubation times, incorporation of ipilimumab (anti-CTLA4) and ideally correlate results with the proportion of PD-1 and CTLA-4 positive T cells in the original sample

Histopathology analysis

Histopathological analysis was carried out in a blinded fashion by an experienced consultant gastrointestinal pathologist who routinely assesses colonic inflammation (E.F.)

Briefly, deparaffinized sections from a randomly selected proportion of patients with checkpoint inhibitor induced colitis, ulcerative colitis and health were stained utilizing a Vector kit and standard haematoxylin and eosin protocol (**Materials**). Brightfield images at 10X were acquired with a Zeiss Axioscanning widefield microscope. Anonymized images were analysed by the histopathologist to quantify the number of lymphoid follicles.

The data was corrected for image area (calculated using the area tool on opensource ImageJ software) and analysed, generating figures utilizing Graphpad Prism software.

Image analysis

Image analysis was performed by N.K.A in conjunction with T.G. using the Visiopharm Integrator System (VIS) platform (v 2019.07.3). Image analysis protocols were implemented as Analysis Protocol Packages (APP) in VIS. Several APPs were designed using threshold classification to quantify the slides

Briefly, images generated by T.G. from a random selection of checkpoint inhibitor, healthy and ulcerative colitis samples as described above were anonymized and transferred electronically to N.K.A, who performed the analysis as per methodology decided by T.G. This blinded analysis was transferred back to T.G., who performed the analysis on all data using Graphpad Prism software, using an unpaired t-test for non-parametric data (Kolmogorov-Smirnov test), with significance being defined as $p < 0.05$. Data are presented as mean with error bars denoting standard error of mean.

Methods

CD103 E-cadherin DAPI and Cl-caspase3 E-cadherin DAPI analysis The total E-cadherin positive and negative area in pixels was determined using the DAPI area as a baseline (DAPI stains all nucleated cells). On the same image, the area in pixels positive for CD103/cleaved caspase3 in E-cadherin positive areas (Intra Epithelial Lymphocytes (IEL)/epithelial apoptosis) and E-cadherin negative areas (Lamina Propria (LP) lymphocytes/stromal apoptosis) was measured. The area was divided by the total area occupied by the epithelial or stromal cell population in order to correct for section-to-section variation.

FABP1-associated CD163 DAPI and iNOS2 DAPI analysis From previously published data as well as our own transcriptomic analysis, we determined FABP1 was upregulated in a graded fashion from the crypt base to the crypt top[19]. We first validated this in immunofluorescence of longitudinal sections of crypts, and standardized the fluorescence signal intensity to distinguish crypt tops from bases across sections, in order to be able to analyse sections where the orientation was not optimal. For each section, utilizing this intensity of signal, we determined whether the crypt area detected was from crypt base, mid-crypt or crypt top. For each area (crypt base, mid-crypt and crypt-top), we designated a standard 'test zone', which was delineated by half the average distance between crypts in pixels (to avoid double counting cells next to adjacent crypts), annotating this as a 'peri-crypt zone'. Within this test peri-crypt zone, we then counted the area positive for CD163 (FABP1-high crypt top, mid-crypt or FABP1-low crypt base M2 macrophages) or iNOS2 (crypt top, mid crypt or crypt base nitric-oxide synthetase activity) as for each respective analysis, correcting for total area using DAPI as before.

FOXP3-associated CD163 macrophages FOXP3 presented a challenge in detection given its relative paucity, which meant even with autofluorescence quenching, small false positive spots were present in each tissue. To circumvent this, we utilized a data analysis approach that ignored such false positives. Briefly, around each CD163-positive area (M2 macrophage) we designated a standard 'test zone' double the radius of an average M2 macrophage (reasoning this represented a maximum diameter within which cells could be expected to be interacting, with any secreted cytokine approaching ~1/50th of the concentration at the centre). Within this 'test zone', we only counted a FOXP3-positive Treg cell if the FOXP3 signal was wholly ensconced within a halo of DAPI (i.e. intranuclear), discounting false positive autofluorescence that did not obey nuclear boundaries. Analysis was again carried out with correction for DAPI-derived section area.

Bioinformatics analysis

Bioinformatics analysis was conducted by A.A.A, C.L. and R.B. in discussion with T.G. and A.A. The methods described herein are taken from their descriptions of the processes involved.

Raw sequencing data processing

All raw sequencing data was converted to from bcl to fastq format using Illumina bcl2fastq software, with upto one mismatch allowed in each sample index barcode. Raw sequence reads were quality checked using FastQC software[179]

Raw 10X scRNA-Seq, CITE-Seq and spatial transcriptomics data processing

For each sequenced scRNA-Seq pool, unique molecular identifier (UMI) counts were processed and aligned against hg38 human reference genome (10x reference: refdata-gex-GRCh38-2020-A) using Cellranger software from 10 × Genomics[180]. Matched protein CITE-Seq and hashing antibodies were processed together with scRNA-Seq as matched feature barcoding libraries. Antibody tag UMI counts were summarised using a joint feature barcoding reference containing sequences from Totalseq-C hashing and protein expression targets, individual tag sequences being derived from the Biolegend website.

Hashed sample de-multiplexing

Hashing antibody UMI count matrices were filtered to keep only 10x cellular barcodes from droplets passing QC based on mRNA expression profiles, as described below. Non-hashing antibody counts and hashing tags not present within any given pool (as hashed sample numbers varied between reactions between 3-5) were also filtered out for each pool individually. Each filtered matrix was used to demultiplex samples[181]. Counts were first normalized using centred log ratio transformation and an initial clustering solution was obtained using clara k-methods clustering with k = 1 + number of samples in the pool. A negative binomial distribution was then fit for each tag and a positive tag threshold was defined as 99th percentile of the normalized UMI counts, with cells below this threshold considered negative for the tag. Cell sample-of-origin was then assigned for each cell based on individual hashtag thresholds with doublets/multiplets defined as cells positive for multiple tags and filtered out from further analysis. A minor fraction of all cells were found negative/below tag threshold for all hash tags and were also filtered out, following inspection of their mRNA-cluster distributions. Untagged cells correlated with lower total mRNA content cells and did not segregate with any particular cluster and thus likely contained unstained/dying cells or free nuclei that have lost their cytoplasm during sample processing. In each case, we then further examined whether sample demultiplexing was correct by ascertaining that the expression of sex-specific genes, such as XIST, segregated correctly with sample-of-origin assignments.

10x scRNA-Seq data analysis

Raw UMI count matrices were imported into R for processing. In order to distinguish cells from empty droplets, cells were called using the ‘emptyDrops’ function from DropletUtils[182]. Raw count matrices were corrected for Illumina index swapping using ‘swappedDrops’[183].

In order to remove poor quality dead or dying cells, we filtered out droplet barcodes for which 1) a high percentage of total UMIs originated from mitochondrial RNAs as well as 2), a low total UMI barcode count overall. The thresholds were derived individually for cells within each compartment as total RNA content and mitochondrial RNA content are highly cell type dependent.

For each individual 10x reaction, R package Seurat[184] was used to normalize expression values for total UMI counts per cell. Highly variable genes were identified by fitting the mean-variance relationship and principal-component analysis used to reduce dimensionality. Scree plots were used to determine the number of principal components to use for clustering analyses for each pool. Cells were then clustered using Louvain algorithm for modularity optimization using a kNN graph as the input. Cell clusters were visualized using a UMAP algorithm[185] with principal components as input and n.neighbors = 30, spread = 1 and min.dist = 0.1.

Cells from separate pools and reactions were merged together, with batch and protocol effect signals being corrected using the harmony algorithm[186]. Merged pool clusters were compared with individual pools to ensure cellular heterogeneity was not lost in the process of batch correction.

CITE-seq protein analysis Count matrices were imported into R as a separate assay in Seurat objects. Only those cells passing QC based on RNA expression analysis were retained for analysis. Count data was normalised using a centred log ratio transformation within each reaction type (i.e. sorted CD3s separately to lamina propria cells).

Normalised tag expression density distributions were used to define cut-offs for positive and negative staining. This was particularly useful in the context of certain cell phenotypes (e.g. Th17 and Tc17 T cells) that are more difficult to distinguish transcriptomically.

scRNA-Seq cell populations were annotated using a combination of known marker gene and protein expression profiles (summarised in **Supplementary Table 1** for each population/cluster) and using previously published scRNA-Seq reference atlas datasets in the colon and PBMCs[19,20,29,187,188].

Trajectories for pseudotime analysis were calculated using the Monocle 3 algorithm[189] for batch-corrected data as processed above. Seurat objects were converted to Monocle 3 cell_data_set objects and CD4+ and CD8+ cells populations clustered separately. The start of the trajectory was denoted as the node within the naïve cell cluster and pseudotime calculated along the trajectory.

Principal component (PCA) analysis was carried out on all de-hashed cells passing QC. These were used to calculate pseudobulk counts for each sample by summing across all UMI counts for each gene for each cell for a given sample. Immune, epithelial and stromal cells were each analysed separately to avoid transcriptional changes in one lineage masking/affecting the changes in another. Count data was normalised using sample normalisation size factors which were computed using the DESeq2 R package[190] median ratio method. The top 1000 most variable genes were then selected and used to compute the PCA for each sample.

Condition specific changes To detect condition specific clusters, the count for each cluster was normalised to the total number of cells within that compartment, with the proportion of cells compared using a two sided Wilcoxon test, with p-values <0.05 considered significantly different. For cell type populations which exist on a continuum rather than discrete clusters (e.g. T cell subtypes), we carried out graph-based differential abundance analysis using the miloR R package[191] with k = 10 nearest neighbours used for neighbourhood definitions.

Transcription factor (TF) modules were detected using the R package SCENIC across all datasets. We utilized the RcisTarget database that contains transcription factor motif scores for gene promoters and transcription start sites. The cell gene expression list was filtered to include only those genes in the database. The remaining genes were used to compute a gene-gene correlation matrix for co-expression module detection using a random forest GENIE3 algorithm[192]. TF network analysis was done using the SCENIC R package[193] to detect co-expression modules enriched for the target genes of each TF. The AUCell package[193] was used to compute a score for each TF module in each individual cell. We identified condition or cluster-specific TF modules using generalised linear models, with p values being adjusted for multiple testing using Benjamini-Hochberg multiple testing correction.

Receptor-Ligand interactions Interactions between all single cell clusters identified from scRNA-seq data were inferred using the ‘Cellchat’ R package[194]. Each condition was analysed separately and then the probability of interactions in between conditions compared to impute significant condition specific changes. Circos plots were used to visualise specific interactions.

GO pathway analysis Gene sets for Gene Ontology[195] and REACTOME pathways[196] were downloaded. The R package ‘AUCell’ was used to calculate an activity score for each cell, blocking for gene detection rate as a co-variate. P values were adjusted for multiple comparison (Benjamini-Hochberg).

Methods

TCR analysis Single cell TCR clonotypes were assembled using Cellranger VDJ software. As VDJ enrichment libraries were not available for intra-epithelial T cells, TCR sequences were inferred from 5' transcriptomic data using TRUST4 software[197]. In both cases, single-cell barcoes were used to link VDJ clonotypes to gene expression data. The VDJTools[198] software was used to compute TCR repertoire statistics. Assembled TCR CDR3β amino acid sequences were used to query VDJdb[199], McPAS[200] and TBAdb for overlap with known and/or public TCR sequences. GLIPH2[201] was used to predict antigen specificity of TCR clonotypes. To study clonal dynamics and expansion, clonotypes were defined by CDR3αβ nucleotide sequences. T cell clonality was defined by computing Shannon entropy and overlaps of TCR repertoires between populations was calculated using Morisita's index using the R divo package. As CD4s were much less clonal, a much larger sampling per individual would be required to infer clonal dynamics.

Spatial transcriptomics data analysis

Raw UMI count spot matrices, spot-image co-ordinates and scale factors were imported into R, with filtering out of spots that did not have overlying tissue. Tissue spots with low RNA content were filtered out as these were likely to represent under-permeabilization. The majority of such spots were section specific/corresponded to tissue artefacts.

Raw UMI spot counts were normalized using a regularised negative binomial regression[202] to better account for variability in total spot RNA content. Clustering was performed using the Louvain clustering method and visualised as UMAPs.

For integrative data analysis, individual slides were integrated using the harmony algorithm. Merged data clusters were compared with those obtained from individual slides to ensure no heterogeneity was lost. Regions were annotated depending on their transcriptional signal, and cross-checked with H&E features as determined by a gastrointestinal histopathologist blinded to gene expression.

Crypt axis score Due to the nature of biopsy specimens, orientation of the tissue section on the slide is extremely variable, so it is extremely difficult to determine the depth of the epithelial cross section in some slides from histology alone. Each ST spot was therefore scored using a transcriptomic epithelial crypt axis score[19] which allowed us to annotate each epithelial spot with its crypt depth (which is not possible by histopathology alone), and validated in resection sections where crypts were longitudinally orientated. 'Upper crypt' spots were defined as the topmost two spots (corresponding to ~100uM) of the upper crypt layer.

Cell type prediction Cell type prediction (within each spot) was calculated using factor analysis in Seurat using the single cell RNA-seq data as reference. Data from separate compartments was merged into a unified reference dataset. Two levels of cell type annotations were retained – Broad (T cells/B cells etc.) and sub-clusters (Naïve, Tregs, Tc17 etc.) For broad cell type predictions, the first 30 components of the scRNA-seq and ST datasets were aligned, whereas the first 60 components were used for finer sub-clusters. We checked that spot composition was robust, we used two additional published methods to check our predictions – SPOTlight[203] and RCTD [204]. Positive spots were classified as those with a cell type prediction probability > 0.

Cell type co-occurrence analysis We calculated all pairwise cell type prediction probability scores across all slides. Undirected, edge-weighted cell type networks were constructed from the correlation matrix from significantly positively correlated pairs ($p < 0.01$) using the 'igraph' R package[205], after filtering out cell self-correlation signals and low correlation edges with an $r < 0.15$. Networks were visualized using 'ggraph' R package using a force-directed Fruchterman-Reingold layout.

Spatial ligand-receptor analysis To identify cellular signalling events, ligand-receptor databases were downloaded from 2 sources[206,207]. Given that the gene detection rate for ST is lower than for scRNA-seq, we hypothesized we would encounter greater dropout. Therefore for each ST spot we also considered (weighted) ligand-receptor scores for adjacent spots depending on their distance from the spot analysed. We then randomly shuffled all spot locations across all slides across 100 permutations to calculate an empirical 'background' distribution that would be expected if there was no location specificity. Then for each spot and ligand-receptor pair, a p-value was computed (with a multiple testing Benjamini-Hochberg correction) based on the difference between the observed and this empirical 'background' expression. Spots were considered significant at a p value of <0.05.

GO pathway analysis Gene sets for Gene Ontology[195] and REACTOME pathways[196] were downloaded. The R package 'AUCell' was used to calculate an activity score for each ST spot, blocking for gene detection rate as a co-variate. P values were adjusted for multiple comparison (Benjamini-Hochberg).

Additional resources

Supplementary data mentioned in the thesis is available on the **Mendeley Data** Portal.

Accession:

https://data.mendeley.com/datasets/k4zwxxm8zf/draft?a=0d8e5a1f-9905-4121-a0e9-d06fd3b22b92

DOI: 10.17632/k4zwxxm8zf.1

Chapter 1

Clinical features of checkpoint inhibitor-induced gastrointestinal disease

Introduction

Background

Briefly, immune checkpoints were discovered in the 1990s as a class of cell surface receptors that are predominantly expressed on T cells (with a small degree of expression reported on NK, Myeloid and B cells)[208]. Their expression is constitutive on a small number of cells in health, but is markedly induced on exposure to antigen/inflammation. Activation of these receptors acts in a negative feedback loop through a variety of mechanisms to inhibit inflammation[72]. In the last decade, inhibition of these pathways in a clinical context through the use of "checkpoint inhibitor" drugs has become standard of care for the treatment of multiple forms of cancer[88]. Off-target effects most commonly include colitis (checkpoint-inhibitor induced colitis/ CC colitis) among a protean list of autoinflammatory conditions[105], which being a novel clinical entity, has been poorly characterized till date.

Immunotherapy in clinical use

Anti-cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) (ipilimumab, tremelimumab), anti-programmed cell death protein-1 (PD-1) (nivolumab, pembrolizumab, cemiplimab), and anti-programmed death ligand-1 (PD-L1; avelumab, atezolizumab, and durvalumab) are the three types of checkpoint therapy currently in widespread clinical use. LAG-3 inhibitor therapy (relatlimab) was introduced in early 2022.

Immunotherapy is widely used as first line treatment in metastatic melanoma, either as combination therapy (anti-PD1 and anti-CTLA4) or monotherapy (anti-PD1). For other cancer types such as renal, prostate and non-small cell lung cancer (NSCLC), it is currently employed after failure of more conventional regimens, where anti-PD1 inhibition is most commonly used, either on its own or in combination with chemotherapy.

The use of checkpoint inhibitors as single treatments and in combination regimens for a growing number of malignancies is expected to increase[89,209]. Drugs that block other checkpoint inhibitor pathways (such as TIGIT) are in clinical development. The first agent blocking LAG-3 was demonstrated to be effective in the treatment of metastatic melanoma in early 2022[92,210]. In cancers where combination therapy (either anti-PD1/CTLA4 or anti-PD1/LAG-3) has been trialled, it has proven more efficacious at inducing and maintaining remission in comparison to monotherapy anti-PD1 or anti-CTLA4 inhibition[210,211]

Current understanding of checkpoint inhibitor-induced colitis

Although combination therapy is more effective, it also (insofar as the most data is available for anti-PD1/CTLA4 regimens) related with a considerably greater incidence and severity of irAEs (including colitis)[82,83].

To identify and assess the severity of diarrhoea and colitis, oncological guidelines were developed based on clinical trial data which advocated using the National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE)[152,159,212,213]. The GI CTCAE is primarily based on clinical symptoms (e.g. stool frequency over baseline for assessing diarrhoea; and abdominal pain, blood in stools and peritoneal signs), which in turn are indicative of life-threatening consequences. The efficacy of CTCAE in terms of diagnosis and prognostication of CC colitis has not been established. Additional immunosuppression could be offered early in the illness course if we could identify patients who are likely to have a protracted course at disease onset.

Clinical efforts have focused on discovering objective markers that appropriately measure disease severity and improve clinical decision-making in idiopathic inflammatory bowel disease (IBD). For ulcerative colitis (UC), these include either mono-modal validated endoscopic scores like the Ulcerative Colitis Endoscopic Index of Severity (UCEIS)[214] or histopathology scores such as the Nancy index[215], or combined clinical and biochemical scores like Truelove and Witts criteria[157] for acute severe ulcerative colitis, all of which correlate with outcomes such as the need for rescue therapy or colectomy[216,217]. At the time of initiating our study, no such markers were available for checkpoint inhibitor-induced colitis (CC), although since then, endoscopic and histological evaluation has been shown to correspond with illness outcomes[218,219].

Regarding treatment, as per oncology guidelines, corticosteroids are recommended as a first-line treatment for CC colitis, with anti-TNF therapy (infliximab) reserved for non-responders[152,212,213]. In reality, there is a lot of diversity in how infliximab is started, how many doses are given, and how often it is given. There is no consensus on when or how to escalate therapy for CC colitis patients with both steroid-responsive and steroid-dependent illness. There is also evidence that infliximab use can cause cancer progression[220]. We therefore sought to determine the effects of infliximab therapy on patients.

There are also case series of patients being given drugs other than anti-TNF agents as rescue therapy – this includes drugs such as Vedolizumab[132,221] (that inhibit trafficking of immune populations), as well as JAK-STAT inhibitors[131] – and we looked for data in our analysis to see if we could derive any insights about these patients.

Aims

Across two tertiary referral centres, we studied a cohort of 1074 ICI-treated patients to characterise the real-world incidence, assessment, and therapy of CC colitis in order to:

1. Determine risk factors for development of disease

2. Determine factors that would aid in prognostication – assessing both current CTCAE criteria as well as IBD markers of severity – biochemical (CRP, Albumin and Haemoglobin) and endoscopic – (UCEIS and Nancy Scores)
3. Describe the effect of current treatment strategies on management and an optimum management strategy if possible.

Results

We identified 1074 patients who were treated with immunotherapy between 2012 and 2018 across two tertiary centres, of whom 134 developed colitis. Being a retrospective analysis, not all data points were available for all patients, and the extent of missing data (where present) is detailed in each graph.

Incidence of CC colitis

The risk of developing CC colitis, as can be seen in **Fig 1.1A**, was significantly higher in those individuals who were treated with a combination of anti-PD1 and anti-CTLA4 (combination) therapy as opposed to anti-PD1 therapy or anti-CTLA4 therapy alone (monotherapy). Moreover, CC colitis occurred significantly earlier in patients on combination therapy (**Fig 1.1B**). Other factors such as patient gender (**Fig 1.1C**) or age of initiating immunotherapy (**Fig 1.1D**) were not significant. Smoking status was only of importance for patients on anti-PD1 therapy (**Fig 1.1E, Fig 1.1F**), where those who had never smoked appeared to be at significantly higher risk of developing colitis (**Fig 1.1E**). Prior IBD did not clearly increase the risk of developing colitis (data not shown), but as this only accounted for 8 patients, the study was under-powered to conclude this definitively. Other series have shown a moderate risk, although numbers were still small, particularly in the context of colitis[118]

Clinical course of CC colitis – prior clinical practice

Existing standard of care

A key component of assessing CC colitis is determining its severity – historically it has been utilized to escalate treatment of the colitis and determine whether or not to restart immunotherapy.

Fig 1.1 - Incidence of CC colitis

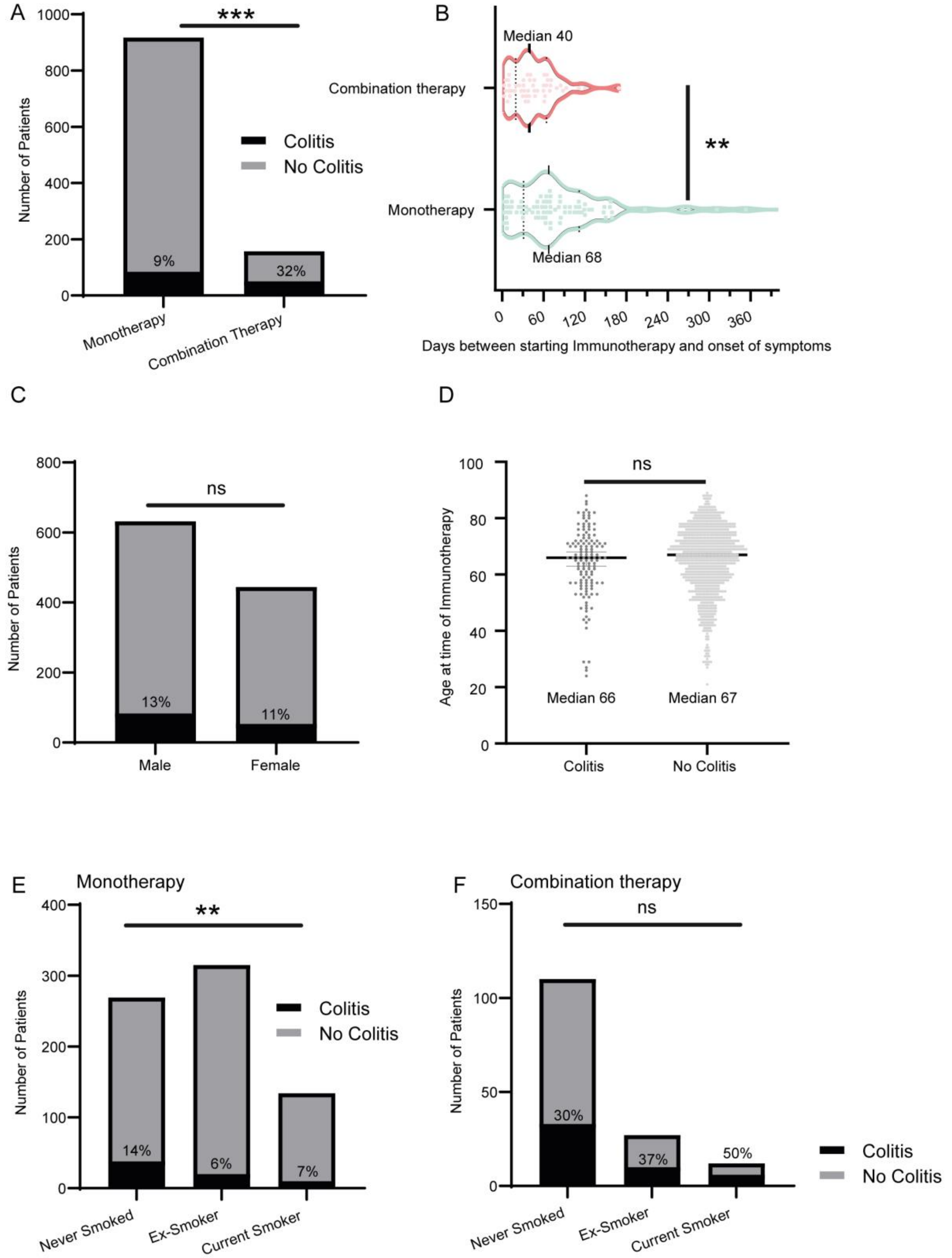


Figure 1.1 Incidence of CC Colitis. (A) Incidence of colitis by immunotherapy regimen. (B) Interval in days between initiation of immunotherapy and development of colitis in patients, split by therapy. (C) Odds of developing colitis by gender. (D) Age distribution of patients given immunotherapy developing colitis versus not developing colitis. (E) All patients on anti-PD1 monotherapy - odds of developing colitis by smoking status. Data not available for 22% of patients (F) All patients on combination anti-PD1 and anti-CTLA4 therapy - odds of developing colitis by smoking status. Data not available for 5% of patients A,C: Fisher's exact T test B: Mann-Whitney U test, D: T test, E,F: Chi-squared test. p<0.05(*), p<0.01(**), p<0.001(***), p<0.0001(****). Monotherapy = anti-PD1 or anti-CTLA4 therapy alone, Combination therapy = simultaneous anti-PD1 and anti-CTLA4 therapy

The Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE)(Version 5.0, **Fig 1.2A**)[159] were developed and in use at the time we performed our study. Endoscopic and histopathological assessment were not part of the protocol; systemic markers such as C-reactive protein (CRP) were variably utilized to quantify the degree of colitis, in line with idiopathic IBD management.

The mainstay of treatment was immunosuppression through steroids. Patients were usually started on 1mg/kg of prednisolone (with clinician variation in induction with intravenous methylprednisolone), following which they were brought back to clinic for reassessment on a weaning course. Individuals with persistent or re-flaring symptoms were treated with re-escalation of doses, followed by repeated weans. Recalcitrant cases were escalated to receiving anti-TNFalpha therapy (infliximab), but the threshold for deciding when cases fell into this category was not standardized and dependent on the treating clinician. Immunotherapy was halted for patients with grade 2 or above colitis. If the colitis was severe, immunotherapy was either downgraded from combination therapy (with simultaneous anti-PD1 and anti-CTLA4 agents) to monotherapy (with anti-PD1 therapy alone) or not restarted at all.

We observed that a significant number of patients (78%) were treated with steroid monotherapy (**Fig 1.2B**) despite many of these patients requiring prolonged courses (upto 219 days; median length 50 days) (**Fig 1.2C**). Given that the standard steroid treatment duration for an inflammatory bowel disease flare is 60 days, this was our first indication that the existing standard of care could possibly be improved.

Existing markers of colitis severity – CTCAE grading

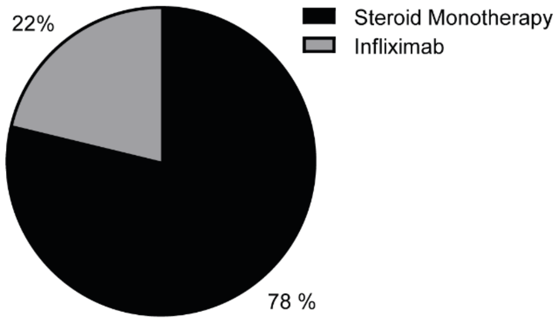
As per the CTCAE grading of colitis (**Fig 1.2A**), the distribution of severity of colitis could be assessed in this cohort of patients and is represented in **Fig 1.2D**. Grade 1 colitis accounted for 28% of cases, with the remainder (Grades 2, 3 and 4) accounting for the majority (72%) of CC colitis cases.

Fig 1.2 - Characterizing Severity of Colitis

A

Grade	Definition
Grade 1	Asymptomatic
Grade 2	Abdominal pain, mucus or blood in stool
Grade 3	Severe abdominal pain, peritonism
Grade 4	Life threatening consequences
Grade 5	Death

B



C

D

E

F

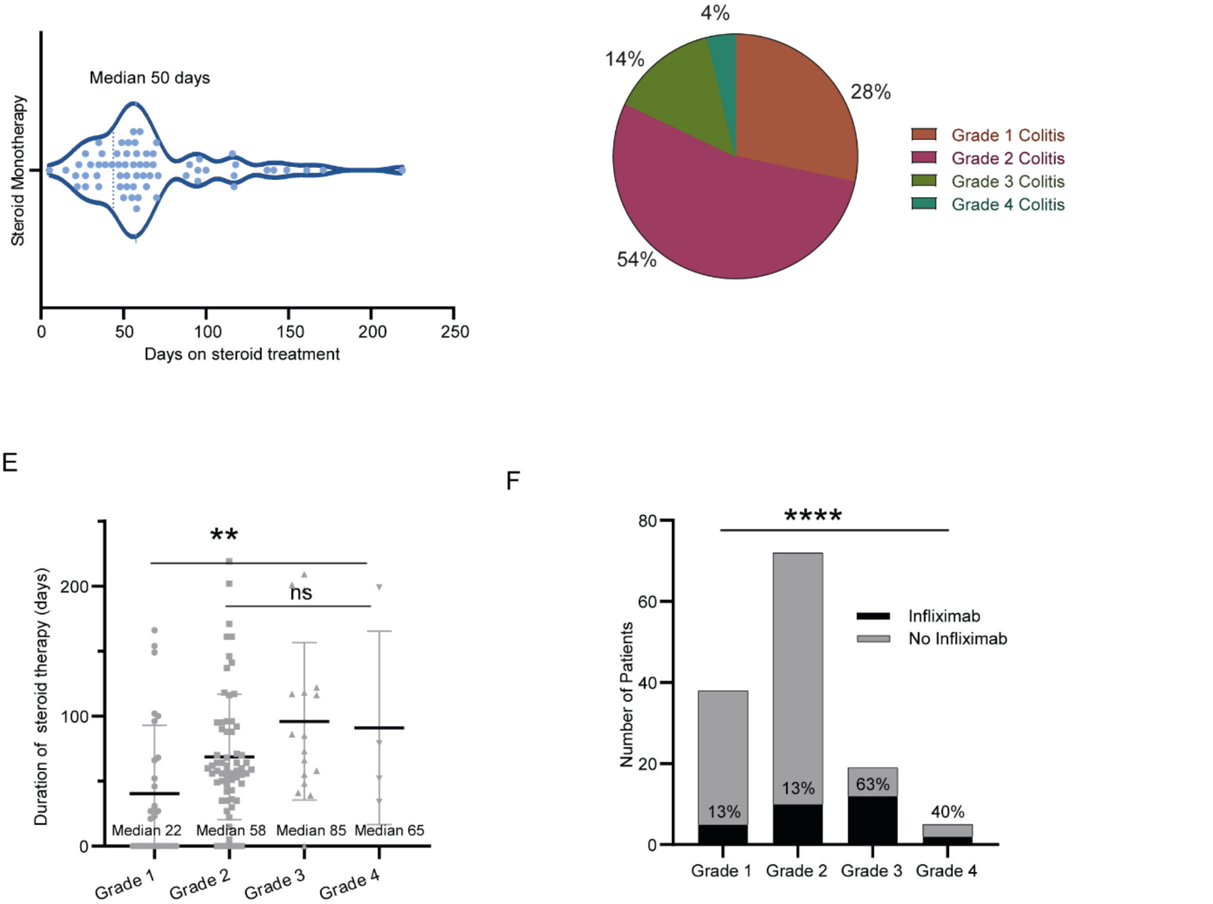


Figure 1.2 (A) CTCAE (version 5.0) criteria for defining severity of CC colitis. (B) Treatment of patients who developed CC colitis. Data unavailale for 2 (1.5%) patients. (C) Duration of steroid therapy in days in patients with CC colitis treated with steroids alone. Data not available for 20 patients (19%) (D) Proportional distribution of severity of CC colitis in this cohort of patients. (E) Duration of steroid therapy in patients with CC colitis of different CTCAE grades. Data not available for 20 patients (15% of total) F: Proportion of patients with different CTCAE grades of CC colitis requiring infliximab for resolution of colitis. E: Kruskal-Wallis one-way ANOVA test, F: Chi-squared test for trend. $p<0.05$(*), $p<0.01$(**), $p<0.001$(***), $p<0.0001$(****). Monotherapy = anti-PD1 therapy alone, Combination therapy = simultaneous anti-PD1 and anti-CTLA4 therapy

Comparison of treatment outcomes with existing CTCAE metrics of severity

Steroid treatment duration correlated with grade of colitis as assessed by CTCAE criteria, however, this was predominantly driven by the difference in steroid duration between Grade 1 colitis and the remainder. Grades 2-4 of colitis could not be distinguished by the duration of treatment required (**Fig 1.2E**); no Grade 5 colitis was seen in our cohort.

If we assessed the requirement for infliximab, again we observed a correlation with treatment requirement (**Fig 1.2F**), with more patients with severe CTCAE grades of colitis requiring infliximab for resolution. Being a retrospective study, this is open to interpretation.

Novel markers of colitis severity – treatment based assessment

Given the morbidity caused by prolonged treatment with high doses of steroids in this cohort (including but not exclusively osteoporotic crush fractures, proximal myopathy, diabetes), coupled with cyclical and extended courses of steroids being in use for the treatment of colitis, we decided to utilize the length of steroid course as a proxy marker for disease severity. We reasoned that milder inflammation would require a shorter steroid course. Similarly, patients who did not respond to steroids and required infliximab for resolution were likely to have more severe disease. We therefore decided to compare these outcome markers of severity (length of steroid treatment and infliximab use required for resolution of symptoms) to attempt to risk stratify patients. Only three patients in the cohort underwent a colectomy, so this could not be utilized as an outcome measure.

Given that CTCAE grading of CC colitis did not distinguish between the majority (72%) of patients in terms of duration of steroid treatment required, we wanted to assess if there were other clinically measurable factors that might better differentiate between different severities of CC colitis.

Given all patients started on steroids received a standardized course, we divided patients into 3 categories – those requiring a single ‘standard’ weaning course of steroids for resolution of colitis (steroids < 60days), those requiring repeated or prolonged steroids (steroids > 60days) and those non-

Fig 1.3 - Identifying patients at risk of more severe colitis

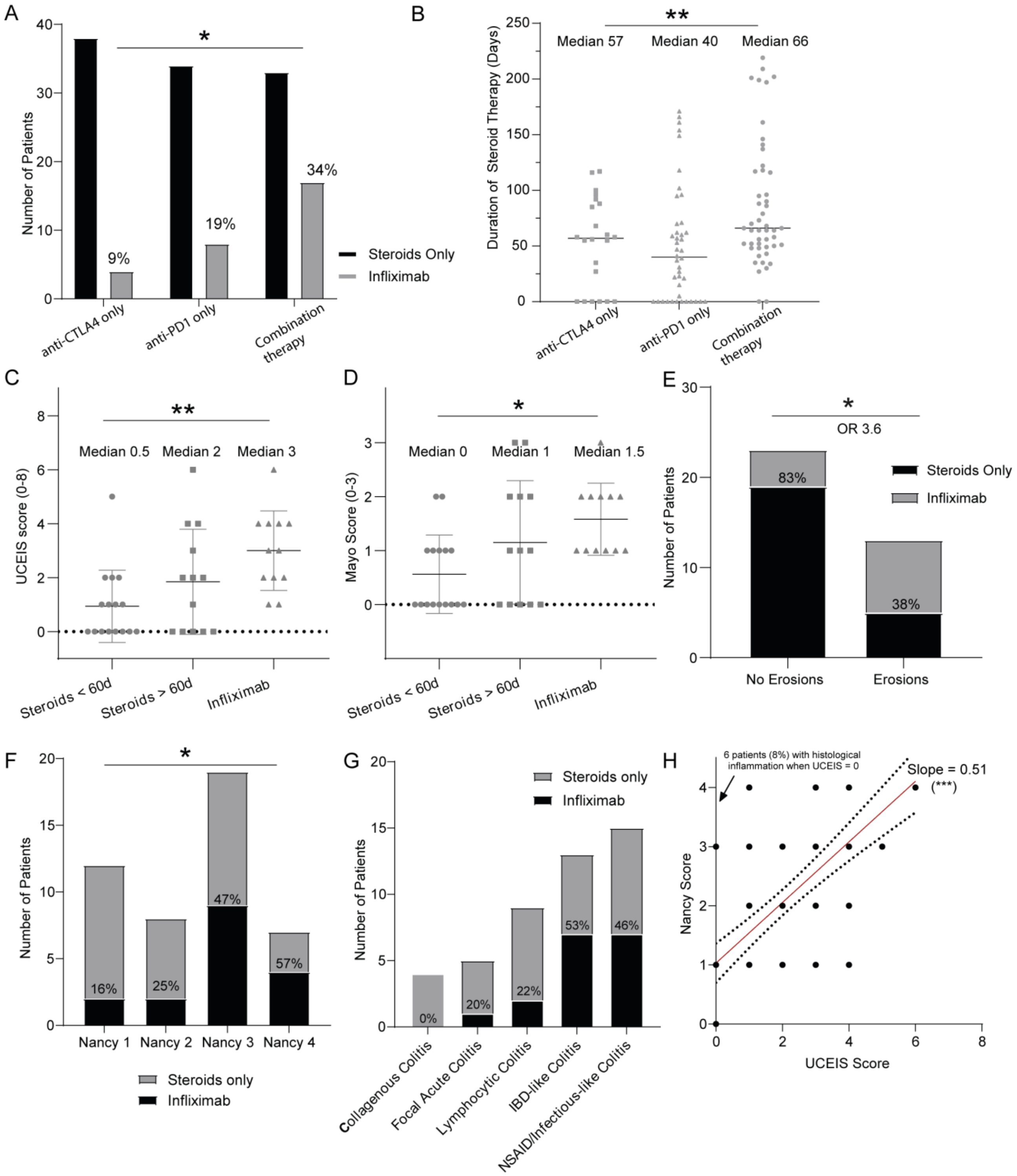


Figure 1.3 (A) Treatment required for resolution by immunotherapy regimen (B) Duration of steroid monotherapy by immunotherapy regimen (C) Correlation between UCEIS endoscopy score and treatment outcome. UCEIS score 0-8, 0: no inflammation, 8, most severe inflammation. Data available for 41 patients (D) Correlation between Mayo endoscopy score and treatment outcome. Mayo 0: no inflammation, Mayo 3: Most severe inflammation Data available for 41 patients. (E) The presence of erosions at endoscopy was predictive of the need for Infliximab. F: Increasing histopathological severity also correlated with outcome. Nancy 0: no inflammation, Nancy 4: most severe inflammation. Data available for 46 patients. G: Correlation of type of histopathology with outcome. Insufficient numbers for statistical analysis. H: Correlation of histopathology with endoscopic appearances. 80 endoscopic-histopathology pairings A:Chi-squared test. B,C,D: ANOVA (Welch) E: Fisher's exact test. F:Chi-squared test for trend. H: Simple linear regression. p<0.05(*), p<0.01(**), p<0.001(***), p<0.0001(****). Combination therapy = simultaneous anti-PD1 and anti-CTLA4 therapy

responsive to steroid treatment (requiring infliximab for resolution). We then proceeded to assess different clinical factors against this tripartite standard of colitis severity.

Identifying patients at risk of more severe CC colitis

Immunotherapy regimen: Combination therapy was not only more likely to require infliximab for resolution (**Fig 1.3A**), but also required significantly longer steroid treatment than either anti-PD1 or anti-CTLA4 therapy alone (**Fig 1.3B**)

Endoscopy: A subset of patients within the study underwent endoscopy during their flare of colitis (as this was not part of the treatment protocol). Post-hoc blinded scoring (**Methods**) was performed as part of the analysis and revealed a strong correlation with outcome. Patients requiring longer courses of steroids or infliximab had more severe endoscopic appearances (Mayo >1; UCEIS ≥3; **Fig 1.3 C,D**). In particular, the presence of erosions was strongly correlated with the requirement for infliximab (Odds Ratio 3.6; **Fig 1.3E**).

Histopathology: Histopathological inflammation as assessed by the Nancy histological index (commonly used for UC assessment) was measured in a blinded post-hoc analysis (**Methods**). Increasing inflammation (as quantified by increasing Nancy scores) correlated with the requirement for infliximab (**Fig 1.3F**). Diverse patterns of histopathological inflammation were noted (**Fig 1.3G**) – namely collagenous colitis (9%), focal acute colitis (11%), lymphocytic colitis (20%), IBD-like colitis (28%) and infectious or NSAID-associated colitis (32%). Certain types of histopathology (IBD-like and NSAID/infectious-like) were associated with an increasing need for infliximab (**Fig 1.3G**), but there was an insufficient number of patients in each category for assessing statistical significance.

Endoscopy and histopathological inflammation were significantly correlated by linear regression analysis (**Fig 1.3H**), but notably, some patients with normal endoscopic appearances still had inflammation by histopathology.

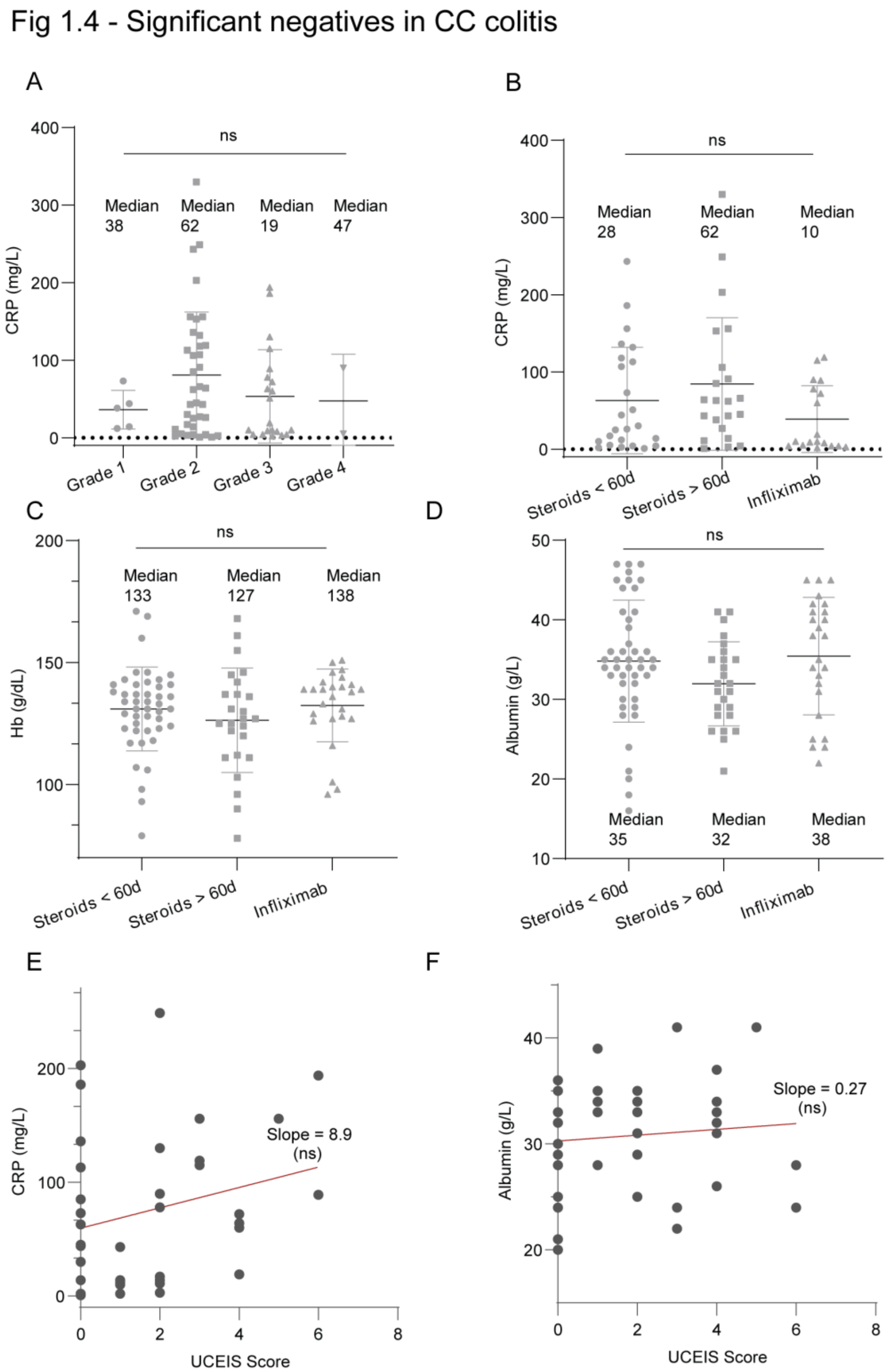


Figure 1.4 (A) No correlation between systemic C-reactive protein (CRP) and grade of CTCAE colitis. p = 0.24 (B) No correlation between systemic CRP and severity of colitis as assessed by treatment required for resolution. p = 0.12. (C) No correlation between systemic haemoglobin and severity of colitis as assessed by treatment required for resolution. p = 0.45 (D) No correlation between systemic albumin and severity of colitis as assessed by treatment required for resolution. p = 0.12. (E) No correlation between CRP and endosocopic inflammation as assessed by UCEIS score. Line - linear regression (E) No correlation between Albumin and endoscopic inflammation as assessed by UCEIS score. Line - linear regression A-D: ANOVA (Welch) E,F: Linear regression. p<0.05(*), p<0.01(**), p<0.001(***), p<0.0001(****), ns = not significant.

Significant negatives

Systemic markers of inflammation: In UC, systemic markers of inflammation such as increased C-reactive protein (CRP) correlate with more severe disease and can be used to prognosticate outcome. However, in CC colitis, we found no such link either with the severity of colitis (**Fig 1.4A**) or with treatment requirement (**Fig 1.4B**). Low Haemoglobin and Albumin are also considered systemic markers of severity in UC, and both of these were also not correlated with treatment requirement (**Fig 1.4C, 1.4D** respectively). Finally, we also assessed whether endoscopic severity of inflammation (in the subset of 80 cases where this information was available) correlated with systemic markers of severity, and this too, was not found to be the case (CRP – **Fig1.4E** or Albumin – **Fig 1.4F**)

Treatment of CC colitis

As previously mentioned, a significant proportion of patients received prolonged steroid monotherapy for treatment of CC colitis, with significant consequent morbidity. We wished to ascertain whether there was any evidence for managing this differently.

Intravenous Methylprednisolone: Clinicians occasionally gave high dose intravenous methylprednisolone as induction prior to starting a course of oral steroids (without a clear pattern). We assessed whether this reduced the overall duration of steroids required for treatment colitis. In the context of steroid only treatment, induction with IV methylprednisolone had no effect on overall treatment duration (**Fig 1.5A**). There was possibly an effect if patients went on to receive infliximab (**Fig 1.5B**), however there were insufficient numbers (only one patient who received infliximab did not receive methylprednisolone) to make a definitive conclusion.

Anti-TNF alpha therapy (Infliximab)

Infliximab was given at a median interval of 33 days after initiation of steroid therapy, i.e. approximately 2 weeks before the end of the first course (**Fig 1.5C**). Although this is appropriate for

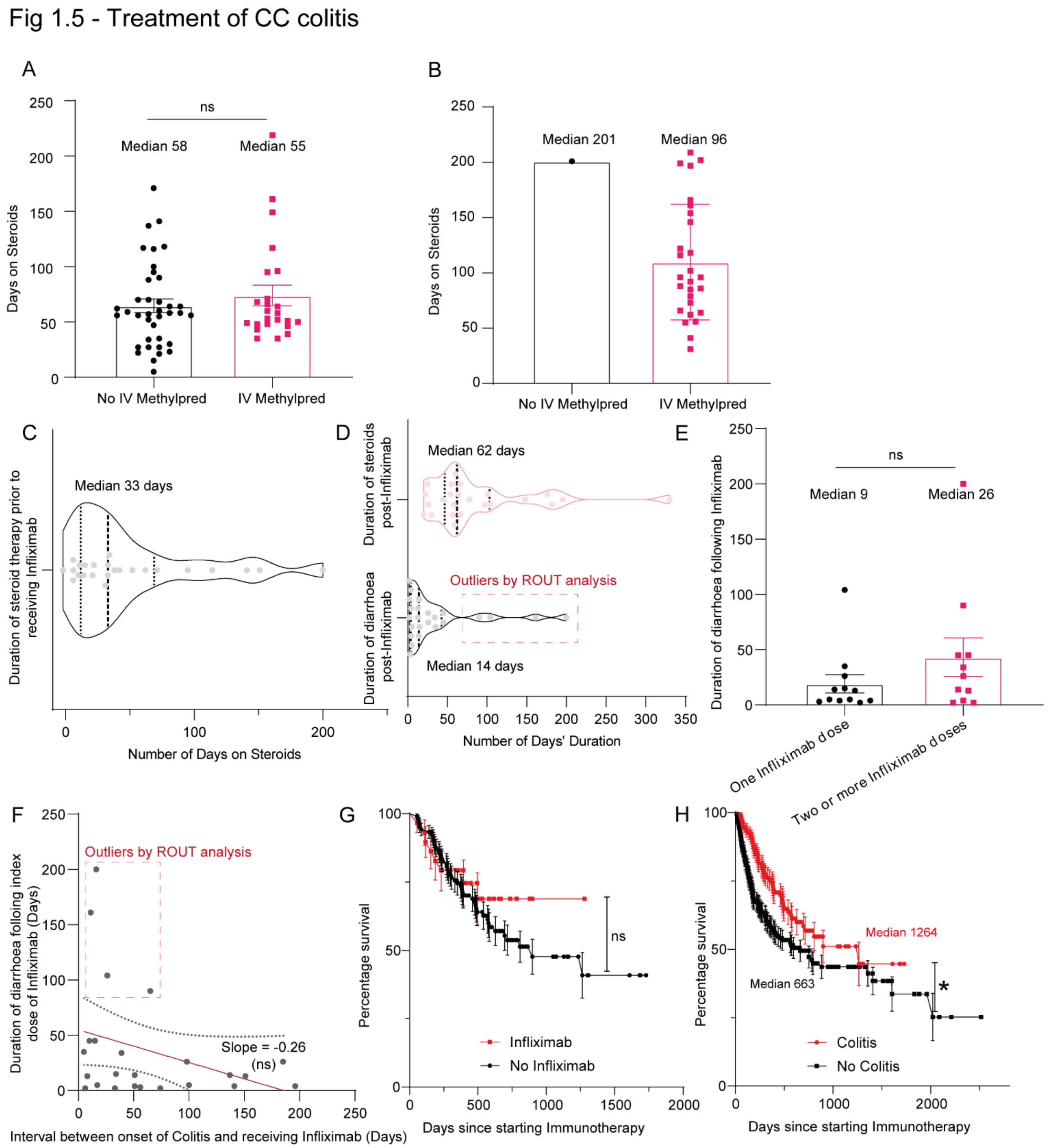


Figure 1.5 . (A) CC colitis treated with steroids only - duration of treatment by whether or not induction methylprednisolone given. p = 0.4 (B) CC colitis treated with infliximab - duration of treatment by whether or not induction methylprednisolone given. Insufficient numbers for statistical analysis. (C) Duration of steroid therapy prior to patients receiving Infliximab (D) Duration of steroid therapy and diarrhoea following patients receiving infliximab. ROUX analysis (threshold = 1% utilized to identify outliers) - dashed box. (E) Duration of diarrhoea following single or multiple doses of infliximab. p = 0.32 (F) No correlation between a delay in receiving Infliximab and more resistant disease. Slope = linear regression, p = 0.26 (G) Survival of patients receiving immunosuppression with Infliximab non-inferior to those patients treated with steroids alone. p = 0.40 (H) Survival of patients who develop colitis is better as compared to those patients who do not, p = 0.01. Data available for 326 patients. A: T test (Welch's correction) E: T test (Mann-Whitney U). E: F:Linear regression (least squares). ROUT analysis for outliers (threshold Q=1%). G,H: Log Rank (Mantel-Cox) test p<0.05(*), p<0.01(**), p<0.001(***), p<0.0001(****), ns = not significant. IV = intravenous. Methylpred = Methylprednisolone.

non-responsive disease, there were a significant number of patients for whom the interval on steroid monotherapy was much longer (a maximum 209 days).

Once Infliximab was given, the median duration of steroid therapy was a single weaning course of steroids (**Fig 1.5D**, **upper violin plot**), and the diarrhoea resolved rapidly (median duration 14 days, **Fig 1.5D, lower violin plot**). This suggested that it was a successful treatment, however there was a 'tail' of 4 patients (14%) who were outliers and had diarrhoea for longer than 50 days after infliximab therapy (suggestive of resistant disease).

In three of these four patients, immunosuppression in addition to steroids and infliximab was used and improved symptoms (vedolizumab in two and mycophenolate mofetil in one); all were melanoma patients who had combination therapy.

This effect did not appear to be dose-dependent (**Fig 1.5E**) where multiple doses of infliximab did not have an additional effect in shortening the duration of diarrhoea (as compared to a single dose). In addition, a delay to receiving infliximab was not correlated with reduced infliximab efficacy (**Fig 1.5F**).

Survival Analysis

Finally, it was also possible to assess the effect of treatment and development of colitis on survival. Immunosuppression of patients who developed colitis with Infliximab did not adversely affect their survival (**Fig 1.5G**).

More broadly, as development of colitis often necessitates halting of immunotherapy as well as immunosuppression with steroids (and/or Infliximab/Vedolizumab/Mycophenolate), we decided to assess the effect of developing colitis on patient survival overall. Despite the caveats mentioned above, as well as morbidity from steroid treatment and from colitis itself, patients who developed colitis had an improved survival compared to those who did not (**Fig 1.5H**).

Discussion

The goal of this analysis of 1074 patients given immunotherapy was broadly, to identify risk factors for development of CC colitis, determine factors that would aid in prognostication as well as an optimum management strategy. When combined, this would also ideally yield insights into the mechanisms behind CC colitis.

Identifying patients at risk of colitis : Combination therapy with anti-PD1 and anti-CTLA4 agents was the most significant risk factor for developing colitis, which also occurred at a significantly earlier interval than monotherapy-induced CC colitis. This would either suggest a different underlying process driving the two colitides or a cumulative effect of pathway blockade. Smoking is protective in UC, through a combination of immunosuppressive and microbiome effects[222]. Similarly, smokers and ex-smokers appeared to be at a slightly lower risk of developing CC colitis, but this effect was only observed for monotherapy. A potential explanation could be that in combination therapy, pro-inflammatory drivers overwhelm this immunosuppression. However, there are several potential caveats with our conclusions and these results should not be over-interpreted at present. Firstly, smoking status was not recorded for 22% of patients overall. Secondly, smoking habits were more likely to have been recorded for patients with lung cancer given the natural clinical associations with the disease than for patients with renal cancer or melanoma receiving immunotherapy. Thirdly, this was self-reported data and not anonymized at the point of collection, and patients can conceivably be seen to be less likely to volunteer accurate information on smoking when receiving treatment for cancer (social desirability bias). Ideally, a prospective anonymized analysis that focused on this question would improve our confidence in the results. Gender, age and prior IBD all not affecting the risk of developing CC colitis demonstrates it doesn't follow a conventional auto-inflammatory pattern. There may however be a correlation of checkpoint colitis with intra-tumour lymphocyte activation, given that patients who develop colitis have a survival advantage despite requiring cessation of immunotherapy, immunosuppression,colitis and steroid-related morbidity.

Prognostication of CC colitis : Given that CC colitis is treated with repeated courses of high dose steroids (until resolution of diarrhoea or escalation to anti-TNF alpha therapy), we utilized treatment duration and escalation as a marker of severity. Existing CTCAE metrics of determining CC colitis severity are subjective and do not distinguish between the majority of patients (72%) who have symptomatic (Grade 2-4) colitis. Although CTCAE colitis grading is correlated with the requirement for Infliximab, this was likely because immunotherapy toxicity guidelines[152,212,213] and local treatment protocols advise escalation to infliximab for grade 3/4 colitis if symptoms persist on steroid treatment. Reassuringly, we observed that grade 3 & 4 colitis patients were more likely to be treated with infliximab.

Utilizing treatment as a proxy for colitis severity Although more discriminatory than assessing the frequency of diarrhoea, which can have multiple causes, this method of assessment is still subjective and prone to bias. As the incidence of CC colitis is likely to increase, and comparative analyses will be required for assessing treatment efficacy, the field could benefit from a more objective assessment measure of severity, or certainly from standards for repeat endoscopy or faecal calprotectin assessment to assess healing. We currently also have not captured patients' experience of this disease, how it or the treatments employed affected their quality of life (QOL). Anecdotally, patients report that immunotherapy associated diarrhoea has a significant impact on their quality of life through associated insomnia, fatigue and effects on mood. Therefore, as we have done for IBD, determining and assessing key core outcome sets[223] is vital for improving patient care and addressing gaps in our understanding of the disease.

Biomarkers of severity Systemic markers such as CRP, Albumin and Haemoglobin did not correlate with outcome, but tissue inflammation as assessed endoscopically or histologically did. This is in contrast to UC, where an assessment of systemic inflammation is an integral part of prognosticating in acute severe colitis with Truelove-Witts and Travis criteria. This has led to a change in practice whereby endoscopy and liaison with gastroenterology is now recommended for prognostication in patients presenting with CC colitis[224]. Faecal calprotectin was not measured routinely in these patients as it was historically not part of the assessment framework proposed by the CTCAE guidelines, or usually managed by gastroenterologists. We did not therefore have enough datapoints to draw any conclusions. Subsequent analysis published by other groups, however, suggests that it is representative of disease severity (endoscopy was used as a gold standard)[225] and could be utilized as part of a holistic assessment.

Management of CC colitis : Our analysis demonstrated that a single course of steroid monotherapy is only sufficient for managing 42% of patients (N.B. for 15%, steroid duration was not available). For the remainder, Infliximab is successful rescue therapy, leading to rapid resolution of diarrhoea within a median of 14 days. When coupled with a weaning steroid course, 86% of patients demonstrated a resolution of diarrhoea within 50 days. If patients did not respond to a single dose of Infliximab, from our data, control was only achieved with the addition of other agents such as Vedolizumab, Mycophenolate or Tofacitinib (rather than higher doses of Infliximab). Immunosuppression with 1-3 doses of Infliximab did not adversely affect cancer control with an impact on mortality at a population level.

Insights into mechanism of CC colitis: Taken together, these results were suggestive that monotherapy and combination therapy-induced CC might be mediated through distinct pathways. Tissue-localised (as opposed to systemic) inflammatory populations may be primary drivers of CC colitis, and TNF alpha is likely to be elevated and playing a key role in perpetuating inflammation. However, the stochastic time to onset of colitis after initiation of a patient on immunotherapy, along with considerable variability in histopathological patterns suggests that there may be different initiating events (e.g. a response to self-antigen versus a response to a change in microbiome etc.) even if the pathway downstream of inflammation is a common one.

In order to try and explore these factors in more detail with an aim to mechanistic understanding and improved therapeusis, we proceeded to more in-depth characterization of ulcerative and checkpoint inhibitor-induced colitis (**Chapters 2** and **3**).

Chapter 2

Exploring CD8+ T cell behaviour in idiopathic ulcerative colitis

Introduction

Ulcerative colitis (UC) is a multisystem inflammatory disorder, thought to be driven by mucosal inflammation, but with well described associations with dermatological, ophthalmic and arthropathic manifestations. Immune-mediated tissue damage is thought to be initiated and perpetuated by a combination of immune dysfunction, barrier failure, and dysbiosis, all on a background of genetic predisposition[226–228].

Multiple approaches have been utilized to understand the function of different facets of the complex interactions between epithelium, stroma, microbiome, resident immune system and circulating cells that form the basis of homeostasis in healthy colon. Historical approaches such as multi-parameter flow cytometry and microarray based transcriptomic analysis[229,230] were instrumental in understanding which cell populations and pathways were altered in inflammation, but were constrained by virtue of being limited in the number of parameters that could be measured and biased by *a priori* thought. Multiple animal models of colitis[231] were key in clarifying the functional effects of different pathways and cellular populations, but none encapsulate the native disease process in its entirely, highlighted by the translational failure of certain therapeutic approaches[45,232].

This changed with the development of Next-Generation Sequencing (NGS) technologies coupled with technical advances such as droplet-based single-cell transcriptomics (sc-RNAseq). Albeit RNA-based, it transformed our understanding of cellular behaviour, offering quantitative unbiased transcriptional profiling of a wide variety of cell types. When leveraged with bioinformatic analysis that reduced the resultant high dimension datasets into clusters that defined cellular states based on their transcriptomic similarity rather than pre-conceived understanding, it revealed new cell types, novel cellular function, cellular development and interactions in diseases that had been thought well characterized[19,20,24,25,29,233,234].

This technology was successfully put to use by many groups, comparing health, UC inflamed and non-inflamed states in order to describe stromal, epithelial and immune cell behaviour, and start to impute mechanisms and function.

In prior analyses, CD8 T cells in the peripheral blood and mucosa had been recognised to be prognostically important in UC (and Crohns'), suggesting the population could be performing a key mechanistic role[57]. There was however, no in-depth and unbiased characterization of their heterogeneity, transcriptional regulation, and effector function

Aims

Utilize novel scRNA-sequencing paired with single-cell TCR information to unbiasedly interrogate the prognostically relevant population of CD8 T cells in colitis:

1) Assess different phenotypes of CD8 T cells in inflamed UC (UC) and healthy controls (HC), identifying disease specific populations, creating an atlas data resource.
2) Utilizing single-cell T cell receptor (TCR) information, infer T cell relationships and the dynamic changes between phenoytpes in inflammation
3) Identify the potential targets of pathogenic populations and attempt to understand the function of these populations

Results

Defining CD8 T cell populations in colonic tissue

Using droplet-based scRNA-seq, we first evaluated sorted colonic CD8+ T cells from three healthy volunteers and three UC patients (**Methods**, **Fig. 2.1A,B**). We combined gene expression data from 8,581 cells that we recovered for clustering analyses after quality control (**Methods**). This revealed 14 CD8+ cell populations (**Fig. 2.1C**), which were annotated as previously described subtypes of CD8 T cells dependent on cluster-specific specific gene expression (**Supplementary Table 2.1**). We identified naive, memory, tissue-resident memory (TRM), effector, and double-positive (DP) CD8+CD4+ cells, as well as populations with innate-like features such as mucosal-associated invariant T cells (MAITs) and intraepithelial lymphocytes (IELs). We also identified a novel IL26+ CD8+ cell population.

We were able to establish that these populations were recovered across multiple donors rather than being patient specific (**Fig 2.1D**), but there were clear differences between the clusters present in health and disease (**Fig 2.1E**).

In order to validate transcriptional signatures and add in cell-surface protein information, we repeated and extended this analysis by performing scRNA-seq combined with antibody-tagged cell-surface protein quantification (CITE-seq, **Methods**). Oligonucleotide-tagged antibodies were added to T cells, which were amplified along with the cell-specific unique molecular identifier (UMI) tag, giving a second metric by which to cluster cells (complete antibody list in **Materials**).

We performed this for a larger validation cohort of 12 patients (n = 7 UC, 5 HC), where we recovered 9,602 additional cells over 4 hashed runs (**Methods, Fig 2.1F, Fig Supp 2.1A**).

In order to ensure confidence in our conclusions, for both the index and validation cohorts, patients' samples were only selected for single-cell analysis if they were on minimal immunosuppressive therapy (typically oral or topical aminosalicylates alone) to control their ulcerative colitis at the time of collection. In particular, we avoided patients on steroids or biological therapy (80% biologic naïve

Fig 2.1 - CD8 T cells in health and inflamed ulcerative colitis

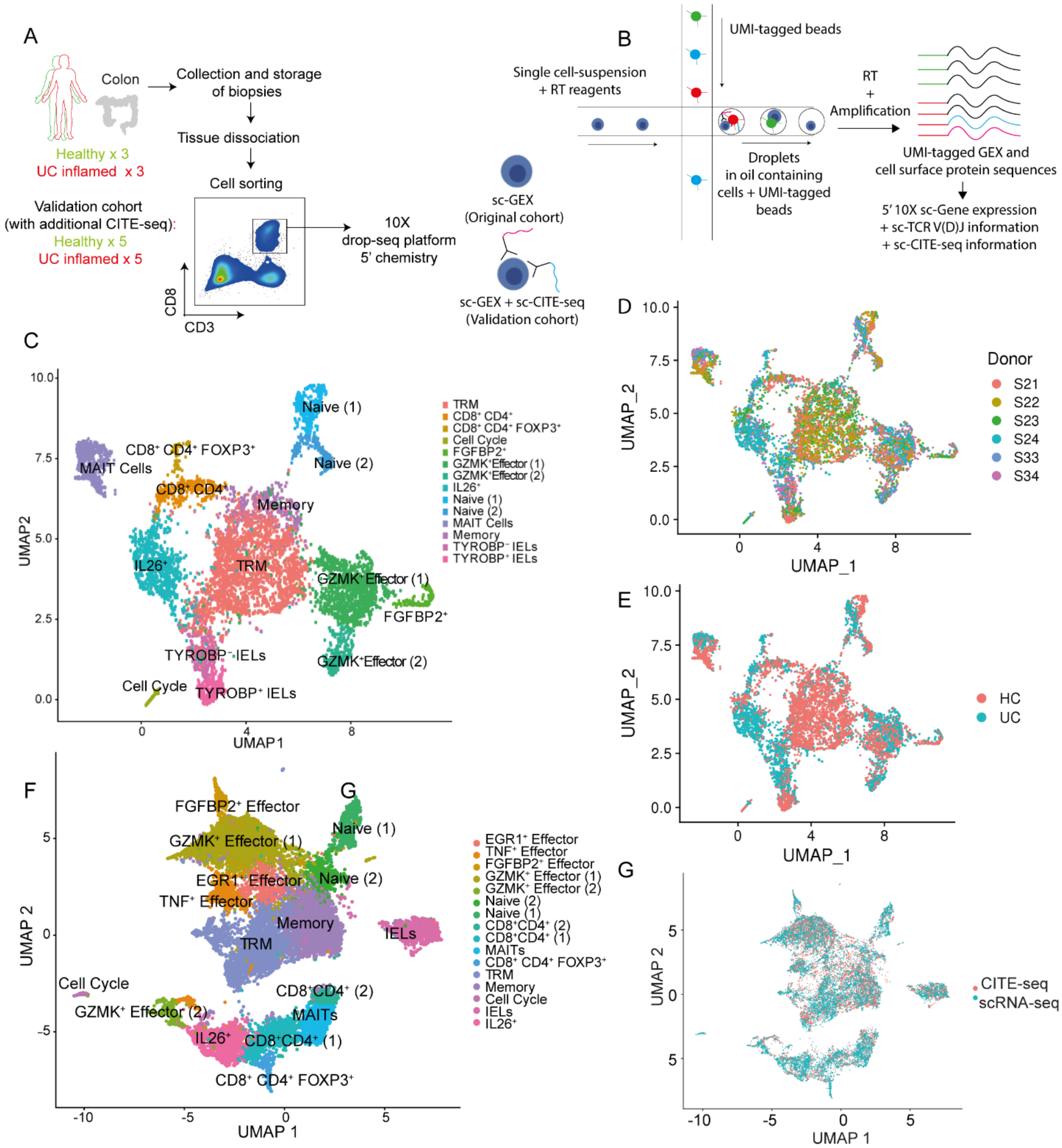


Figure 2.1 . (A) Experimental setup - initial and validation cohorts with additional cell surface protein (CITE-seq) information in validation cohort. (B) Schematic of 10X drop-seq platform, demonstrating how microfluidics are utlized to create single cell-bead droplets that then execute inta-droplet cell lysis and reverse transcription to get a transcriptional snapshot of each cell. sc-TCR information is derived from sc-gene expression using amplifying primers that target the TCR sequence. Protein information is derived from the oligo tags attached to antibodies (C) UMAP embedding of initial cohort demonstrating 14 CD8 T cell subtypes in the colon across health and inflamed UC. (D) UMAP embedding of initial cohort by sample (3 x HC and 3 x UC). (E) UMAP embedding of initial cohort by sample type: healthy control (HC) or inflamed ulcerative colitis (UC). (F) UMAP embedding of validation cohort with additional CITE-seq information demonstrating 16 CD8 T cell subtypes - some additional sub-clusters were delineated for CD8 effector cells (G) Overlay of initial (scRNA-seq) and validation (CITE-seq) cohort-CD8 T cells demonstrating consistency in all cell types recovered, including the double positive, IL26 and FOXP3+ CD8 subsets

in this cohort), whilst also ensuring a representative sample of mild to severe disease severity (Mean UCEIS 4.5; all patient characteristics summarized in **Appendix D**).

Combining CITE-seq with scRNAseq allowed us to reproduce our original findings with slightly different clustering. Interestingly, by increasing our total cell pool, we were able to define two additional sub-clusters of effector cells : *EGR1+* and *TNF+* effector cells (detailed in **Supplementary Table 2.1 , Fig 2.1F, Fig Supp 2.1A, C**), which had clustered together with the GZMK+ cluster in our initial cohort (**Fig Supp 2.1E**). It allowed us assess the expression of specific proteins with phenotypic relevance in addition to those that are poorly expressed at the transcript level, such as exhaustion markers (e.g. and PD-1) and memory markers (CD45RO) (**Fig Supp 2.1F**). We were also able to cluster cells by cell-surface antibody expression (**Fig Supp 2.1B**) and superimpose gene expression (GEX) clusters (**Fig Supp 2.1C**), confirming the existence of the novel IL26+ population, as well as the CD4 CD8 double positive and FOXP3-positive CD4+ CD8+ T cells. We were able to confirm that the novel IL26+ cluster was CD4-, whilst bearing residency markers (CD103) and markers of activation and antigen exposure (CD69, PD1) (**Fig Supp 2.1D**). In summary, the overlap of this second cohort was sufficient to validate our original findings (**Fig 2.1G**), as well as highlighting sub-clusters not clearly evident in the first iteration.

Disease specific CD8 T cell population changes in UC

We first interrogated our dataset for disease specific changes in UC.

As stated earlier, we saw significant alterations in UC (**Fig 2.2A**), where we present data from the original (scRNA-seq) and validation cohort (CITE-seq) side by side. The novel population of IL26+ cells, consistently increased in UC in both the original and validation datasets, comprising 18% (up to 29%) of all CD8+ cells recovered in UC.

As could be expected, GZMK effector cells were increased in inflammation, comprising a mean of 25% of CD8 T cells in UC. In line with this, we saw an increase in MAIT, Cycling populations, with reciprocal proportional decreases in memory cells and TRM-like T-cells (which made up about 45% of all recovered cells on average in healthy individuals but only about 10% of CD8+ cells in UC).

Fig 2.2 - CD8 T cell dynamics in inflammation and health

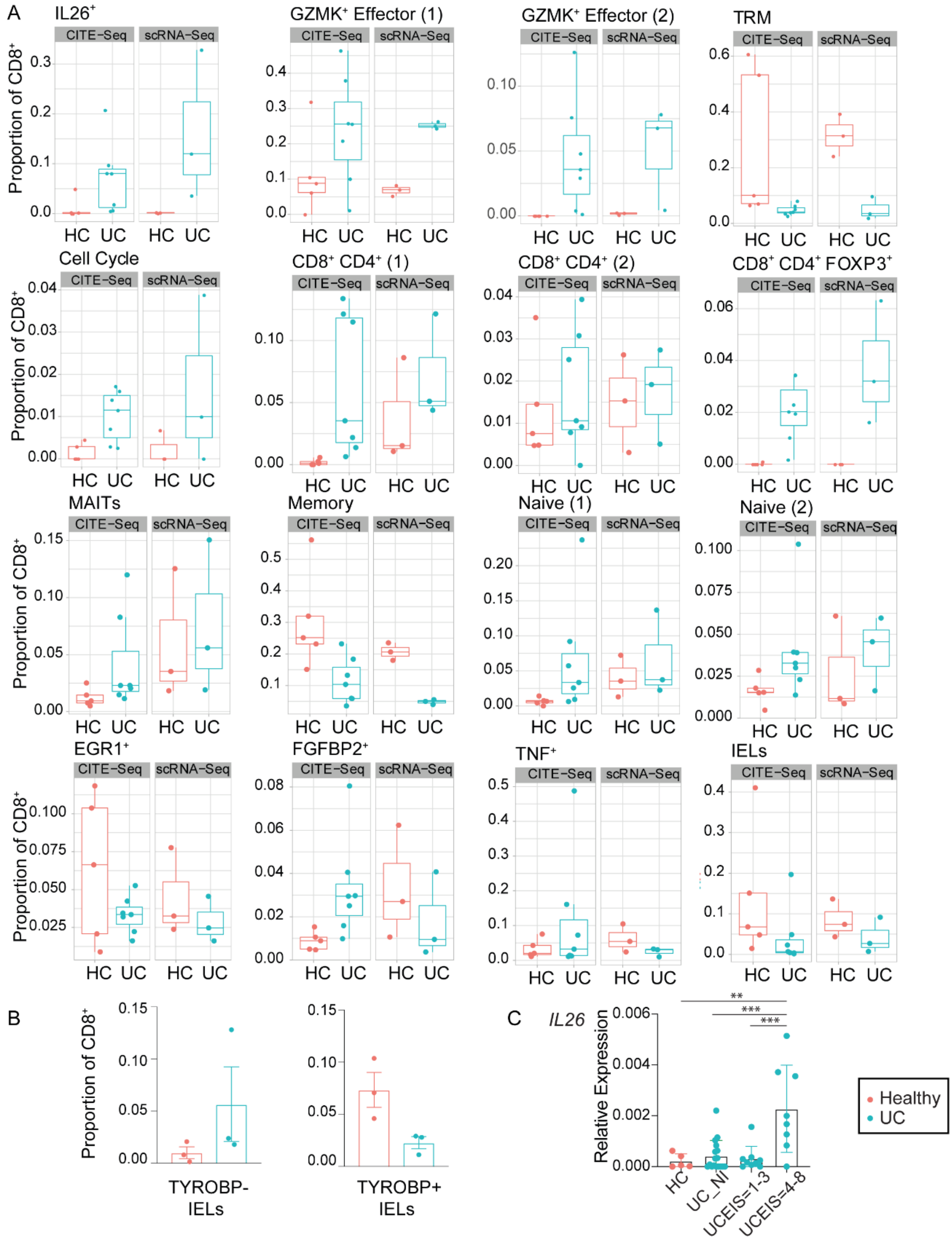


Figure 2.2 (A) Proportion of all 16 clusters of CD8 T cells in tissues across health (HC) and inflamed ulcerative colitis (UC). Data presented for both initial (scRNA-seq) and validation (CITE-seq) cohorts separately. (B) Intraepithelial lymphocytes (IELs) proportions by both induced (TYROBP-) and natural (TYROBP+) phenotypes. (C) qPCR quantification of *IL26* expression in the colon of HC (n=5) and UC (n=17) by degree of biopsy inflammation. UC_NI: Non-inflamed UC; UCEIS 1-3: Mild; UCEIS 4-8: Severe. ANOVA (Tukey's multiple comparison), Mean + SEM shown. **p=0.0018, ***p=0.0002 (UC_NI vs UCEIS 4-8), ***p=0.0005 UCEIS1-3 vs UCEIS 4-8, housekeeping gene *GAPDH*, no line = non-significant result.

Although forming a very small proportion of the CD8 T cells overall, the CD4+CD8+ double positive and FOXP3+ CD8+ populations also increased consistently in inflammation. Finally, although the proportion of IELs as a whole remained unchanged (**Fig 2.2A**), this was an aggregate result, as we observed a decrease in naturally occurring IELs (TYROBP+) and an expansion of induced TYROBP- IELs in inflammation (**Fig 2.2B**).

The only slight discordant result was for the effector FGFBP2 population, where we saw that with increased n, we saw an increase in inflammation, whereas this had not been clear in the original smaller dataset.

We confirmed the increase in IL26 by performing RT-PCR of whole biopsy samples from patients with different degrees of inflammation as measured by UCEIS scoring (scale 0-8, high score = more severe inflammation) (**Fig 2.2C**). Moreover, we could see in this independent cohort that the degree of increase of *IL26* increased with the degree of inflammation.

Imputing function for CD8 T cell population – single gene analysis

We then performed a deep dive analysis of the transcriptomic signatures of all CD8 T cell populations in order to understand their function in more depth. Expression of key genes across all the recovered subsets is presented in **Fig 2.3A**.

For example, we see that the FGFBP2+ and GZMK+ effector cells, although both expressing cytotoxins (e.g. *GZMA*), can be delineated depending on the subtype of other granzymes, expressing *GZMH* and *GZMK* respectively. FGFBP2+ cells may play a greater role in antimicrobial activity given their higher expression of alarmin and lipid membrane damaging agent granulysin (*GNLY*)[235].

We can see that the double positive CD4+ CD8+ populations are both positive for permissive cytokine secretory receptor *IL1R1* as well as *TNFRSF4*[236], a gene linked with IBD, activated and memory T cell phenotype, and is associated with a pathogen (viral/bacterial) response. These cells also produce some *IL26* and express *KLRB1* both of which are associated with an activated Tc17 type phenotype[237].

Fig 2.3 - Imputing CD8 T cell subset function - single gene analysis

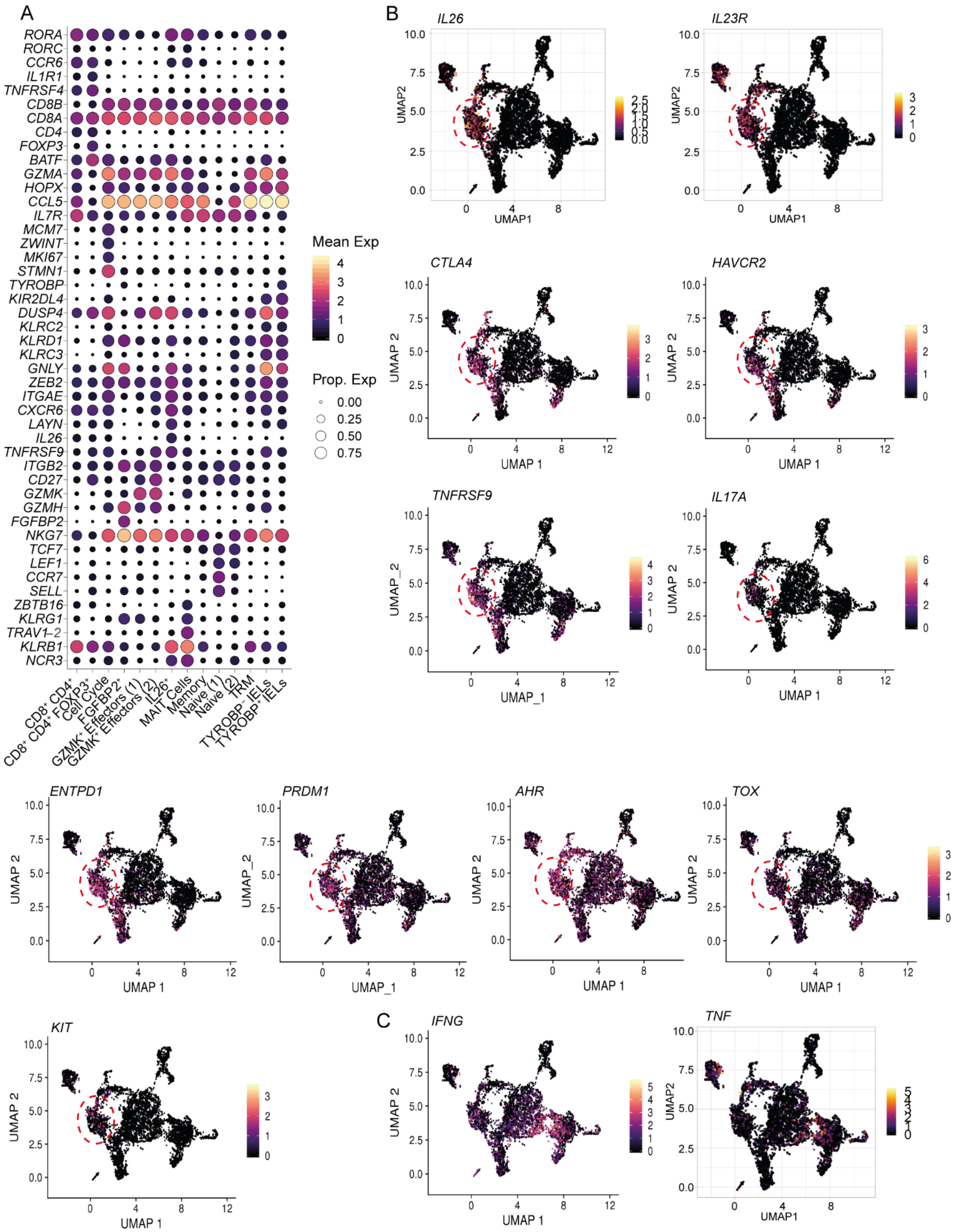


Figure 2.3 (A) Bubble plot of the expression of a selection of relevant genes across tissue CD8 T cell clusters detected in inflamed UC (UC_I) and health (HC) as determined by single-cell RNA sequencing. The percentage of cells within the cluster and relative expression of the gene are demonstrated (B) Multiple plots, each demonstrating the expression of a selection of relevant genes across the UMAP representation of all tissue CD8 T cells in UC_I and HC.The IL26+ cluster is highlighted with a dashed red line (C) Expression of *IFNG* and *TNF* across the UMAP representation of all tissue CD8 T cells in UC_I and HC.

IL26+ and MAIT cells both express the NK cell receptor NCR3, suggesting that there may be a degree of plasticity between adaptive and innate immune functions in these cells.

Looking in more detail at our novel UC-associated cluster, IL26+ cells (**Fig 2.3B**, red circle) could be identified by expression of cell surface *IL23R* (once MAIT cells had been excluded). They had some transcriptional characteristics in common with activated GZMK+ populations and induced IELs, including significant expression of co-inhibitory/exhaustion receptors such *CTLA4* and *HAVCR2* (TIM3), as well as high levels of *TNFRSF9* (4-1BB) that is known to drive pro-inflammatory interactions with antigen-presenting cells[236]. In line with previous indicators, these cells also upregulated *IL17A*.

Unusually, IL26+ cells also expressed a number of markers typically found in lineage-negative type 3 innate lymphoid cells, including transcription factors c-KIT and AHR.

Finally, we could also assess the expression of both *IFNG* and *TNF*, thought to be key effectors in the inflammation across these populations (**Fig 2.3C**). Although there was mid-level expression of both by the IL26+ population, the highest producers were split between resident TRM and effector GZMK+ cells, which in our validation cohort, formed the EGR1+ effector cluster.

Imputing function for CD8 T cell population – TF and GO pathways

In order to better understand the underlying regulatory networks that may control peripheral CD8+ plasticity, as well as how these may be perturbed in IBD, we performed gene co-expression analysis. We then scored each cell for the activity of gene modules which were both co-expressed with and enriched for transcription factor cis-regulatory motifs, identifying 273 active transcription factor activity associated circuits. Hierarchical clustering of the activity of these networks highlighted cell type specificity groups, highlighting both new and known well established biological processes (**Fig 2.4A**). For instance, KLF2, LEF1 and KLF3 were active specifically in naïve cells. KLF2 is known to regulate chemokine receptor expression in naïve T-cells[238]. FGFBP2+ cells also showed high activity of KLF2 network[239]

Fig 2.4 - Imputing CD8 T cell subset function - TF, pathway and GO network analysis

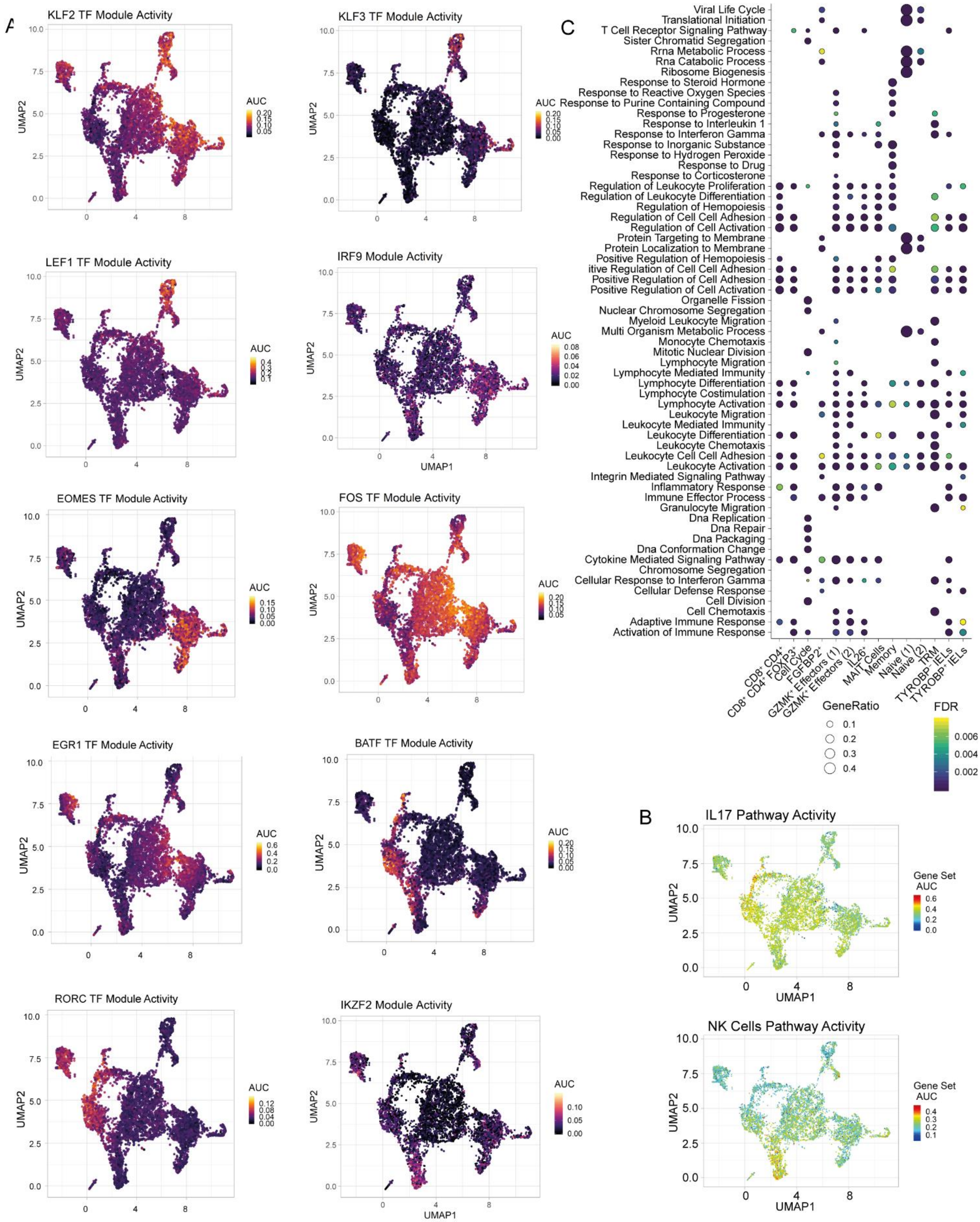


Figure 2.4 (A) Transcription factor network overlay on the UMAP of all cells recovered across UC and Health. (B) Selected pathways overlay on CD8 T cell UMAP, calculated using area-under the ROC curve analysis of canonical pathways databases. (C) Bubble plot of GO pathway analysis by clusters recovered in UC and Health.

The largest cluster of TF-associated networks showed specificity to the small cluster of cycling cells, which included MYC[239] (figure not shown).

GZMK+ effector populations showed specific activity of EOMES, a TF that may compliment T-Bet for full effector phenotype differentiation of T-cells[240], as well as IRF9, an interferon signalling regulator that has been shown to prevent T-cell exhaustion in chronic viral infections[241].

EGR1 and EGR2 module activity was strictly localised to the cell transition gradient between effector-memory pool and GZMK+ effectors, demarking the EGR+ effector cluster that was more clearly delineated in the validation cohort. These cells also showed high, localised expression of TNF (**Fig 2.3C**), as well as activity of FOS, FOSB, FOSL1, JUN and JUNB network activity, transcription factors that are part of recognized immediate early response networks that can be rapidly induced by a variety of stimuli[242].

IL26+ cells and double-positive CD4+CD8+ cells showed a large overlap between active transcription factor networks, including BATF and RORC. RORC, but not BATF was also particular to MAIT cells.

Finally, the highest activity of HELIOS network could be attributed to the TYROBP+ natural IEL population, suggesting a possible regulatory role[243].

Using the curated canonical GO and REACTOME pathways database gene sets we used area-under-the-ROC-curve analysis (**Methods**) to score and identify cells with active gene signatures in a more unsupervised manner, independently of cell clusters. We found localised activity of IL17 pathway in double positive and IL26+ cells (**Fig 2.4B**). IEL (TYROBP + and -) clusters and some IL26+ cells were highly enriched for NK cell and NK-cell mediated cytotoxicity pathways. (**Fig 2.4B**).

We also investigated cluster-specific pathways by performing Gene Ontology (GO) enrichment analysis (**Fig 2.4C, Methods**). Many clusters shared strong enrichment for similar processes, including T-cell activation, differentiation, and proliferation, as well as cytokine production. Cycling cells exhibited a strong mitosis-related GO term enrichment, while GZMK+ effectors and GZMA+SELL-CCR7- cells showed enrichment for chemotactic and cell-migration related processes.

Terms enriched in IL26+ cluster markers included regulation of innate immune response and lymphocyte co-stimulation, amongst others.

Taken together, our findings show the breadth of variability among colonic CD8+ T cells, including IL26+ cells with hybrid innate and adaptive characteristics and double positive regulatory CD8+ T cells.

Differential changes in transcriptomic behaviour in inflammation

We went on to identify changes that occur in inflammation on a cluster-by-cluster basis. On doing so, we identified 997 differentially expressed genes (**Fig 2.5A**). The majority (615) of identified DEGs were significantly differentially expressed within a single cluster only, with just 34 genes that were significantly differentially expressed in more than four clusters. These cell-type independent response genes encompassed loss of expression of *SPINK2*, *FOS* and *CD160*, but upregulation of *TNFRSF9* and *CTLA4* in TRM, effector and IEL populations (**Fig 2.5A**).

GO enrichment analysis of upregulated genes highlighted strong signatures for both type I and type II interferon responses, T-cell activation, cytokine production, cell killing, and upregulated innate immune response pathways (**Fig 2.5B**)

Our TF activity analysis also highlighted gene modules that were not only cell-state specific, but also showed differential activity levels in IBD (**Methods**) including ETV7, STAT3 and PRDM1/BLIMP-1 (**Fig 2.5C**)

We hypothesize that these results reflect differential pathway activity in effector populations in IBD, where we observe higher levels of co-stimulation by CD28, TCR signalling, antigen processing, antigen presentation, PD1 and CTLA4 pathways than in health.

Fig 2.5 - Changes in cellular behaviour in inflammation

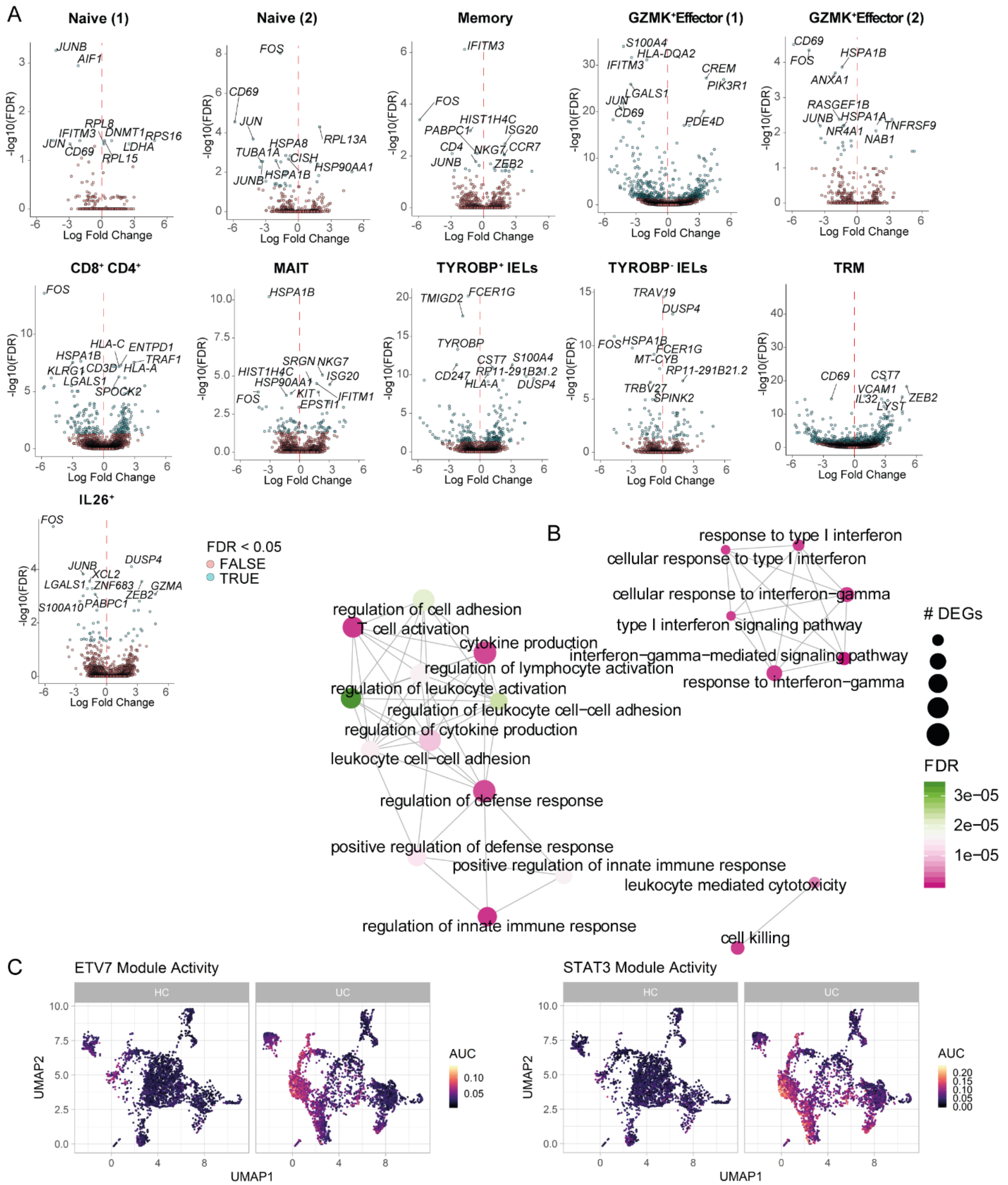


Figure 2.5 (A) Differentially expressed genes on a cluster by cluster basis between ulcerative colitis (UC) and health. (B) GO Pathway analysis of differentially expressed genes in UC versus health (C) Selected differential transcription factor module activity different between UC and health, superimposed on UMAP of CD8 T cells in UC and health. FDR: False discovery rate

Determining functionally relevant populations in ulcerative colitis

In order to determine which of these changes in inflammation might be more relevant for disease pathogenesis, we looked to previously published data for IBD genetic susceptibility, reasoning that populations expressing those genes were more likely to harbour culprits.

As our single cell analysis revealed high resolution expression data, we superimposed the data for UC susceptibility from GWAS studies onto our high-resolution transcriptomic dataset for CD8 T cells (**Methods**), also incorporating previously published sc-RNA data for epithelial and stromal cells[19,29]. We discovered that GWAS genes from UC-associated regions were selectively enriched in active inflammation (**Fig 2.6A**), particularly in CD8 T cells, although some targets were also to be found in undifferentiated stem-like epithelial cells. IL26+ and IELs showed the most significant signal enrichment (**Fig 2.6A**), driven by genes including *KIR3DL2* (rs17771967), *IL26* (rs2870946), and *IL23R* (multiple risk alleles).

Given that increased epithelial cell damage is one of the hallmarks of IBD, we sought to assess how IBD alters the cross-talk between CD8 T cells and epithelial cells. To this end, we identified possible intercellular ligand-receptor interactions between these two populations: UC-specific changes in CD8 T cells in this dataset, combined with published single-cell RNA seq data from epithelium in UC and Health[19] (GEO: GSE116222). Using 2648 known interactions encompassing growth factor as well as cytokine signalling events amongst others (**Methods**), we calculated possible loss (downregulation of either ligand or receptor expression) or gain (upregulation of either ligand or receptor expression) for each CD8+ sub-population (with the 'target' being matched expression across epithelial subsets).

On doing so, we identified 1716 altered cross-talk events in UC (**Fig 2.6B**), with 104 unique R-L pairs across 22 cell types, with gains in interactions constituting the majority of these events (n = 1575), constituting between 15-51 interactions in health and 14-80 in UC. Of note, the populations showing the most changes in interactions in UC included the novel IL26+ cluster of cells, as well as the TNF/IFN rich cell cycling population. The highest number of interactions was with the stem cell subset, which

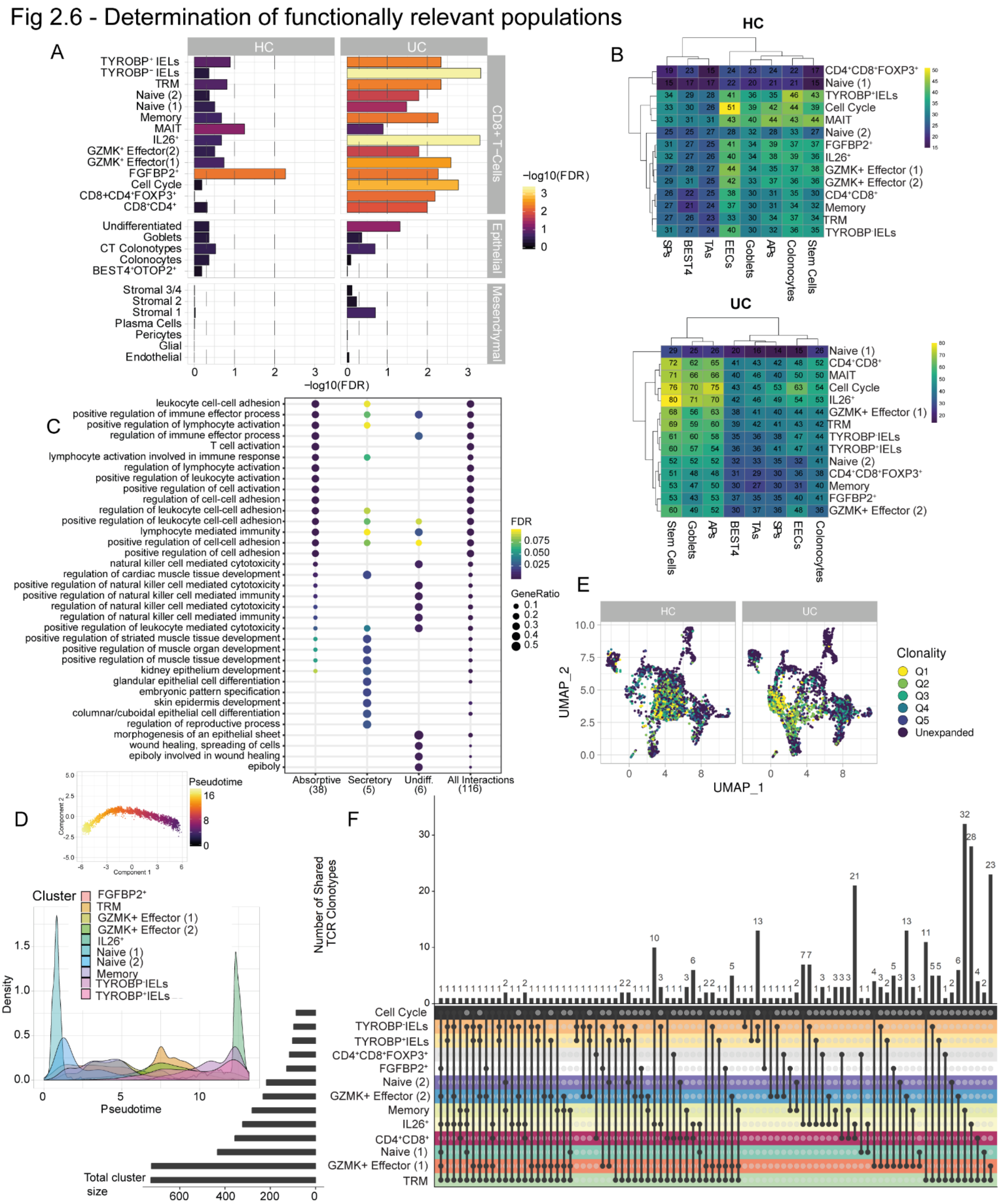


Figure 2.6 (A) Overlay of genes identified as UC susceptibility loci as per GWAS analyses on transcriptomic signal from the scRNA dataset CD8 T cells in health and UC in this study, as well as previously published datasets of epithelium and stromal cells. FDR: False discovery rate. (B) Ligand-receptor analysis in health and UC between CD8 T cell populations and epithelial cells identifying the sum of all differentially up- or down-regulated ligand-receptors on a cluster-by-cluster basis. The majority of these interactions were increase in receptor-ligand signals in UC (C) GO pathway representation of the most signficantly up-regulated receptor-ligand signalling events in UC as compared to health. (D) Pseudotime analysis of transcriptome of CD8 cells recovered in this dataset (double positive and MAIT cells excluded, n = 3HC, 3UC). Cells fall into a linear trajectory (top), with cluster density along this trajectory (bottom), with Naive cells (bottom, left) progressing through TRM and GZMK+ effector subsets, terminating in the IL26 and IEL populations (bottom, right), (n = 3HC , 3UC). (E) Superimposition of degree of clonal expansion on the UMAP of all CD8 T cell subsets, divided by health or ulcerative colitis. Unexpanded – clones in <4 cells. Q1- top 20% of the most expanded clones per sample. Q5 – the bottom 20% of the most expanded clones. Q2-Q4 represent the middle three quantiles (n = 3HC, 3UC). (F) Upset plot showing TCR clones shared between different clusters, each shared clone visualised as an 'interaction', indicated by a black circle joined by black lines. Each unique clone is counted only once, regardless of degree of expansion. The barplot at the top indicates the total number of shared TCR clones for those cluster intersections. Cluster intersections without shared TCRs are omitted for clarity.

is known to be functionally altered by inflammation[38,244]. GO enrichment analysis of the most significantly enriched of these interactions (**Fig 2.6C**) encompassed chemotaxis, leucocyte activation and adhesion as well as epithelial healing. As evidence indicated these effector populations were possibly functionally relevant, we next wanted to understand how they might arise. As we could see from our data that CD8 subsets occupied different points on a transcriptomic continuum, we sought to re-create their development trajectory, reasoning that similar cells are more closely related. In order to reconstruct this, we performed pseudotime analysis (**Methods**) on all recovered cells, excluding MAIT and double positive cells (as it is likely that they arise distinct to the other subsets). On doing so, cells fell into a linear trajectory, with naïve cells at the start progressing into central memory, effector memory and GZMK+ effector cells, with IL26+ cells placed at the end (**Fig 2.6D**)

As our dataset also yielded single-cell TCR information, we could also use this to analyse clonal dynamics and lineages between subsets (**Methods**). We found that naïve, MAIT and DP T-cell populations exhibited highly diverse clonal structures, with most cells expressing a unique TCR CDR3 sequence pair in both UC and health (**Fig 2.6E**), in keeping with our assumptions regarding their development and role. In health, cells of the TRM phenotype showed the most clonality. In contrast, in UC, we found that IL26+, induced TYROBP- IELs and GZMK+ effector cells, in decreasing order, were the most clonally expanded (the majority of the largest clones in each sample displayed these phenotypes).

We also examined the degree of sharing of TCR clonotypes between clusters and their degree of overlap. In total, 320 out of 3835 unique clonotypes occurred in more than one cluster (**Fig 2.6F**). These constituted the clonally expanded populations, with a total of 2438 cells sharing their clonotype with cells in other clusters (**Fig 2.6F**). As could be expected, TRMs shared the most clonotypes with other cell states, most commonly with the GZMK effector (32) , Memory (23) and IL26+ (28 unique TCRs) clusters. Any clone observed in more than two different clusters also appeared in the TRM population. The largest number of “nodes” as well as the largest “triplet” (10 unique clonotypes) involved the cycling cluster, with the triplet identifying a network between the cycling, IL26+ cells and TRM cells. IL26+ cells also shared a number of clones with the induced TYROBP- IEL (7) and CD4+CD8+ double positive populations (21). As expected, the two GZMK+ effector populations shared a significant number of clones (13), as did the two types of IEL populations (13). Thus, TCR repertoire analysis confirmed our transcriptomic analysis and revealed the transitional journey of CD8 T cells through diverse states in UC.

We also compared the TCR sequences against publicly available databases (**Methods**). The expanded clones were diverse (we did not detect any shared specificity groups) and had largely private TCR repertoires, although we detected 93 TRB CDR3 sequences shared between at least 2 donors. These included known public TRB sequences for common EBV, CMV and influenza-derived peptides, potentially reflecting past exposure to common antigens, and none of these were expanded in UC (data not shown for sake of clarity). Despite their sophistication, such analyses have their limitations – public dataset repositories are historically dominated by viral antigen-focused research, do not take into account the particulars of MHC expression by the individual, and have limited representation of self or cancer-related antigens. A lack of such antigen targets or shared specificity in this analysis therefore cannot be taken as definitive proof that self-antigens are not implicated in UC pathogenesis, but simply reflects the incomplete nature of these libraries and our current prediction capabilities concerning protein binding using bioinformatics.

Taken together, both the pseudotime and clonal dynamic analysis indicated that unexpanded naïve cells occupy the start of a trajectory that terminates in an expanded, effector or post-effector IL26+ population of cells that is actively interacting with the epithelium.

Exploring cellular interactions in inflammation

We assessed the ligand-receptor interactions for key populations in more detail, and we present the data from selected populations, including the IL26+ (**Fig 2.7A,B**) and GZMK+ effector populations (**Fig**

Fig 2.7 - Exploring cellular interactions in inflammation

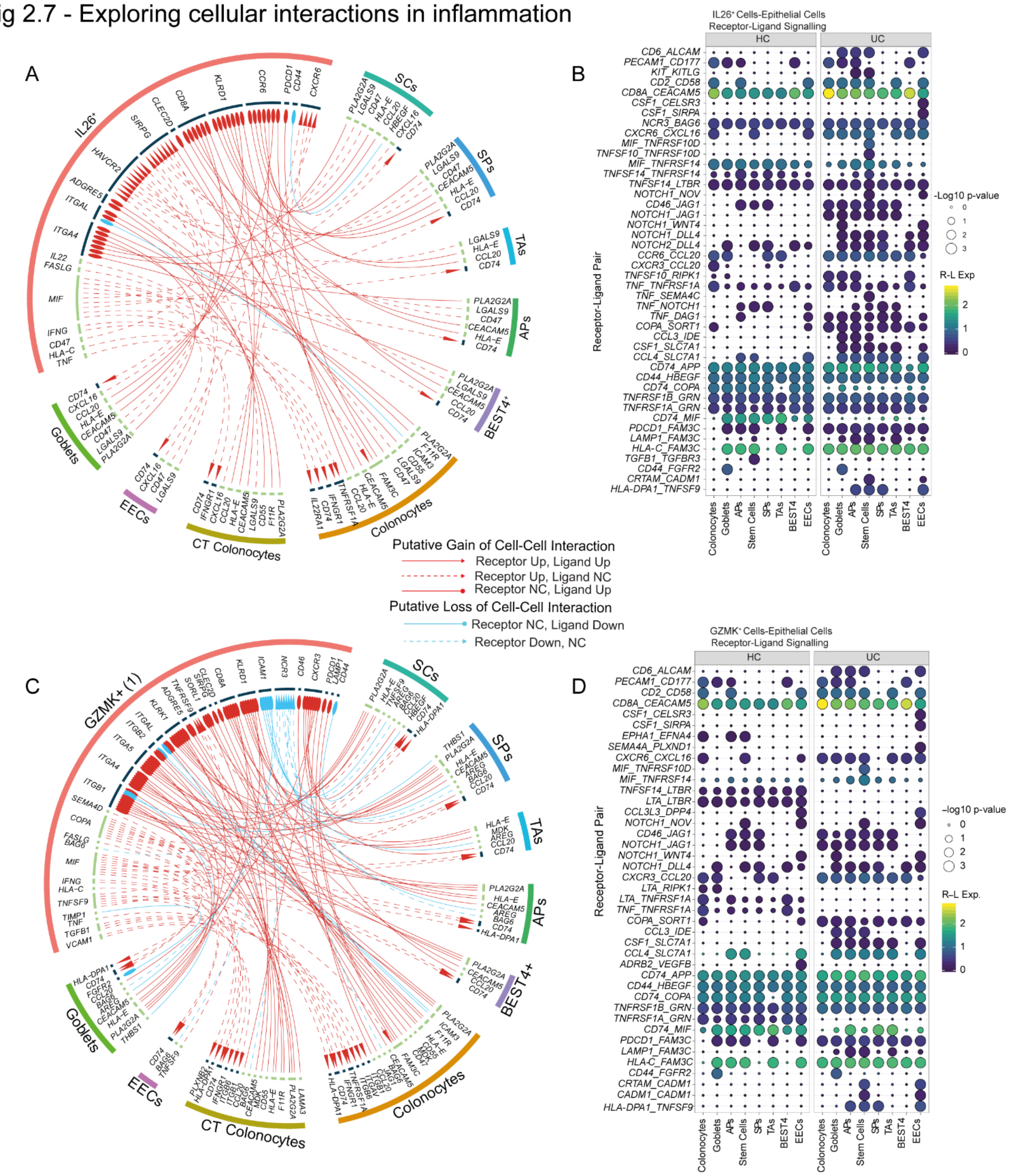


Figure 2.7 (A) Circos plot showing all putative gain (red) or loss of (blue) cell-cell interaction events in inflamed UC as mediated by receptor-ligand pair signalling between IL26+ T-cells and epithelial cell sub-types. (B) Dotplot of selected significant paracrine receptor-ligand interactions between IL26+ cells and epithelial cells discovered using CellphoneDB (empirical permutation p-value) (C) Circos plot showing all putative gain (red) or loss of (blue) cell-cell interaction events in inflamed UC as mediated by receptor-ligand pair signalling between GZMK+ effector T-cells and epithelial cell sub-types. (B) Dotplot of selected significant paracrine receptor-ligand interactions between GZMK+ effector cells and epithelial cells discovered using CellphoneDB (empirical permutation p-value). Abbreviations: SCs: Stem Cells. SPs: Secretory Progenitors. TAs: transit-amplifying cells. APs: absorptive progenitors. CT Colonocytes: Crypt-top colonocytes. EECs: enteroendocrine cells. NC: no significant change.

2.7C,D). As previously mentioned, the majority of these interactions were gain of function interactions (red) as opposed to loss of interaction (blue).

There were multiple cytokine signalling interactions that were increased in UC, for example, *CXCR6-CXCL16* for multiple epithelial subsets including stem cell and secretory cells and IL26+ and GZMK+ effector subsets. This may act to mediate cellular migration at a local level, with the cumulative effect dependent on the relative proportion of different populations. We could see that the non-classical MHC molecule *HLA-E*, a ligand for NKG2 family receptors, including inhibitory *KLRC1* or co-stimulatory *KLRC2* was strongly induced in multiple epithelial populations in UC, with more significant up-regulation in both secretory and absorptive crypt-top cells. *HLA-E* may participate in αβ T cell receptor mediated recognition of peptides derived from stress-related or pathogen-derived proteins[245]. We speculate this plays an important role to regulate mucosal CD8+ cell response in IBD.

While both secretory and absorptive cell types participated in many of these putative signalling events, it is interesting to note that we also detected lineage-specific alterations. For example, these included *IL18-IL18R1/IL18RAP* and *TNF-TNFRSF1A/TNFRSF21* signalling between absorptive, but not secretory cells, and multiple CD8+ sub-clusters.

With respect to the novel IL26+ population, taken together with our previous analysis, our data indicated that although this cluster had many interactions with epithelium in common with previously characterized clusters e.g. effector GZMK+ cells. In order to clarify the overall effect, we sought to identify the functional effect of IL26 in inflammation.

IL26 in ulcerative colitis – detecting source and target cells

In order to ascertain the effect of IL26, we first sought to identify the cells producing it.

A single clone of anti-IL26 antibody (AK-155) is extensively validated in literature[246,247]. Although initially promising, we determined that this clone was inaccurate at detecting IL-26 expressing cells by FACS in our populations of interest. This was because we detected a large number of IL-26 expressing

CD8 cells in peripheral blood (**Fig 2.8Ai**), where we know from public datasets that IL-26 expression is minimal. To confirm our conclusion, we utilized a HDLM-2 cell line known to over-produce IL-26 in culture[248]. We were able to demonstrate that despite HDLM-2 cells expressing significantly more IL-26 RNA than PBMCs (by a factor of ~1500, **Fig 2.6Aii**), AK-155 detected similar amounts of IL26 protein in both HDLM-2 cells and PBMCs despite antibody concentration titration (**Fig 2.8Aiii**).

AK-155 was similarly ineffective at detecting IL-26 using immunohistochemistry (IHC), demonstrating non-specific binding patterns despite titration and variation of antigen retrieval methods, as well as equivalent binding in health and UC (**Fig 2.8B, left panel, top and bottom**). We also attempted using in-situ hybridization (ISH) - we could see that the frequency of the IL26 transcript is low in a typical section of UC (**Fig 2.8B, right panel, top and bottom**), detecting 1-2 positive spots per section. For our CyTOF analysis of these populations, we therefore decided to use surface IL23R as a marker of IL26 after excluding MAIT cells using appropriate gating[2].We next tried to identify the cells expressing the cognate receptor for IL26. This is described as a heterodimer of IL10RB/IL20RA[247], although some evidence exists that this may not be the complete repertoire of receptors acted upon by IL26[249]. We first assessed the expression of the heterodimer by sc-RNA seq in published datasets for epithelium, stroma and epithelium. While IL20RA had a low detection rate, IL10RB localised to colonocytes in the epithelium and was widely expressed by mesenchymal cells, but not CD45+ cells, in both health and UC (**Fig 2.8C**).

IL10RB/IL20RA expression was also assessed by FACS and discovered to be co-expressed in a small proportion of epithelial (1.9%), mesenchymal (1.73%), and immunological compartments (3.7%) (**Fig 2.8Di**). We used a known positive control SW480[250,251] cell line to confirm the expected expression pattern (**Fig 2.8Dii**) was as we observed.

By both sc-RNA seq and FACS data, the expression of the heterodimer did not increase in UC as compared to health (**Fig 2.8C,Diii**).

Fig 2.8 - Determining the effects of IL26 - detecting cells of origin and effect

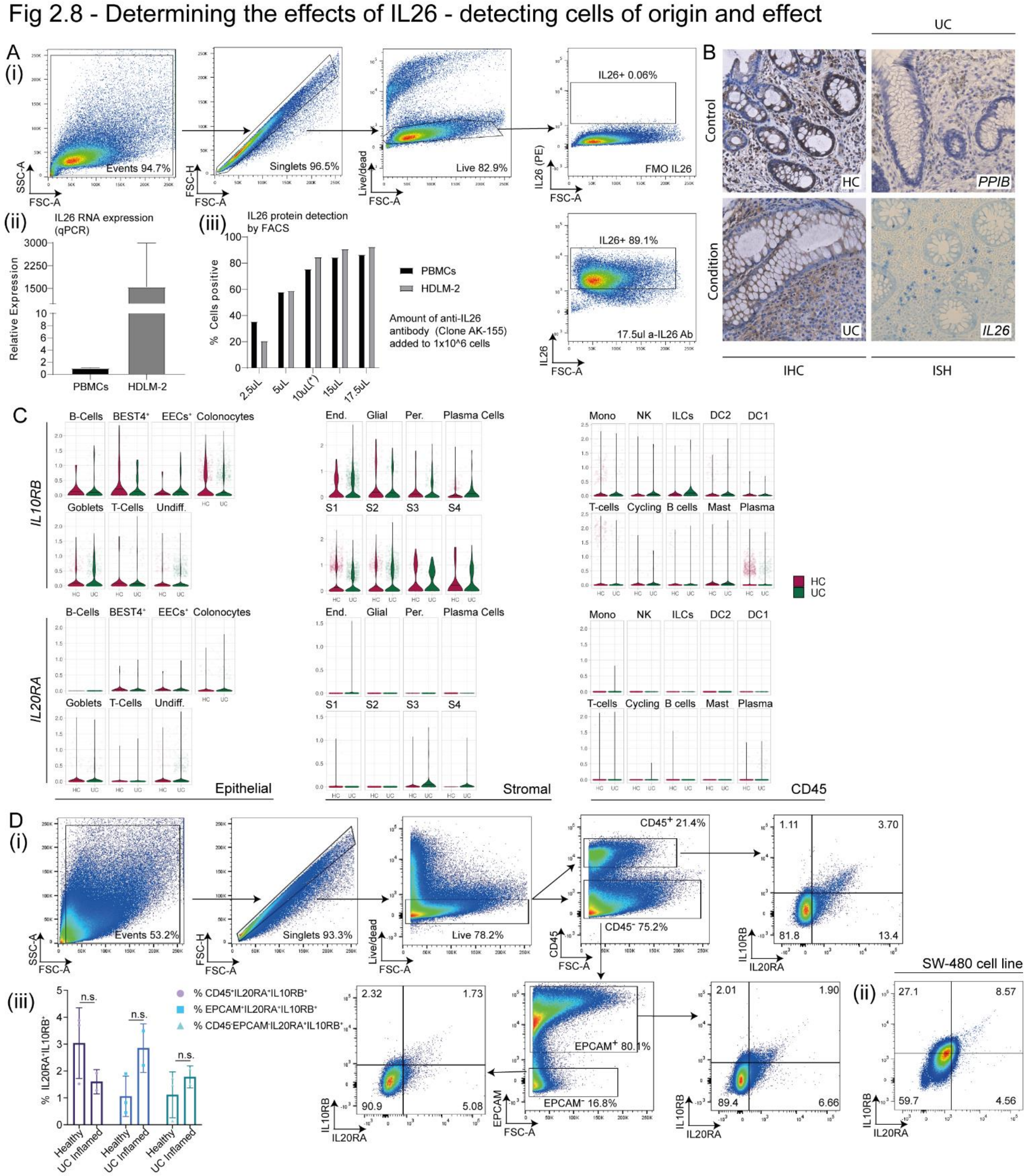


Figure 2.8 (A) (i) FACS analysis of IL-26 expression using AKT-155 antibody clone in PBMCs, demonstrating significant positivity as assessed by FMO analysis, a result at odds with publicly available transcriptomic information. (ii) qPCR analysis of IL26 expression by RNA in PBMCs and the HDLM-2 cell line, reference gene: *RPLP0, GAPDH*. (iii) FACS analysis of proportion of PBMC or HDLM-2 cells positive for IL-26 using a titration series of AK-155 antibody, gating as per (i), demonstrating AK-155 binds non-specifically. (B) IHC and ISH analysis of IL26/*IL26* expression in UC as assessed using AK-155 antibody/RNAscope RNA probe respectively as compared to healthy control or *PPIB* control gene. All images at 10X magnification. (C) Violin plots of sc-RNA seq analysis of expression of known IL26 receptor, heterodimer *IL10RB/IL20RA* in epithelial, stromal and CD45 fractions in health and UC (D) FACS analysis of expression of IL26 receptor, heterodimer of IL10RB/IL20RA at protein level across (i) epithelium (CD45-EPCAM+), stromal (CD45- EPCAM-) and CD45 fractions in representative sample (HC) (ii) in EPCAM+ fraction of SW480 cell line known to express the receptor per published literature (positive control) (iii) across health and ulcerative colitis, n = 3 HC, n = 2 UC in epithelial (p=0.09, t=2.43), immune (p= 0.25, t=1.42) and mesenchymal compartments (p = 0.39, t=0.99). two tailed unpaired t-test. DF=3. Mean and SD are shown, Abbreviations: PBMCs - peripheral blood mononuclear cells, IHC: immunohistochemistry, ISH: In-situ hybridization. n.s: not signficant.,qRT-PCR: Quantitative real-time polymerase chain reaction. HC: healthy controls, UC: ulcerative colitis, Undiff: Undifferentiated, ILCs: Innate lymphoid cells, DC: dendritic cells, NK: Natural killer cells, End: Endothelium, EEC: Enteroendocrine cells, Per: Pericytes, S1-S4: Stromal 1 - 4.

Taken together, this data indicated it would be challenging to isolate sufficient numbers of IL26+ cells from a typical biopsy sampling approach in UC in order to perform co-culture experiments. It also indicated that the likely target population lies within the epithelial compartment, although the possibility of action on stromal and immune/myeloid compartments through as yet undiscovered receptors could not be excluded.

Determining the possible functional effects of IL26

In order to try to understand the effect of IL26 on the colonic epithelium, we attempted a number of different strategies.

The *IL26* gene is absent in wild-type mice, however, its cognate receptor (IL10RB/IL20RA) is present, and appears to be capable of inducing signalling following stimulation with IL26 [252,253]. A bacterial-artificial-chromosome (BAC) induced IL-26 expression model in mice (hIL26-Tg) had been developed by another group, where mice were engineered to express a 190-kb human BAC transgene containing the human IL26 gene, with expression confirmed in the small intestine and colon[163]. We collaborated with this group in order to assess the possible effects of IL26 on colitis.

For the experimental model, we exposed human IL-26 expressing (hIL-26Tg) and BAC-negative, sibling wild type (WT) C57BL/6J (B6) mice to 2.5% DSS in drinking water for a period of 6 days (n = 3WT, n = 4 hIL-26Tg mice). In addition, hIL-26Tg mice were injected with neutralizing anti-IL26 antibody (developed and validated by our collaborators) or control antibody on days 0 and 3 (**Fig 2.9A,** details in **Methods**). Mice were sacrificed on day 6, with RNA extracted from colonic tissue and subjected to bulk RNA sequencing (RNA-seq) and analysis.

We initially confirmed that we could observe IL26 expression in colonic tissue in hIL-26Tg mice and that this increased significantly on induction of DSS colitis (data not shown). Under baseline conditions, hIL26-Tg mice and WT mice clustered closely together by PCA analysis (**Fig 2.9B**), but we still identified 295 significantly (< 5% FDR) differentially expressed genes between WT and hIL-26Tg conditions (**Fig 2.9C**). Of note, even at baseline, in WT mice we observed a relatively higher expression

Fig 2.9 - Determining the effects of IL26 - functional experiments

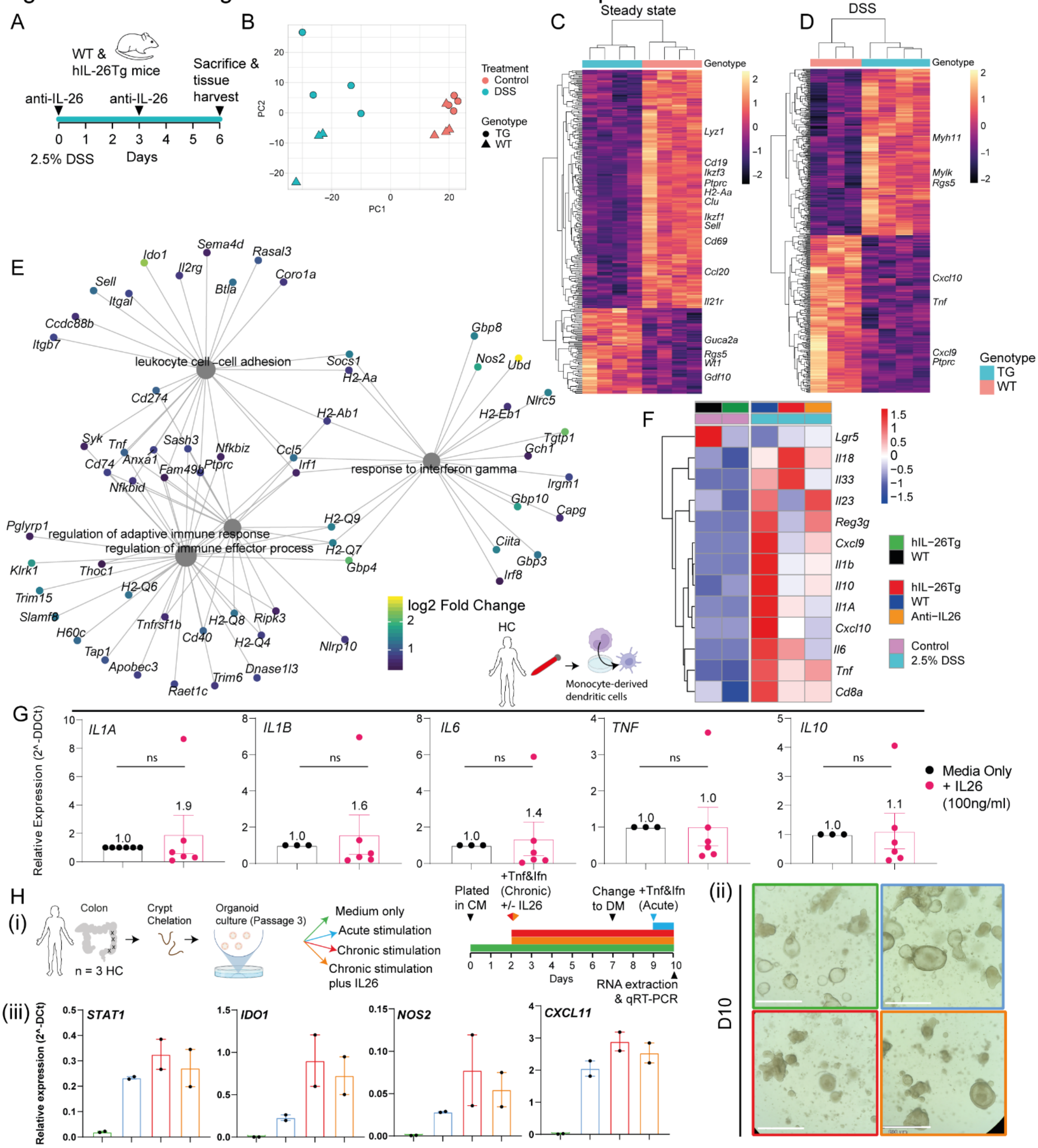


Figure 2.9 (A) Experimental overview of induction of dextran sulphate sodium (DSS) colitis in C57BL/6J (Wild-type, WT) and littermate human IL-26 producing transgenic (hIL-26Tg) mice versus controls with no colitis. Combined with sub-study where hIL-Tg DSS mice received either intraperitoneal (i.p.) injection of anti-IL26 antibody or mIgG1,k isotype control antibody. (B) Principal component (PCA) analysis of mRNA from bulk RNA-seq of WT (n = 4 DSS treated mice, n = 3 controls) and hIL-26 (n = 4 mice in each group) colonic tissue. (C) Heatmap comparing differentially expressed genes between WT and hIL-26Tg mice not subjected to DSS challenge (steady state) (D) Heatmap comparing differentially expressed genes between WT and hIL-26Tg mice subjected to DSS challenge (colitis). Both (C) and (D): Wald test, Benjamini-Hochberg multiple testing correction. (E) Cnetplot showing the most enriched GO terms in significantly downregulated genes in DSS challenge IL-26Tg mice as compared to WT mice under DSS challenge (n=4 hIL-26Tg, n=3 WT). Hypergeometric over-representation test, Benjamini-Hochberg multiple testing correction. (F) qRT-PCR analysis of selected genes in colonic tissue from mice under steady state and 2.5% DSS challenge, with anti-IL26 or control mIgG1,k isotype control injection given on days 1 and 3 as per experimental overview. (G) qRT-PCR analysis of selected genes in human PBMC-derived monocytes from HC differentiated into dendritic cells, exposed to IL-26 or vehicle control, showing no significant differences in inflammation signals, Mean + SEM, paired ratio t-test, p = 0.52 (*IL1A*), 0.54 (*IL1B*), 0.88 (*IL6*), 0.67 (*TNF*), 0.49 (*IL10*) (H) (i) Experimental overview of human colonic organoids exposed to medium alone (control), acute (TNF 1ng/ml and IFNg 1ng/ml for 24 hrs), chronic (TNF and IFNg, same concentration, 8days) or chronic plus IL-26 (10ng/ml) inflammatory conditions. (ii) Appearance of a representative organoid culture from all four arms on day 10 of experiment, 10X magnification, scale bar = 890uM (iii) qRT-PCR of lysate from colonic organoids from all four arms on day 10. n = 3 HC pilot, however, one sample developed an infection, so data from two surviving cultures is presented.CM: Conditioning medium, DM: Change to differentiation medium. IL26: Interleukin 26, D10: Day 10, HC: Healthy controls

of immune signatures – e.g. CD45 cells (*Ptprc*), B cells (*Cd19*), activated T cells (*Cd69*) as well as lymphocyte differentiation markers (*Ikzf1, Ikzf3*), as well as lysozyme (*Lyz*) typically produced by macrophages in the colon, with a corresponding reduction in non-immune cell markers (*Rgs5*, pericytes; *Guca2a,* BEST-4 cells). Taken together, it suggested that perhaps even under baseline/steady-state conditions, IL-26 may reduce immune infiltration and signalling.

In DSS-induced inflammation, these differences became even more evident (**Fig 2.9B, 2.9D**), with 473 differentially expressed genes between WT and hIL-26Tg mice. We observed significantly lower pro-inflammatory cytokine (*Tnf*) and chemokine signalling (*Cxcl9, Cxcl10*) in hIL-26Tg mice as compared to WT mice, with a reduction in immune cell burden (*Ptprc*) in colitis.

GO pathway analysis of significantly downregulated genes in DSS challenge hIL-26 mice showed reductions in leucocyte adhesion, adaptive and effector immune responses (**Fig 2.9E**), with subsequent reductions in epithelial and stromal response to interferon gamma (e.g. *Nos2*), all in keeping with an effect to reduce inflammation. Finally, we performed qRT-PCR analysis of the arm of hIL-26Tg DSS-challenge mice treated with neutralizing IL-26 antibody on days 0 and 3 (**Fig 2.9F**) and observed that many of the inflammation-ameliorating effects of IL26 were reversed, with increases in *Cxcl9*, *Il23* and *Tnf* signalling in IL26 blockade.

We then wished to see if we could replicate some of these findings in a human model. Given that many of these transcriptomic changes could be explained by differences in myeloid cell behaviour, we initially explored the effect of IL-26 on monocyte-derived dendritic cells (**Fig 2.9G, Methods**). This simple model did not demonstrate any clear observable effects of IL-26 incubation on any key inflammatory transcripts (**Fig 2.9G**) that we had seen using our mouse model.

Finally, we decided to investigate the possibility that IL-26 was acting directly on the epithelium, given our scRNA-seq data had indicated multiple interactions with the epithelial compartment, as well as detecting its receptor by FACS and sc-RNAseq on epithelial cells. We performed a pilot experiment with human colonic epithelium organoids (**Fig 2.9Hi, Methods**), simulating acute and chronic inflammation as well as assessing chronic inflammation in the presence of IL26. We could see that 'chronic' inflammation appeared to induce more budding than 'acute' stimulation, with smaller, denser organoids, which was partially ameliorated in the presence of IL26 (**Fig 2.9Hii**). Transcriptionally, we could see clear elevations downstream of JAK-STAT pathway activation (*STAT1, IDO1*), evidence of epithelial stress (*NOS2*) and chemokine production (*CXCL11*) in both 'acute' and 'chronic' inflammation, signals which are well described in epithelium in UC as compared to health, suggesting this model replicates these features of UC (**Fig 2.9Hiii**). Interestingly, we saw a trend for key transcriptomic changes in 'chronic' inflammation reflective of epithelial damage and inflammatory cascade activation improving in the presence of IL26 (**Fig 2.9Hiii**).

Discussion

The goal of our analysis was to characterize CD8 T cells in ulcerative colitis and health utilizing novel single-cell RNA sequencing in an effort to discover novel disease associated states, understand CD8 T cell function, clonal dynamics and interaction with epithelium.

UC associated disease states: Our analysis identified multiple novel cell states in UC and health, including highly activated EGR1+ cells, two populations of effector cells (GZMK+, FGFBP2+), and confirmed the existence of double positive CD4+CD8+ and CD4+CD8+FOXP3+ cells, as well as a new IL26+ population. The IL26 population bore many of the gene signatures that have been identified to increase the risk of UC by GWAS studies, suggesting a functionally relevant role for this population. It confirmed previous paradigms in UC, such as an expansion of effector cells and a relative reduction in TRM resident cells, but also demonstrated new ones, such as an expansion of IL26+ cells, CD4CD8+ cells and induced TYROBP- IELs in inflammation.

Understanding CD8 T cell function Combining transcriptomic with GO pathway and transcription factor (TF) analysis allowed us to hypothesize function for these cell states. Consequently we were able to link multiple TF modules to cell states, including LEF1 (Naïve), EGR1 (activated), IKZF2 (IEL) and BATF/RORC (IL26). Concerning the IL26+ population in particular, we were able to see high expression of exhaustion/antigen experience markers (*PDCD1, HAVCR2*, *CTLA4*). They also however demonstrated moderate activation of the IL17 pathway, believed to be pathogenic, as well as upregulating *GZMA* in inflammation.

Understanding T cell clonal dynamics Single-cell TCR combined with pseudotime analysis of transcriptomic data allowed us to hypothesize how T cells underwent state changes in tissue. IL26+ cells occupied the end of the trajectory (that originated with naïve cells), whilst also being highly clonal, sharing multiple clonotypes with cycling, resident, IEL and GZMK effector cells, again supportive of the hypothesis that they represent a key CD8 T cell phenotype in inflammation. TRMs shared multiple clones with cycling, GZMK effector, double positive and IEL populations, which coupled with the relative depletion of TRM cells in UC suggests that not only do cells shuttle actively between these phenotypes, but that cells with resident properties play a relatively small role in UC, and circulating/re-circulating cells may have a greater role. Double positive cells were a distinct, definite phenotype that shared multiple clones with CD8+FOXP3+, TRM and IL26 clusters, possibly implying they occupy a transitional state between resident and more tolerogenic FOXP3+ populations that increase in UC. Interestingly, there was also relatively little sharing between the FGFBP and GZMK effector populations despite phenotypic similarity, suggesting differences in origin, though this may be disproven with larger sample sets.

Understanding interactions with epithelium : Through ligand-receptor analysis, we uncovered multiple possible interactions of diverse CD8 T cell subsets with epithelium, which we went on to explore using mouse and human colonic models of disease, focusing specifically on the novel IL26 population. We were able to show, that in line with our hypothesis, IL26 appeared to demonstrate a protective role in the context of inflammation, and this may be being partially mediated through the epithelium, although clearly, further replication and characterization of these preliminary results would be required to confirm a possible therapeutic role for IL26.

Taken together, this analysis helped better define CD8 T cell roles in UC, describe novel populations and impute function. We utilized novel organoid culture systems to model the effect of IL26 on epithelia, with promising pilot data, and this approach could be developed further to understand the mechanism and possible therapeutic implications of this, and other cytokines at a much lower cost than current approaches.

A significant caveat however is that this analysis could not delineate whether any of these changes were specific to ulcerative colitis or whether they also occur in all inflammation, regardless of aetiology. We also are ignorant of the effect of the peripheral blood compartment, as well as trafficking between blood and tissue. We also could not look beyond CD8 T cells to the effect of myeloid and CD4 cells on epithelial tissue. In order to improve on this, we proceeded to compare ulcerative colitis and checkpoint inhibitor-induced colitis (**Chapter 3**).

Chapter 3

Comparing ulcerative colitis, checkpoint inhibitor-induced colitis and health

Introduction

Checkpoint inhibitors are a novel class of drug that entered mainstream clinical use for the treatment of a wide spectrum of metastatic cancer in the last decade. They exist in a separate class to prior strategies for the treatment of cancer, as previously stated, acting to rouse tolerant 'exhausted' immune cells within the tumour microenvironment as well as central lymphoid tissue to induce a host inflammatory response to cancer[96,100]. In many patients, this is of sufficient magnitude and duration to induce remission for many years, an unprecedented outcome in metastatic disease. The discovery revolutionized treatment, earning their discoverers the Nobel prize.

However, it was quickly recognized that immune checkpoint blockade was associated with induction of an inflammatory response in sites without a tumour load, indicating these pathways play an important role in maintaining immune homeostasis. The most common cause of substantial morbidity and ICI discontinuation is inflammation in barrier sites, particularly colonic inflammation (checkpoint inhibitor-induced colitis, CC). It can cause life-threatening consequences, such as colonic perforation, although the incidence is low. Even if colectomy is not required, the severity and persistence of symptoms can be disabling, significantly reducing quality of life[105].

Our investigation of the clinical patterns of disease (**Chapter 1**) demonstrated this immunotherapy-induced or checkpoint-induced colitis (CC) has distinct clinical patterns to idiopathic colonic inflammation, ulcerative colitis (UC). However, it also shared several features with UC, such as exhibiting a chronic inflammation, responding to similar drugs, being mucosal, predominantly colonic (particularly distal colonic), with histological features most commonly resembling UC. As also explained in the introduction, we therefore decided to compare CC with UC, with the aim of identifying processes and pathways unique and common to both, with an aim of yielding mechanistic insights.

Prior work within our group[2,19,24,29] and abroad[20,25,187] utilized unbiased single-cell RNA sequencing to identify new cell types and cell states associated with inflammation in epithelial, stromal and immune compartments, and when combined with transcription factor and pseudotime analysis, allowed researchers to posit function and cell development trajectories. Such unbiased analysis was therefore attractive when considering advancing our understanding of a new disease state (CC), as well as providing a more robust understanding of which changes are unique to UC and versus those shared with any chronic GI inflammatory process.

Multiple models of ulcerative colitis (e.g. DSS and T cell transfer models in mice)[231] have been developed and extensively characterized; despite their individual limitations, they have proved useful in distinguishing correlative events in inflammation from causality. Given how recently CC has occurred as an entity in humans, such model systems are rare/poorly developed[178,254].

CC colitis does not naturally occur in PD-1 or CTLA-4 heterozygous knockout mice (homozygous CTLA 4 knockout mice do not survive till birth)[255,256], and only one model of DSS- augmented CC colitis has been developed so far[257]. As it is unclear to what extent this form of augmented colitis mirrors spontaneous disease observed in humans, our research focused on primary human samples.

Aims

We performed unbiased single-cell RNA sequencing and spatial transcriptomic analysis of patients with CC and UC, comparing it with health, and broadened our scope to consider all possible compartments – epithelium, stromal and immune in order to:

1. Phenotype and compare immune and non-immune cell behaviour in checkpoint inhibitor-induced colitis, health and ulcerative colitis.

2. Understand cellular interaction in health and disease and attempt to identify self-propagative positive feedback networks maintaining chronic inflammation
3. Predict which patients are likely to develop checkpoint inhibitor-induced colitis

Results

Single cell RNA sequencing reveals multiple subsets of immune, stromal and epithelial cells in health and disease

To compare the full range of intestinal homeostasis and dysregulation, we first generated a multi-modal single cell dataset (**Fig 3.1A**). This comprised matched colonic and PBMC samples from patients with active UC (UC_I), paired non-inflamed (UC_NI) from histologically non inflamed areas from the same patients, CC colitis (CC_I), and immunotherapy-treated patients with no colitis (CC_NI) along with healthy controls (HC). Biopsies were taken from the sigmoid/descending colon for both UC and CC.

Of note, the CC_NI colonic samples were not matched biopsies from the same patients because CC_I patients had disease that was patchy and beyond the extent of the endoscopic examination, so obtaining non-inflamed biopsies from the same patient was not possible. This was distinct to the pattern of sampling that was chosen for UC for multiple reasons. Firstly, UC was amenable to clear demarcations between inflamed and non-inflamed tissue; secondly, such analysis was in line with prior literature; and thirdly, it avoided inter-patient confounding factors.

Our dataset therefore does not permit comparisons between UC_I and UC_NI or UC_NI and HC when regarding questions of cell trafficking, and it also does not allow us to draw any conclusions about the time course of resolution in UC or CC colitis. Our comparisons between CC_I and CC_NI are open to inter-patient bias, but we sought to mitigate this by sampling a large number of patients in our analysis.

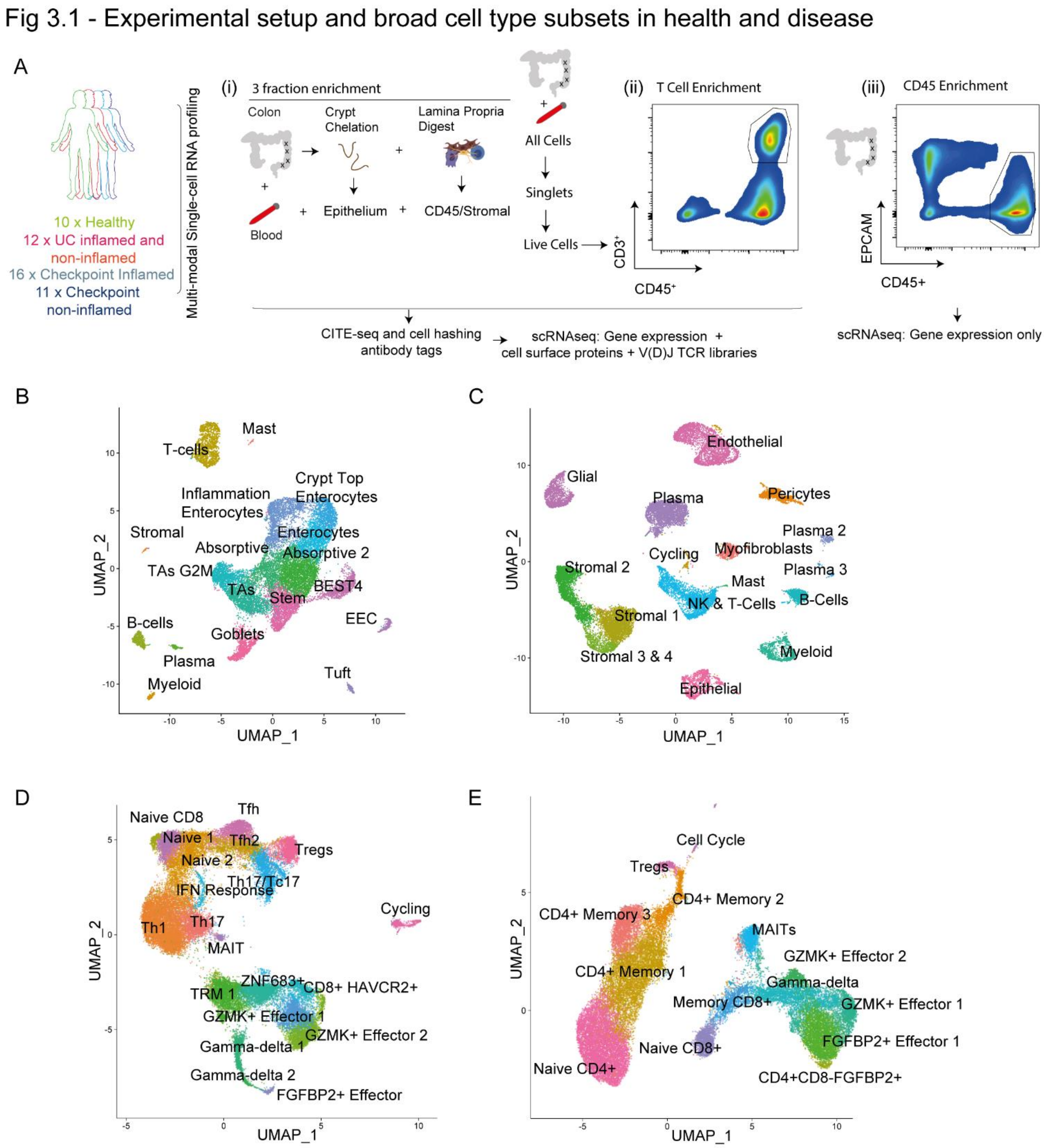


Figure 3.1 (A) Experimental overview of project. A comparison of healthy controls with checkpoint inhibitor-induced colitis, patients given checkpoint inhibitors without colitis and ulcerative colitis (inflamed and non inflamed) across (i) enriched epithelial, enriched stromal and blood compartments utilizing multi-modal single-cell RNA sequencing, yielding unbiased transcriptional information (Gene expression libraries) paired with selected cell-surface protein information (CITE seq libraries) and single-cell T-cell receptor sequences (TCR libraries). (ii) FACS strategy for selective enrichment of T cells in a second iteration of samples, cells from the highlighted gate were sorted for scRNA-seq (iii) FACS strategy for selective enrichment of CD45 cells for a third iteration of samples, cells from the highlighted gate were sorted for scRNA-seq. (B) UMAP visualzation of all epithelial cell subtypes recovered across health and disease by sc-RNA seq (C) UMAP visualization of lamina propria-enriched stromal and CD45+ (immune) populations recovered across health and disease by sc-RNA seq (D) UMAP visualization of all tissue derived T cell populations recovered across health and disease by sc-RNA seq (E) UMAP visualization of all blood-derived T cell populations recovered across health and disease by sc-RNA seq. Abbreviations: CITE-seq: Cellular Indexing of Transcriptomes and Epitopes by Sequencing. TCR: T-cell receptor, TAs G2M: Transit amplifying G-to-M- phase epithelial cells. scRNA-seq: single-cell RNA sequencing.

When selecting patient samples to incorporate in the study, we ensured that the exposure to biological drug therapy was avoided. Therefore, for the experiments interrogating cellular trafficking, none of the patients with UC_I or CC_I were exposed to vedolizumab (which therapeutically binds a4B7 and inhibits transport to the colon). We also attempted, where possible, to minimize exposure to disease-modifying agents overall, so 95.5% of UC and 85% of CC patients included in transcriptomic analysis were biologic naïve, and equivalent numbers of CC_I (35%) and CC_NI (45%) patients were steroid-naive. We also checked for parity in degree of inflammation between UC (median UCEIS 5) and CC colitis (median UCEIS 3.5) as well as median duration of inflammation (UC_I = 34 days, CC_I = 25 days; p=0.13).

We sampled equally from CC_I and CC_NI patients given monotherapy (anti-PD-1 therapy alone) and dual therapy (combination anti-PD-1 and anti-CTLA-4), the two common regimens utilized for immunotherapy treatment of cancer (full patient details in **Appendix D**).

We employed single-cell chemistry that yielded matched RNA-Seq (transcriptome), VDJ-Seq (TCR information) and CITE-Seq information (cell surface protein information; antibody list in **Materials**) in an iterative approach that focused on different fractions (**Fig 3.1Ai-iii**). We optimized protocols (**Methods**) to isolate epithelial, stromal and immune populations from the same patient (**Fig 3.1Ai**), and in successive iterations focused in on the CD45 (**Fig 3.1Aiii**) and T cell fractions (**Fig 3.1Aii**) using fluorescence-activated cell sorting (FACS) based cell enrichment, capturing 155,421 cells in total. These were split across the Epithelial (21,135), Stromal/CD45 (26,907) and Immune (107,379) compartments.

Following data processing, quality control, and batch correction, unsupervised clustering (**Methods**) identified the majority of major cell types expected in the colon and blood (**Fig 3.1B-E**, **Supplementary Table 3.1, 3.2 and 3.3**), which we annotated using cluster marker genes (**Additional resources**) with reference to previously published single cell resources[20,24,25,29,59,187,188,258] (**Methods**).

Epithelial changes in inflammation

The epithelial fraction was generated utilizing a protocol that enriched for the crypt fraction from colonic biopsies. We recovered 21,135 cells across 18 populations, encompassing all significant colonic absorptive and secretory cell populations, as well as immune cells that were in close association with the epithelium (**Fig 3.2A**). Cells were annotated based upon their transcriptomic signatures (**Fig 3.2B**), utilizing previously published datasets as reference[19,20,24]. There were no disease-specific epithelial cell states (**Fig 3.2C**), although there were clear abundance differences between inflammation and health (**Fig 3.2D**).

When examined at higher resolution, inflammation associated enterocytes (denoted by interferon-gamma response genes such as *IFI27* and *DUOX2*) were found in greater abundance in CC_I and UC_I versus their non-inflamed counterparts (**Fig 3.2E**). The relationship was inverted for crypt top enterocytes, likely due to increased cell death and more rapid cell turnover in inflammation. Proportions of other cell types such as enteroendocrine cells and stem cells were similar between inflammation and non-inflamed colonic tissue (data not shown). Mucus-producing goblet cells showed a trend towards being increased in CC_I, which may be relevant given the typically mild appearances of CC_I with low UCEIS scores seen at endoscopy as described in Chapter 1.

A high level overview of the transcriptional behaviour of epithelial cells can be visualised as a principal component analysis with all cells from a disease subtype being grouped together (**Fig 3.2F**). On doing so, we observed clear differences between inflammation and health, with more subtle differences between CC_I epithelium and UC_I. Notably, whereas UC_I and UC_NI are similar, CC_NI is indistinguishable from HC. We can also appreciate that there appear to be differences within CC_I depending on whether samples were treated with mono (anti-PD1 regimen) or dual therapy (combination anti-PD1 and anti-CTLA4 regimen).

We were particularly interested in exploring the differences between CC_I and UC_I given that our current understanding of UC is that intrinsic defects in the epithelial barrier are thought to perpetuate

Fig 3.2 - Epithelial compartment in health and disease

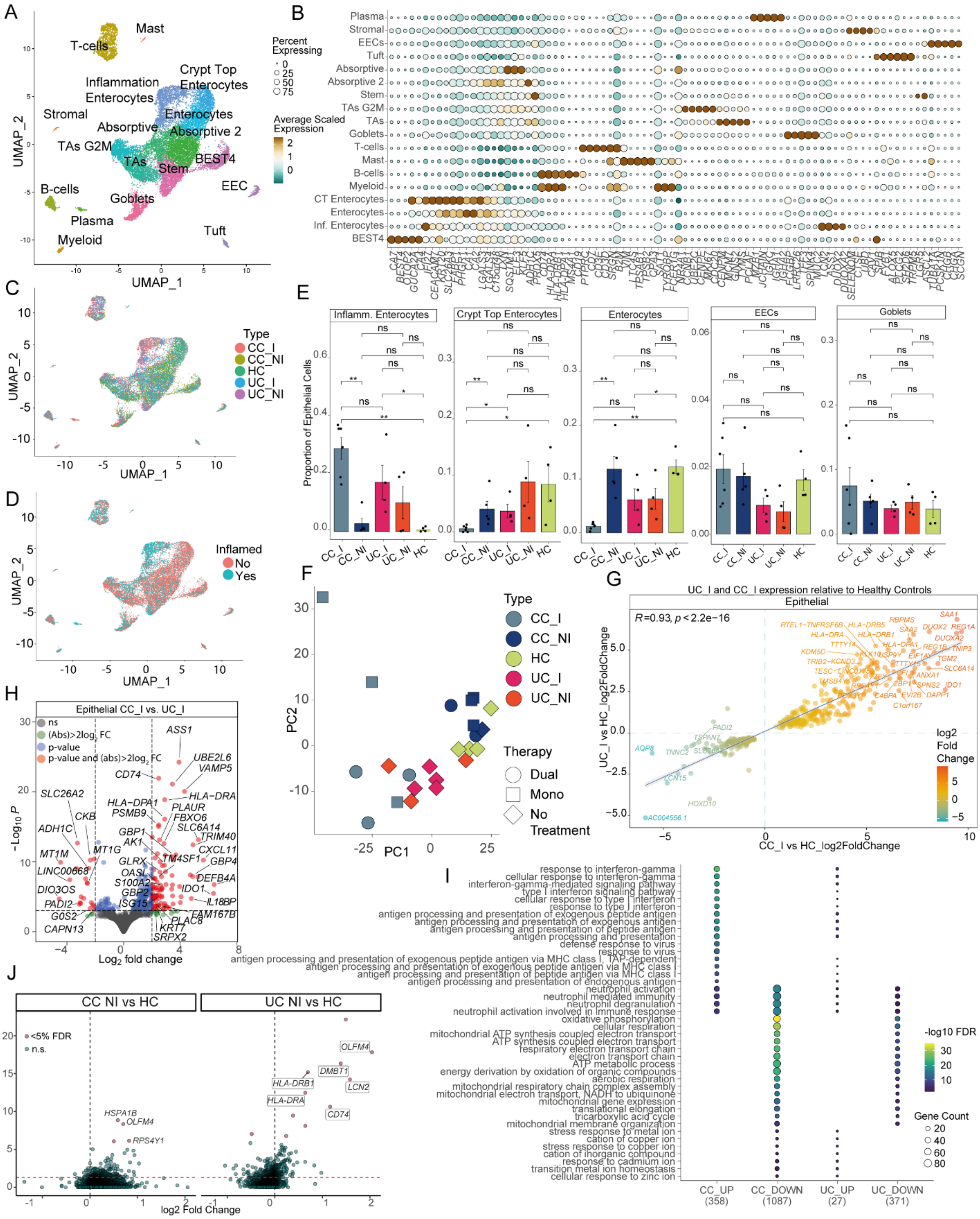


Figure 3.2 (A) UMAP of epithelial subsets in health and disease. (B) Selected gene expression across epithelial subsets in (A). (C) Epithelial UMAP by sample type (D) Epithelial UMAP by inflammation, showing clear inflammation related populations. (E) Selected cell type proportions by condition (F) PCA analysis of global gene expression by sample type, showing CC_NI clustering with HC, UC_NI with UC_I and CC_I distinct from all (G) Plot of genes in epithelial cells up- and down-regulated in CC_I and UC_I (H) Genes selectively enriched and depleted in CC_I and UC_I (I) GO analysis of pathways preferentially up and downregulated in CC_I and UC_I. (J) Plot of genes in epithelium in CC_NI and UC_NI as compared to HC p<0.05(*), p<0.01(**), ns: not significant, Statistical test: Pairwise T test with multiple comparison correction. CC_I : Inflamed checkpoint-inhibitor induced colitis, CC_NI: non inflamed colonic samples from patients given checkpoint inhibitors, HC: healthy control, UC_I: Inflamed ulcerative colitis, UC_NI: Non inflamed ulcerative colitis.

inflammation. Interestingly, the epithelial response was similar in both states, whether up- or down-regulated genes were considered (**Fig 3.2G**), with 2071 dysregulated genes in CC and 766 genes in active UC versus health. The differences were however those of degree, e.g. *IDO1* higher in CC_I versus *DUOX2* higher in UC_I, with both increased in inflammation as compared to health. These differences were true across multiple epithelial cell subtypes. Allowing for this general trend, we looked to identify specific genes upregulated in UC_I or CC_I (**Fig 3.2H**). In CC_I, there was upregulation of 1393 genes compared to UC_I: these encompassed antigen presentation class I MHC *HLA* antigens, interferon response genes (e.g. *ISG15, IFI6*), chemokines (*CXCL9,CXCL11*) and tryptophan metabolism genes (*WARS, IDO1*). GO pathway analysis identified that these diverse genes were downstream of increased interferon gamma and type 1 interferon responses in CC_I (**Fig 3.2I**) as compared to UC_I.

In line with previous reports, UC_NI epithelium resembled UC_I, with UC_NI retaining expression of several inflammation-associated genes (e.g., *LCN2, OLFM4, HLA-DRA*) as compared to HC (but the degree of upregulation was lower in UC_NI)(**Fig 3.2J**). As already mentioned, CC_NI was indistinguishable to health.

As the duration of inflammation for CC_I and UC_I samples was the same (**Appendix D**), these differences could not be attributed to chronicity alone. Taken together, they suggest that there may be a stereotyped epithelial response to inflammation, which exists on a continuum from Health and CC_NI through UC_NI, to UC_I and CC_I.

Stromal populations in health and disease

We enriched for lamina propria and CD45+ cells from the same samples from which we had extracted epithelial crypt cells, and subjected them to the same form of analysis. In line with previous reports[259,260], we identified four distinct fibroblast populations, endothelial, pericyte and glial cells, along with expected populations of immune cells (**Fig 3.3A**). There was also a small amount of expected epithelial carry-over given the nature of the enrichment protocol. Cells were identified by their characteristic transcriptomic profiles (**Fig 3.3B**) with recourse to previously published datasets, and all expected subtypes were recovered.

We went on to further sub-cluster stromal populations of interest.

Higher resolution data analysis (**Fig 3.3C**), revealed how previously described activated fibroblasts[20] actually comprised multiple sub clusters derived from inflamed states of parent stroma one through four. We also describe potentially novel subsets/states within stromal two and three – such as S2 NPY+ and S2 TLL2+ (**Supplementary Table 3.1**). Although their function is at present unclear, utilizing spatial transcriptomics (described later), these appear to form part of lower stromal layers. It may explain why they have not been characterized till date as endoscopy biopsies may only capture a few cells in each dataset.

We also looked at glial (**Fig 3.3D**) and endothelial (**Fig 3.3E**) populations in more detail. Glial cells demonstrated clear activation signatures in inflammation, but we observed no proportional differences between CC_I and UC_I. Similarly, we observed multiple subtypes of arterial, venous and lymphatic endothelium, but the relevance of this sub-division is unclear.

As with the epithelial fraction, there were no disease specific populations (**Fig 3.3F**) but there were clear inflammation-associated differences (**Fig 3.3G**) across all key stromal populations. A pseudobulk analysis, as for epithelium (**Fig 3.3H**) again demonstrated clear differences between inflammation and non-inflamed states. Again, mono and dual therapy CC_I appeared to cluster differently, with mono therapy being more similar to UC_I. Although CC_NI again was indistinguishable from health, in contrast to epithelial differences, UC_NI appeared to have more similarity to HC in the stromal compartment.

Differential gene expression analysis with HC identified 1193 (UC_I) and 3216 (CC_I) dysregulated genes within these populations which were highly correlated between UC and CC (**Fig 3.3I**). Comparing UC and CC inflamed samples directly, we were further able to identify 1371 genes in stromal cells

Fig 3.3 - Stromal compartment in health and disease

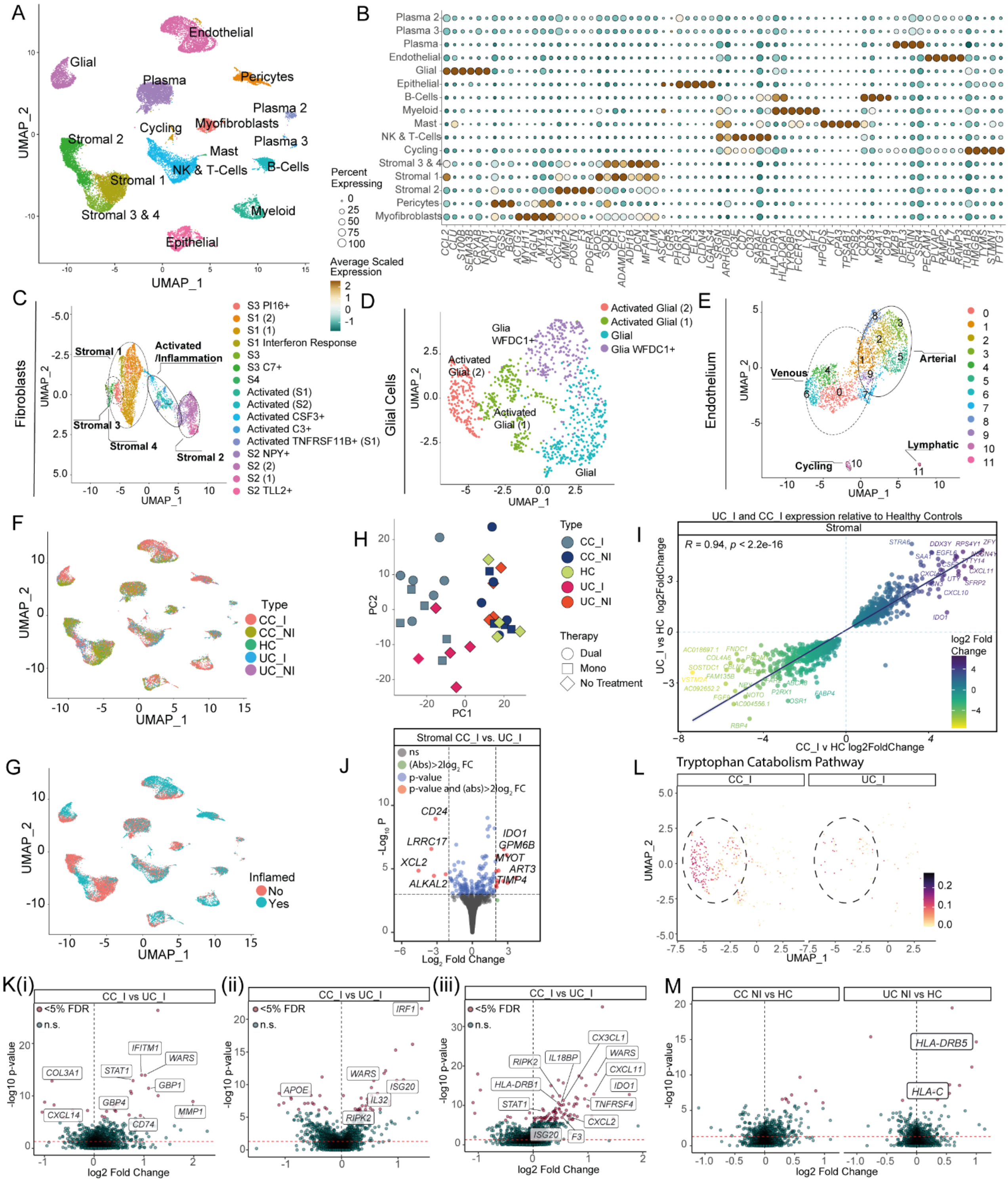


Figure 3.3 (A) UMAP stromal and CD45 subsets in health and disease. (B) Selected gene expression across stromal subsets. (C) Sub-clustering of all Stromal (1-4) cells, demonstrating the activated fibroblast phenotype, which in turn sub-clusters (D) Sub-clustering of all glial cells (E) Sub-clustering of all endothelial cells, demonstrating arterial and venous subtypes (F) Stromal/CD45 UMAP by sample type (G) Stromal/CD45 UMAP by inflammation.(H) PCA analysis of global gene expression by sample type (I) Plot of genes in stromal cells up- and down- regulated in CC_I and UC_I (J) Genes selectively enriched and depleted in CC_I and UC_I stroma (K) Plot of genes up-and down-regulated in CC_I versus UC_I in fibroblast (i), glial (ii) and endothelial (iii) cell subclusters. (L) Plot of tryptophan catabolism pathway activity in glial cell subtype. (M) Plot of genes in stroma in CC_NI and UC_NI as compared to HC, demonstrating CC_I and HC are similar. CC_I : checkpoint inflamed, CC_NI: checkpoint non inflamed, HC: healthy control, UC_I: Ulcerative colitis inflamed, UC_NI: UC non inflamed.

which were differentially regulated between these conditions (**Fig 3.3J**). As for changes in the epithelial compartment, not only were these genes shared across multiple cell types (e.g. glial, endothelium and fibroblasts **Fig 3.3Ki-iii**), but were broadly the same transcriptomic signatures and GO pathways (**Fig 3.3L**) we saw in epithelium in inflammation – chemokines (*CXCL9, CXCL11*) tryptophan catabolism pathway genes (*IDO1, WARS*) and genes downstream of interferon gamma and STAT1 signalling e.g. Class I MHC *HLA* and *ISG* genes.

Non-inflamed stromal behaviour mirrored the behaviour of non-inflamed epithelium, with minor differences between CC_I and HC. UC_NI and HC were more similar than for epithelium, but there were still some signs of residual inflammation, with an upregulation of class I MHC *HLA* genes in UC_NI (**Fig 3.3M**).

Taken together with the epithelial data, it suggests that there are common modules of inflammation activated across diverse cell populations in UC_I and CC_I. However, existence of different cell states within the same sample also suggests diversity of cellular interactions and environments driving these changes, and understanding these would be the next step in improving our understanding of the perpetuating factors in inflammation.

CD45+ immune populations in disease

We performed a similar scRNAseq analysis on CD45 cells isolated as part of our stromal protocol, in addition to supplementing cell numbers by sorting CD45+ cells from tissue biopsy specimens, and CD3+ cells sorted from tissue and blood (**Methods**). As we performed additional TCR and CITE-seq upon the T cells that we isolated, for the sake of clarity, this chapter will first describe the changes within the non-T cell compartment first before dealing with the T cell compartment.

We recovered all expected CD45 subtypes from tissue including myeloid, B cells, plasma cells, natural killer cells and a small number of mast cells (**Fig 3.4A**), in addition to T cells. Much as for the pattern of response in stroma and epithelium, the broad response in UC_I and CC_I as compared to health was the same (**Fig 3.4B**). For example, *S100A8* and *S100A9* increased in both UC_I and CC_I, which

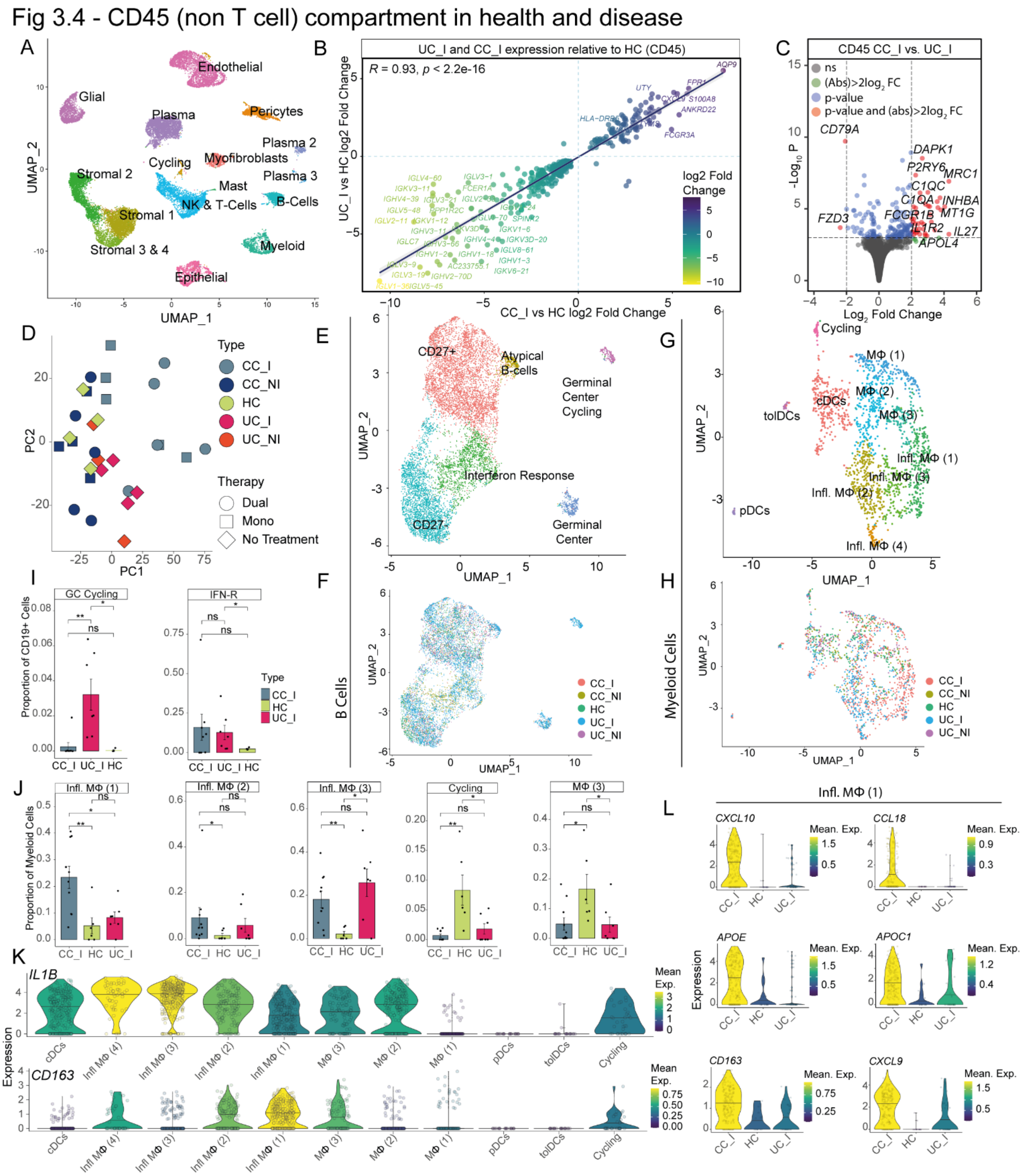


Figure 3.4 (A) UMAP stromal and CD45 subsets in health and disease. (B) Plot of genes in CD45 cells up- and down- regulated in CC_I and UC_I as compared to HC (C) Genes selectively enriched and depleted in CC_I and UC_I CD45+ cells. (D) PCA analysis of global gene expression by sample type in CD45 cells.(E) Sub-clustering of all B cells, demonstrating memory, atypical and interferon response phenotypes (F) UMAP of all B cells by disease type (G) Sub-clustering of all myeloid cells, demonstrating pro-inflammatory and non-activated macrophage and dendritic cell phenotypes (H) UMAP of all myeloid cells by disease type (I) Proportional differences of different B cell subtypes by disease, significantly different populations shown (J) Proportional differences of different myeloid cell subtypes by disease, significantly different populations shown. (K) Expression of selected genes across different myeloid clusters (L) Differential expression of selected genes in inflammatory macrophages (1) subtype between CC_I, HC and UC_I. CC_I : checkpoint inflamed, CC_NI: checkpoint non inflamed, HC: healthy control, UC_I: Ulcerative colitis inflamed, UC_NI: UC non inflamed, cDC: dendritic cells, pDC: plasmacytoid dendritic cells, tolDC: tolerogenic dendritic cells, M Φ: Macrophage. p<0.05(*), p<0.01(**), ns: not significant

suggests faecal calprotectin (heterodimer of S100A8/9) may be equally efficacious at detecting inflammation in both forms of colitis.

There were, as before, some signatures increased to a greater degree in CC_I vs UC_I (**Fig 3.4C**), of which several could be seen to be downstream of an increased interferon gamma/ TNF alpha response in CC_I (e.g. *MT1G*, *FCGR1B*)[261]

PCA analysis of the overall signature (**Fig 3.4D**) revealed similarity between UC_I and UC_NI, HC and CC_NI, with CC_I clustering separately. Interestingly, mono and dual-therapy patient CD45 responses overlapped more than their epithelial and stromal responses.

Given that we demonstrated that UC_I and CC_I were similar in many respects, we reasoned that many of these signals were likely common to inflammation, and so more disease specific processes were likely to be in differentially regulated populations and pathways, which we found were in B cells and Myeloid clusters, the key findings from which we present here.

B cells sub-clustered into clearly defined previously described populations (**Fig 3.4E**) with changes that were particular to inflammation (**Fig 3.4F**). For example, although interferon-response B cells were

enriched in both CC_I and UC_I (**Fig 3.4I**), germinal centre and germinal centre cycling cells were particularly enriched in UC_I.

The differences in inflammation were more marked when considering the myeloid populations. Using previously published datasets[59], we were able to identify multiple subtypes, including dendritic cells, resting and inflammation associated macrophages (**Fig 3.4G**, **Supplementary Table 3.1**). There appeared to be condition specific enrichments (**Fig 3.4H**), with Inflammatory Macrophages subtypes 1-3 enriched in CC_I compared to health, with Inflammatory Macrophages subtype 1 (Infl. MΦ(1)) significantly increased in CC_I compared to UC_I (**Fig 3.4J**). In contrast, MΦ(3) and Cycling Macrophages were relatively increased in health compared to both forms of colitis.

Although the division between Type 1 and 2 macrophages does not hold as clearly in humans as it does in mice, Infl. MΦ(1) appeared to express the M2 (anti-inflammatory) marker *CD163* whereas *IL1B*, a classic marker of M1 macrophages was more highly expressed in Infl. MΦ(3) macrophages (**Fig 3.4K**)[262–264].

When compared across UC_I, HC and CC_I, there were also condition specific differences within cell subtypes. The most marked difference could be seen in the Infl. MΦ(1) subset (**Fig 3.4L**), which in addition to being proportionally enriched in CC_I, also expressed higher levels of M2 anti-inflammatory markers - *CD163* and *CCL18*, along with more unconventional markers such as *APOE,* resembling macrophages that have been described in the context of fat metabolism. They also, unusually, expressed some of the same chemokine markers (*CXCL9, CXCL10*) we identified as being enriched in CC_I epithelium and stroma.

T cell populations in health and disease

Given T cells are the populations that express immunotherapy targets PD-1 and CTLA-4, we went on to examine the T cell populations in health and disease in more detail, looking both at blood and colonic tissue. Combined across all isolation modalities, we recovered 41,144 CD3+ cells in biopsies and 36,176 from PBMCs.

We describe all of the commonly understood subsets in tissue and blood, including rarer populations such as gamma delta cells and circulating MAIT cells (**Fig 3.1D,E**). Combining transcriptomic with CITE-seq data (**Fig 3.5B**), allowed us annotate exhausted cells, resident cells and naïve cells with confidence (**Supplementary Table 3.2, Table 3.3**).

As with the changes in epithelium and stroma, there were marked differences in inflammation, but this was more marked in tissue, and the changes in blood were less significant (**Fig 3.5A, Fig 3.5D**). We could also see reactive changes common to both disease processes, such as Tregs increased in UC and CC tissue as compared to health, while Th1 and Gamma delta cells were both relatively depleted.

Fig 3.5 - T cell compartment in health and disease

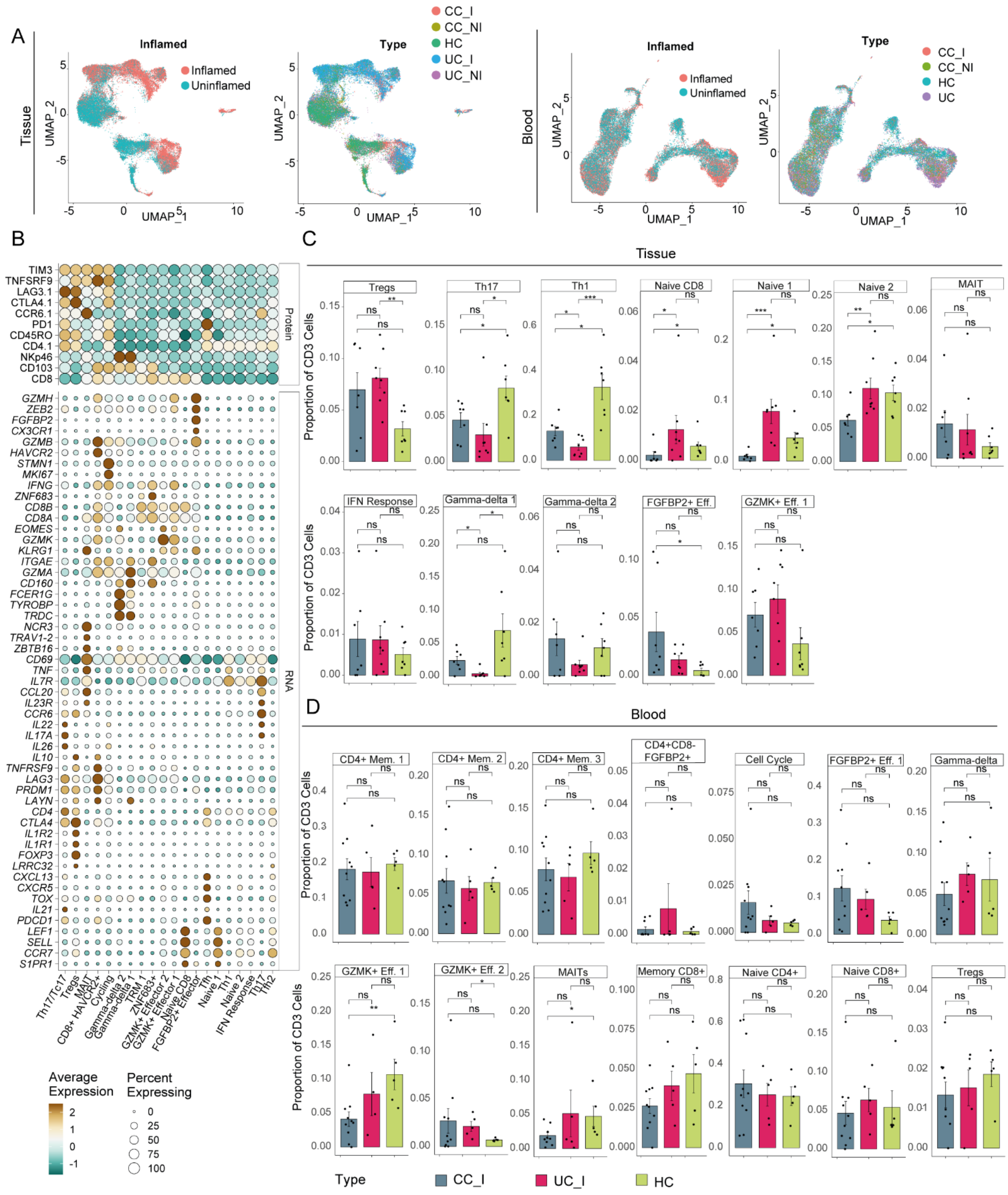


Figure 3.5 (A) UMAP embedding of T cells in health and disease as broken up by compartment - tissue or blood derived, as well as by disease state and by dataset (B) Expression of selected genes and CITE-seq antibody markers across all T cell subsets described here. (C) Proportional differences in tissue between CC_I, UC_I and HC broken down by subset of T cell (D) Proportional differences in blood between CC_I, UC_I and HC broken down by T cell subset. p<0.05(*), p<0.01(**), p<0.001(***), Statistics: Unpaired T test ns: not significant. CC_I: Checkpoint inflamed, CC_NI: Checkpoint non inflamed, UC_I: Ulcerative colitis inflamed, UC_NI: Ulcerative colitis non inflamed, HC: Healthy control, Eff: Effector T cell, Mem: Memory T cell. GZMK+: Granzyme K positive, MAIT: Mucosal associated invariant T cells, ZNF683+: HOBIT/ZNF683+ T cells, HAVCR2: Exhausted T cells.

Interestingly, there were condition specific changes in tissue in checkpoint inhibitor induced colitis and ulcerative colitis. We explore these changes in more detail, including making comparisons with T cells recovered from epithelium

TRM cells play a key role in checkpoint inhibitor-induced colitis

The method of isolation that we employed to enrich for epithelium (**Fig 3.6A, Methods**) also allowed us to selectively enrich for immune cells that were found in close association with these cells – and these were found to encompass T cells, B cells, Myeloid and Mast cells (**Fig 3.2A**). The residual fraction, by contrast, was relatively enriched for immune cells associated with the lamina propria.

Certain populations such as *IL17* expressing CD4 and CD8 T cells were enriched in inflammation as a whole (i.e. CC_I and UC_I) as compared to health (**Fig 3.6B**). However, there were disease specific enrichments, such as T-follicular helper cells (subtypes 1 and 2) as well as *GZMK* effector cells that were enriched in UC_I. In contrast, exhausted (*HAVCR2+*), resident cell (TRM1 and ZNF683+) and cycling populations were more enriched in CC_I, suggesting different driver populations.

We looked at cycling populations in more detail as these were actively stimulated populations (**Fig 3.6C**), and could see that these comprised Tc17 populations in UC_I whereas were derived from ZNF683+ (i.e. resident) populations in CC_I.

Gratifyingly, as we described in our earlier results (**Chapter 2**), in this independent dataset, *IL26* – expressing CD8 Tc17 cells were enriched in UC_I. Moreover, this was specific to the UC inflammatory disease process and not shared with CC_I (**Fig 3.6E**)

When we looked at intraepithelial populations in more detail (**Fig 3.6Di**), we could see that they comprised a higher proportion of the HOBIT/ZNF683+ population in CC_I (whereas the UC_I specific Tc17 population proportion was similar to HC). In addition, these CD8 T cells appeared to be activated and terminally differentiated (**Fig 3.6Dii**), expressing higher levels of *GZMH* and *ZEB2* than in either UC_I or HC. Taken together, it suggested that the ZNF683+ population was stimulated, actively

Fig 3.6 - The role of TRM cells in checkpoint inhibitor-induced colitis

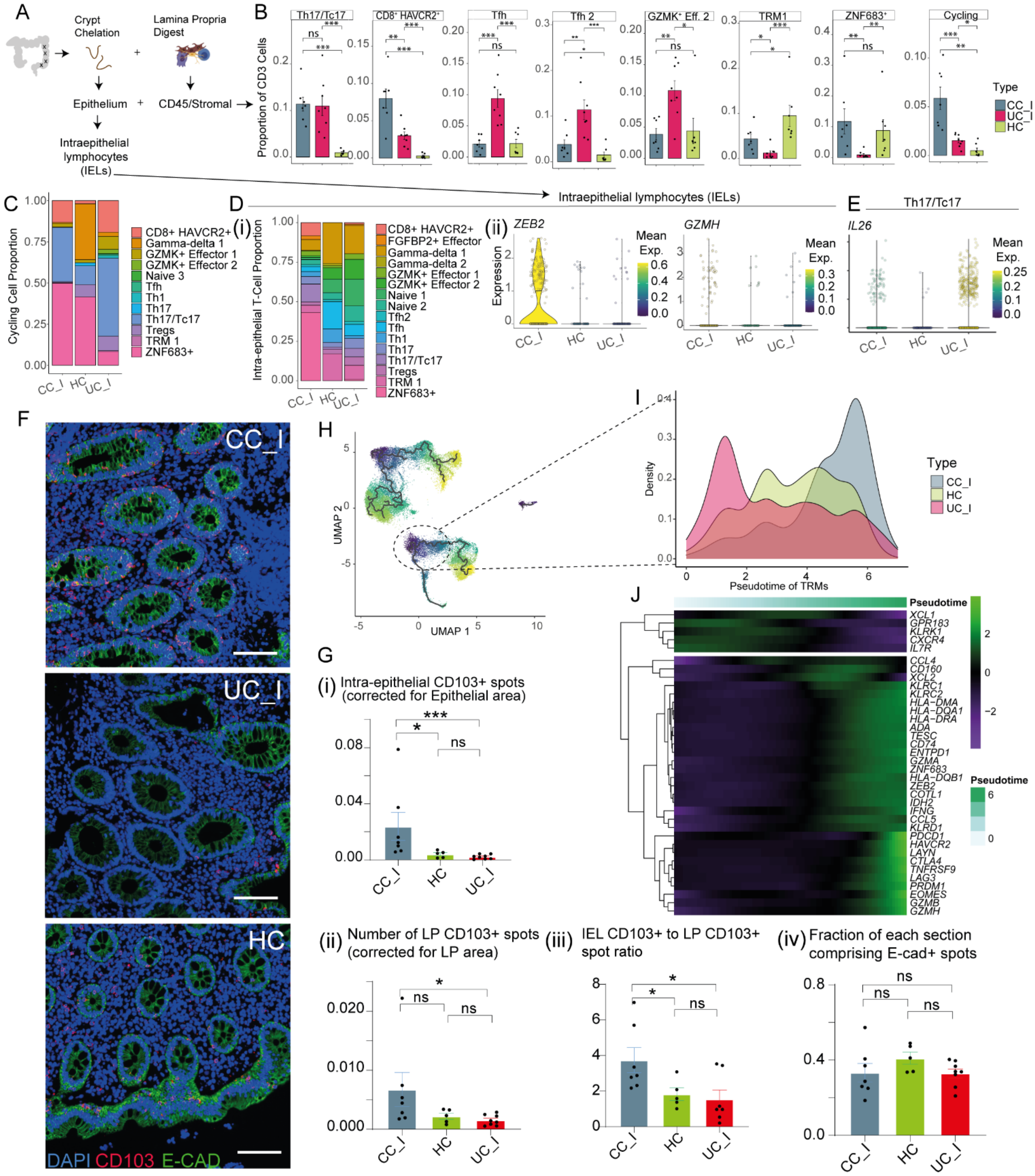


Figure 3.6 (A) Experiment schematic - derivaion of different CD45 populations, enriching for lamina propria derived and epithelial derived cells (B) Proportional differences of key populations between CC_I, UC_I and HC (C) Cycling T cells by cell type and disease condition (D) Intraepithelial T lymphocytes (i) by disease state and cell type and (ii) by selected gene expression across health and disease. (E) Expression of IL26 in Tc17 cells across health and disease, demonstrating an increase in inflammation, particularly in UC_I. (F) Representative immunofluorescence images of resident (CD103+), epithelial (E-cadherin,ECAD) and nuclear (DAPI) co-staining in health and disease. White bar = 100uM (I) Quantification of signals from immunofluorescence images from cohort of patients demonstrating a CC_I specific increase in CD103+ cells in (i) epithelium (ii) lamina propria relative to HC and UC_I. (iii) Epithelial CD103 enrichment outweighs lamnia propria enrichment in CC_I (iv) Control epithelial content is the same in inflammation and health for these sections. (H) Superimposition of pseudotime on UMAP of all tissue T cells (I) Expanded view of pseudotime of TRM population by disease type (J) Expression of selected genes along pseudotime axis of TRMs. p≤0.05(*), p<0.01(**), p<0.001(***), ns: not significant. Statistics for (G): Unpaired T test with Kolmogorov-Smirnov comparison of cumulative distributions for non-parametric data. CC_I: Inflamed checkpoint inhibitor-induced colitis, UC_I: Inflamed ulcerative colitis, HC: Healthy control, Eff: Effector T cell, Mem: Memory T cell. GZMK+: Granzyme K positive, MAIT: Mucosal associated invariant T cells, ZNF683+: HOBIT/ZNF683+ T cells, HAVCR2: Exhausted T cells. TRM: Tissue resident memory. IEL: Intraepithelial lymphocytes, LP: Lamina propria. E-cad: E-cadherin.

expanding, exerting a cytotoxic function i.e. likely contributing significantly to the inflammation; moreover, this population had migrated into the epithelium in CC_I.

To confirm our novel finding of intraepithelial lymphocytosis being a key feature in CC_I, we performed immunofluorescence (**Methods**) on an independent cohort of patients. We utilized CD103 as an antibody marker of ZNF683+ cells (**Fig 3.5B**), E-cad as a marker for epithelium and DAPI for nuclear staining. We observed marked differences in the absolute enrichment of resident CD103+ populations in CC_I, validating our findings by single-cell RNA sequencing (**Fig 3.6F**), and were able to quantify this (**Methods**) to show that resident cells were not only absolutely more abundant in the epithelium (**Fig 3.6Gi**) and lamina propria (**Fig 3.6Gii**), but the ratio of intra-epithelial to lamina propria cells was significantly higher in CC_I (**Fig 3.6Giii**). As control, we confirmed that the proportion of epithelium included in each sample was the same (**Fig 3.6Giv**).

In order to understand the *ZNF683*/HOBIT population better, we performed pseudotime analysis on all tissue derived T cells (**Fig 3.6H**). HOBIT+ cells lay on a continuum between TRM cells and exhausted CD8+ HAVCR2+ populations, supporting their role as a terminally differentiated activated TRM population. When looking at the differences in TRM and ZNF683+ populations across disease states in more detail (**Fig 3.6I**, **J**), we could see that activated *PDCD1, CTLA4* positive *IL7R* negative HOBIT cells were dominant in CC_I. This population expressed multiple markers of activation such as multiple granzymes (*GZMA, GZMB, GZMH*) as well as interferon (*IFNG*), further support of their cytotoxic role in CC_I.

T cell receptor analysis reveals trafficking, amplification and antigen specificity patterns.

We reconstructed TCR repertoires from matched single cell VDJ sequencing data, yielding 40,745 clones across tissue and blood, within the above described cell phenotypes. Using Morisita's index (**Methods**), which takes into account variations in the overall size of repertoires between sub-populations, we calculated the relative strength of pairwise clonal repertoire overlaps for all clusters in blood and tissue within each condition on a sample by sample basis. This allowed us to append clonal sharing network information for both CD4+ and CD8+ populations in blood and tissue, summarized over **Fig 3.7A** and **Supplementary Fig 3.2A**. Taken together, it allowed us to assess local expansions, clonality, phenotype dynamics and trafficking between blood and tissue. CD4 T cell populations comprised largely single clones so inferring relationships was difficult, despite recovering 36,000 cells from PBMCs and 40,000 CD3+ T cells. CD4 data is presented in **Supplementary Fig 3.2A** and we deal with the conclusions from CD8 T cell repertoire analysis in more detail.

In health (**Fig 3.7Ai**), the clusters from tissue and blood generally formed their own communities in and any clone trafficking between blood and tissue cells was secondary. The majority of clone trafficking in health was accounted for by sharing between GZMK+ effector populations and MAIT populations between blood and tissue. In tissue, TRM populations showed the greatest clonal overlap with ZNF683+ cells, demonstrating the continuity of reactivation between these two populations. In blood, clonal overlaps were observed between GZMK-expressing and FGFBP2+-expressing effector cells

In UC (**Fig 3.7Aii**), also summarised as a dendrogram, we observed different dynamics, where the trafficking of all effector populations was substantially increased. There was also an increase between tissue and blood MAIT cell trafficking in comparison to HC. In line with previous reports, exhausted, cycling and Tc17 (also *IL26* expressing) populations were closely linked[2], and shared few clones with circulating cells, suggesting the expansion of Tc17 cells in UC is restricted to tissue.

In CC (**Fig 3.7Aiii**), whilst we also observed an increase in trafficking between effector populations in blood and tissue when compared to HC, this effect was less than seen in UC. Within tissue, exhausted, cycling, TRM and ZNF683+ cells formed their own cluster, with few blood-shared clones, suggesting these cells were forming a local tissue response. In contrast to UC, Tc17 cells were not closely clonally linked to this group and the expanded activated ZNF683+ population was central instead.

Fig 3.7 - T cell receptor analysis in disease and health

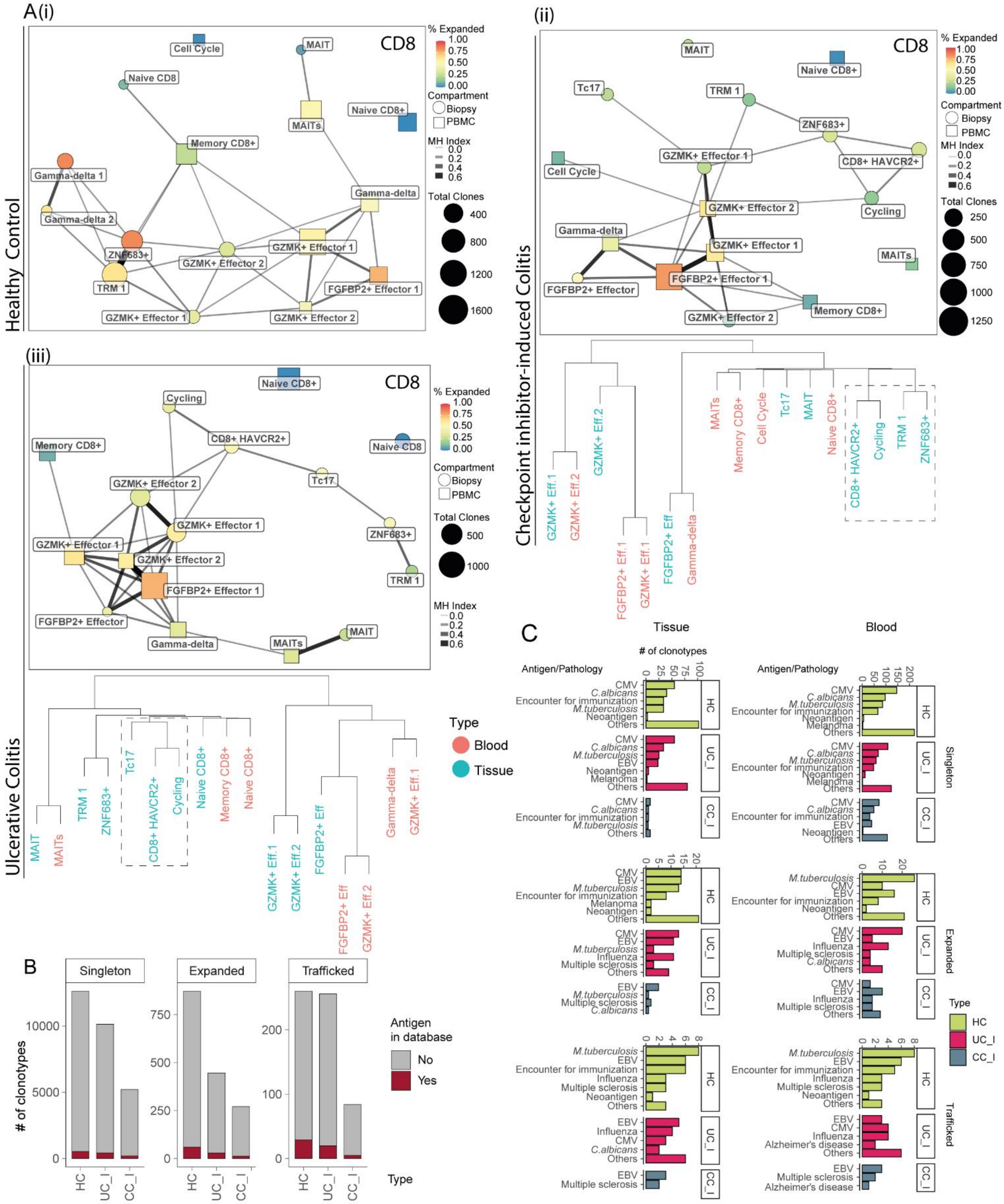


Figure 3.7 (A) Analysis of patterns of CD8 T cell trafficking in health and disease. Combined depiction of size of clonotypes, degree of sharing and degree of expansion. For CC_I (ii) and UC_I (iii), further dendrogram highlighting sharing in between certain clones - Tc17, exhausted and cycling cells in UC and exhausted, TRM, ZNF683 and cycling cells in CC. (B) Comparison of TCR sequences recovered against public databases - the majority of TCR sequences are against unknown antigens (C) Depiction of TCRs for known antigens recovered from the current dataset. CC_I: Checkpoint inflamed, UC_I: Ulcerative colitis inflamed, HC: Healthy control, Eff: Effector T cell, Mem: Memory T cell. GZMK+: Granzyme K positive, MAIT: Mucosal associated invariant T cells, ZNF683+: HOBIT/ZNF683+ T cells, HAVCR2: Exhausted T cells. TRM: Tissue resident memory.

Taken together, this suggests that in CC and UC, both trafficked and locally expanded T-cells contribute to mucosal inflammation; however, in UC T-cell trafficking is substantially increased and in line with relative depletion of TRM and enrichment of effector populations in tissue. In UC, Tc17 and exhausted cells are central to the network of locally proliferating resident populations, while in CC reactivated TRM and exhausted cells are more prominent. Together with the increase in cycling and reactivated TRMs in CC samples, this suggests that while both trafficked and locally expanded T-cells play a role in both UC and CC, local T-cell expansions not only contribute more to inflammation in CC than UC but also involve phenotypically different T-cell populations.

Cross-referencing public TCR databases (VDJdb, McPAS, TBAdb)[199,200], we identified 1,263 unique clonotypes (exact TRB matches) potentially linked to 47 known antigens or pathology in our datasets. This would suggest that the majority of T-cell responses in all conditions are due to antigens not well captured in public databases (**Fig 3.7B**). The public TCRs were largely restricted to non-expanded and non-trafficked i.e. singleton T-cells.

The antigens that we were able to identify likely represent bystander clones (**Fig 3.7C**) in tissue and blood rather than driver populations given that there are no obvious differences between HC, UC_I and CC_I. Notably, we did not observe expansion of T-cell clones previously associated with cancer neo-antigens or melanoma in CC, although this mechanism cannot be ruled out due limitations in available data and private TCR repertoires, as also highlighted in **Chapter 2**.

In summary, when reviewing our findings across epithelium, stromal and immune populations, we observed consistent differences between CC_I and UC_I (along with features common to both). Many of these transcriptional signals were maintained across populations as diverse as crypt cells, fibroblasts, glial and endothelial cells. We wanted to understand cellular organisation of these elements in an unbiased manner and interrogate interactions across disease states, differences in which could yield mechanistic insights into how such diverse responses were emerging.

Unbiased spatial transcriptomics highlights cellular associations in disease and health

In order to further our understanding of cellular interaction, organisation and mechanism, we undertook spatial transcriptomics (ST) of 18 tissue sections (**Fig 3.8A**, **Methods**) from HC, CC_I and UC_I biopsy and resection samples after optimization of the protocol for colonic tissue. The Visium platform yields an unbiased transcriptomic snapshot paired with conventional haematoxylin and eosin (H&E) staining information (**Fig 3.8B**), at a resolution where each 'spot' comprises the conglomerate signal from 5-10 cells. This allowed us to corroborate transcriptomic features corresponded to tissue structures as assessed by a consultant histopathologist.

Across 18 sections (**Fig 3.8B, Fig Supp 3.1A**), following quality control (**Methods**), we retained 25,672 good quality, ST spots for further analysis. Integrative clustering analysis (**Methods**, **Fig 3.8C**) revealed 20 broad spatial regions, which we labelled according to anatomical regions, transcriptomic signatures or representative cell type enrichment, utilizing our single cell atlas as reference (**Fig 3.8D, Fig 3.8E**). We could therefore resolve regions into the co-localised signals of different cell compositions –for example we observed enrichment of stem cells (and correspondingly *SPINK* and *LGR5*) with transit-amplifying (TA) cells (corresponding to deep crypt regions in the mucosa, myofibroblast (and correspondingly collagen *COL1A1* and myosin *MYL9*) enrichment in muscularis mucosa, arterial and venous endothelium, together with pericytes and Stromal 3 (S3) type vascular niche fibroblasts in vascular and peri-vascular spots, while diverse immune populations were most strongly enriched in follicle/tertiary lymphoid structures (**Fig 3.8D, Fig 3.8E**). The majority of the identified ST regions were present in all of the sections, with a bias for submucosal regions being over-represented in the resection samples given increased sampling depth achievable surgically. CC_I samples were represented as biopsies as no patients underwent surgical treatment of their colitis over the duration of the research period.

Interestingly it should be noted that in our ST dataset, we were able to localise cell type signatures that were not present within our scRNA-seq reference dataset, likely due to technical limitations in

Fig 3.8 - Spatial visualization of cellular behaviour in disease and health

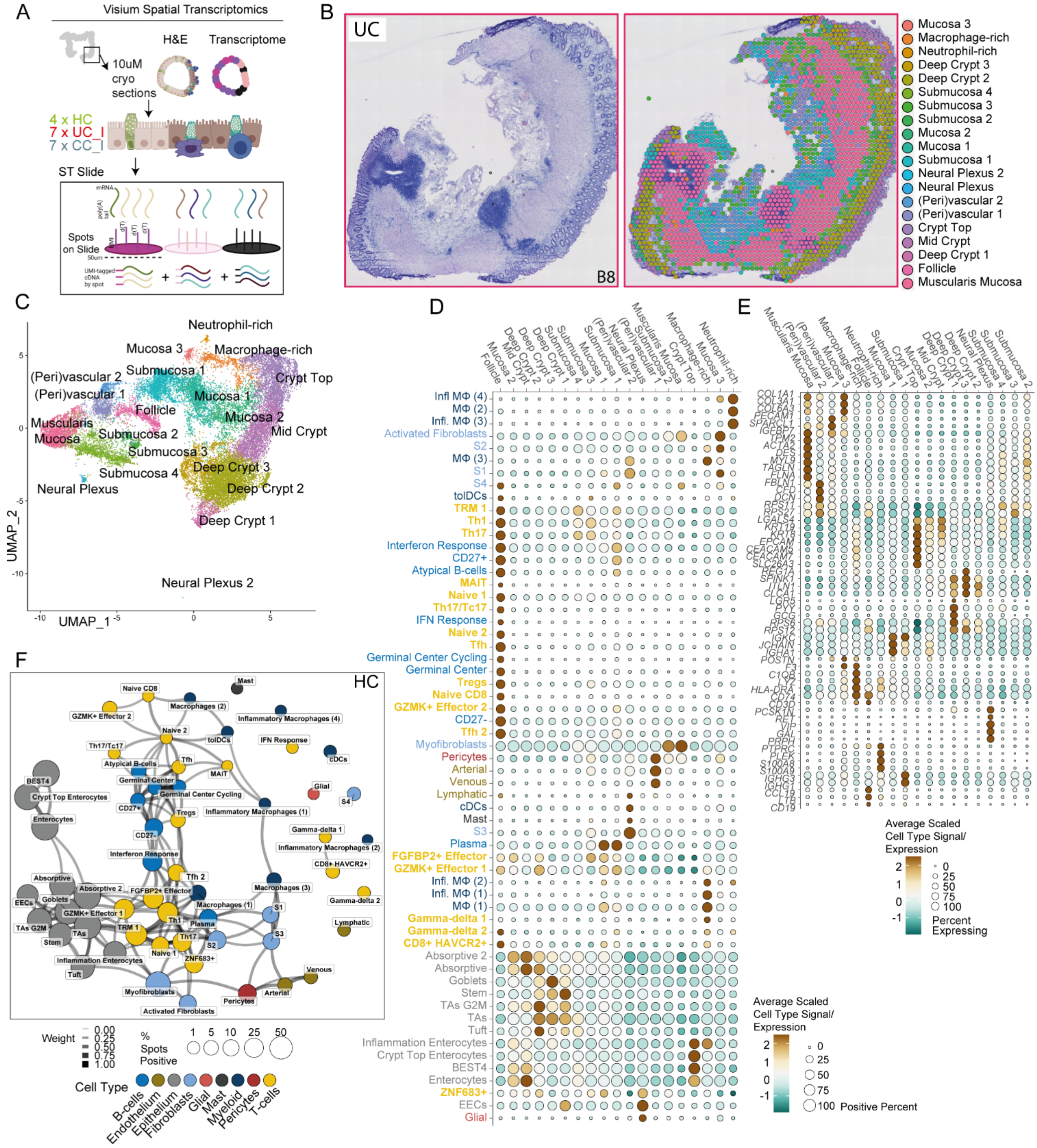


Figure 3.8 (A) Experimental overview and brief schematic representation of 10X Visium spatial transcriptomic technology (B) Representative section of UC_I, matched H&E image with transcriptome-identified spot clusters. ST spots annoted with reference to published genes (section ID: B8, all sections in supplemental figure, high resolution paired H&E in supplementary data) (C) UMAP embedding of spot clusters across all sections as derived from unbiased spatial transcriptomics, annotated according to tissue location/structure or cell type enrichment (D) Dot plot heatmap for cell type enrichment and distribution across all spatial regions identified in (B and C) (E) Dot plot heatmap of top gene expression markers for each spatial region cluster identified by ST, represented in (B) &(C)(F) Analysis of cellular interactions in HC, derived from spatial transcriptomic data. Pairwise analysis visualising cell populations (as determined by sc-RNA seq) most strongly co-occurring within the same spatial locations. Edge width and intensity represents correlation between populations. Edges below r < 0.15 are not shown for clarity. Nodes represent cell types, coloured by broad compartment. Node size represents the percentage of all spots with cell type signal > 0. CC/CC_I: Inflamed checkpoint inhibitor-induced colitis, UC/UC_I: Inflamed ulcerative colitis, HC: Healthy control.

droplet encapsulation due to cell size, shape or RNA content (e.g. neurons – Neural plexus regions characterized by *RET/VIP*, neutrophils – neutrophil-rich spots characterized by *S100A8/S100A9*); or due to sampling/cell rarity (e.g. follicular dendritic cells – follicle regions, Paneth cells – deep crypt 3) (**Fig 3.8E**). This enabled us to extend our understanding of rare cell populations as well as identify relationships that we could not observe by scRNA-seq (such as the association of neutrophils with inflammatory macrophages subtypes 2, 3 and 4).

Cell-cell signal correlation analysis (**Methods**) revealed the co-occurrence of cell types that correspond to colonic tissue architecture, pooled across all samples of that type. We present this diagrammatically as a linked bubble plot, where the size of the bubble denotes the area covered by a cell type (i.e. its relative abundance), with the degree to which cell types co-occur denoted by the weight of the link between two bubbles. In health (**Fig 3.8F**) we can see for example that crypt top cells, such as mature enterocytes and BEST4 cells clustered together, while deep crypt cells like goblets and TA signatures were also highly correlated. We can observe significant co-occurrence of endothelial and pericyte populations together with S3 peri-vascular fibroblasts, while immune populations such as follicular T-cells and B-cells co-localise together within lymphoid structures. This confirmed known cellular relationships and structures in homeostasis, and was a novel method of holistically assessing cellular relationships in an unbiased way, providing a baseline for assessing changes in UC_I and CC_I.

Spatial transcriptomics highlights disease-specific changes in epithelial cellular interaction

A similar analysis of all UC_I and CC_I demonstrated a significantly altered correlation structure (**Fig 3.9A,B**). In UC_I, cells of a particular type such as T cells and B cells, tend to cluster together. This would suggest that the immune response organises itself in more stereotyped patterns in tissue in UC_I, likely in lymphoid structures clearly delineated from the epithelium.

In CC_I, in contrast, we observed a strong co-localisation of several T-cell subtypes with epithelial populations, as well as increased co-occurrence of macrophages and T-cells within the same spots

(**Fig 3.9A**, highlighted in **Fig 3.9B**). This was the first indication that the major alteration in overall spatial organisation of colonic tissue in CC is T-cell and macrophage infiltration in the mucosa.

We also observed Treg cells in close association with tissue macrophages, a feature particular to CC_I (**Fig 3.9A**). Given that from our single cell data we could see there was a predominance of M2 macrophages in checkpoint colitis, this would explain why there was an abundance of Treg cells in close association, as M2 macrophages are known to polarise cells into becoming a Treg phenotype. This would also be supported by the increase in tryptophan catabolism pathway (*IDO1, WARS,* **Fig 3.2H, 3.3J,L,K**) that we saw in CC_I. The tryptophan catabolism pathway is again known to polarise cells towards a Treg type. It is likely however that these T cells were dysfunctional given that they mediate their effects through PD1 and CTLA4 – and either or both of these receptors were pharmacologically inhibited in these patients.

Given that multiple lines of inquiry had highlighted that the crypt-top epithelium was behaving in a unique manner in CC_I, starting from proportionality differences (**Fig 3.2E**) through to differences in T cell interaction (**Fig 3.9A**) as seen above in ST, we decided we decided to look at the crypt-top specifically across disease and health. In order to do this, we developed a bioinformatic crypt axis score that was capable of identifying whether epithelial cells were from crypt base or top based upon their transcriptional signature[24], and applied it to our samples in ST. Utilizing this, we were able to see clear gradients in resection sections, and could use this to orientate biopsies where such differences were less clear (**Fig 3.9C**).

Across disease and health, ST crypt top spots clustered into clear healthy and inflammation - associated regions (**Fig 3.9D**). We were able to assess not only the extent to which different regions were expanded, but how they co-localised. Here again, we saw clear differences between UC_I and CC_I, representing cellular interaction differences which we could not appreciate by sc-RNA seq. INF-4 and INF-3 predominated in UC_I, whereas checkpoint inhibitor colitis was characterised by INF-1 and, in particular, INF-2 (**Fig 3.9E,F**). We could also see differences in organisation, e.g. whereas INF-2

Fig 3.9 - Utilizing spatial transcriptomics to interrogate cellular interaction

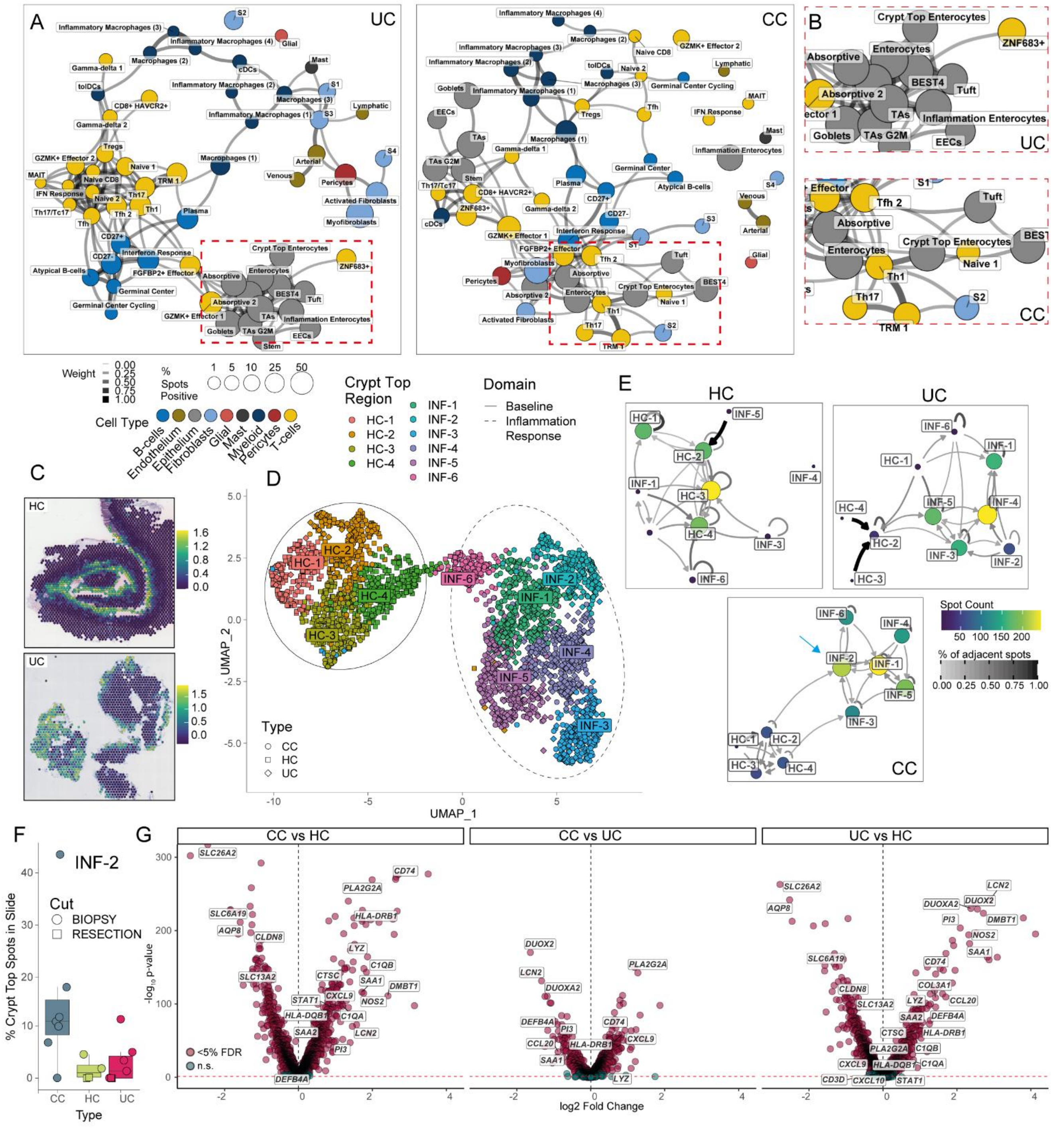


Figure 3.9 (A) Analysis of cellular interactions in UC_I and CC_I, derived from spatial transcriptomic data. Pairwise analysis visualising cell populations (as determined by sc-RNA seq) most strongly co-occurring within the same spatial locations. Edge width and intensity represents correlation between populations. Edges below r < 0.15 are not shown for clarity. Nodes represent cell types, coloured by broad compartment. Node size represents the percentage of all spots with cell type signal > 0 (B) Expansion of highlighted area from (A) demonstrating increased (and differential) interactions between crypt cells and T cells in CC_I vs UC_I. (C) Epithelial crypt axis score as developed by scRNA-seq, distribution across ST spots in a longitudinal cut of a healthy colonic resection (top) and a cross section of a UC biopsy sample (bottom). High intensity = crypt top. (D) UMAP embedding of crypt top spots (crypt axis score >= 1.0) as transcriptional clusters. The ellipses highlight spots from healthy control and inflamed sample (CC and UC) areas cluster together , computed at 95% confidence intervals with a multivariate t-distribution. (E) Cluster network graph visualising degree of spatial adjacency between ST spots from crypt top domains shown in (D). Edge colour and arrow width between two nodes represents the mean fraction of all directly spatially adjacent spots of that type, node size and colour represent the total number of ST spots of a given type. Inter-relationships for all spots in HC, UC_I and CC_I sections is shown. (F) Boxplot visualising the proportion of pro-repair INF-2 crypt top domain as a fraction of all crypt-top spots in ST resection and biopsy sections for CC_I, HC and UC_I. (G) Volcano plots of differentially expressed genes in mucosal crypt top ST spots between all conditions, with degree of log-fold change along the x axis. Selected genes are labelled. CC/CC_I: Inflamed checkpoint inhibitor-induced colitis, UC/UC_I: Inflamed ulcerative colitis, HC: Healthy control.

associated with inflammation-associated regions only in UC_I, it was bordered by health-associated regions in CC_I (**Fig 3.9E**).

When then assessing the transcriptomic changes in crypt tops across disease and health, we saw many of the same changes as we had seen using sc-RNA seq, in a validation of that data in this independent cohort of patients (**Fig 3.9G**). Notably, we saw enrichment of a number of chemokine molecules, including *CXCL9*, *CXCL10* and *CXCL11*. We also saw enrichment of epithelial repair promoters such as *CD74*[265]. The majority of gene signatures, such as upregulation of *HLA* molecules, was downstream and consequent of excess interferon gamma signalling in CC_I. In contrast, in UC_I, we saw a relative enrichment of *CCL20* which is known to be secreted by S4 as part of its interaction with and driving the creation of lymphoid follicular structures.

Epithelial responses drive accumulation of immune populations

We utilized our single cell data to impute cell types that were present in crypt-top regions (**Fig 3.10A**). We could see that INF-3 and INF-4 were enriched for inflammation-associated enterocytes. In contrast, INF-2 and INF-1 comprised of myofibroblasts, stem and transit-amplifying cells, along with enrichment for ZNF683+ cells and exhausted CD8 T cells or plasma cells and Tfh 1/2 respectively. The inflammatory macrophages (Infl Mphage (1)) that were characteristic of CC_I were clustered together with Tregs and ZNF683+ cells in INF-2, additional validation of their interaction and processes that led to their perpetuation.

When we superimposed GO pathway analysis on inflammatory domains, we confirmed that there was an increase of interferon gamma signalling across INF-2, 3 and 4, accompanied by increases in antigen presentation, with reflex increases in macrophage signalling and tissue repair processes in INF-2 (**Fig 3.10B**). We could relate this mechanistically to the composition of these micro-domains. For example, INF-2 areas, promoting epithelial regeneration and repair, although enriched for many activated populations such as GZMK+ effector cells, ZNF683 and Tc17 cells, was also enriched for regeneration-promoting S2 fibroblasts and CD163-bearing Inf Mphages. INF-6, by contrast, was enriched for

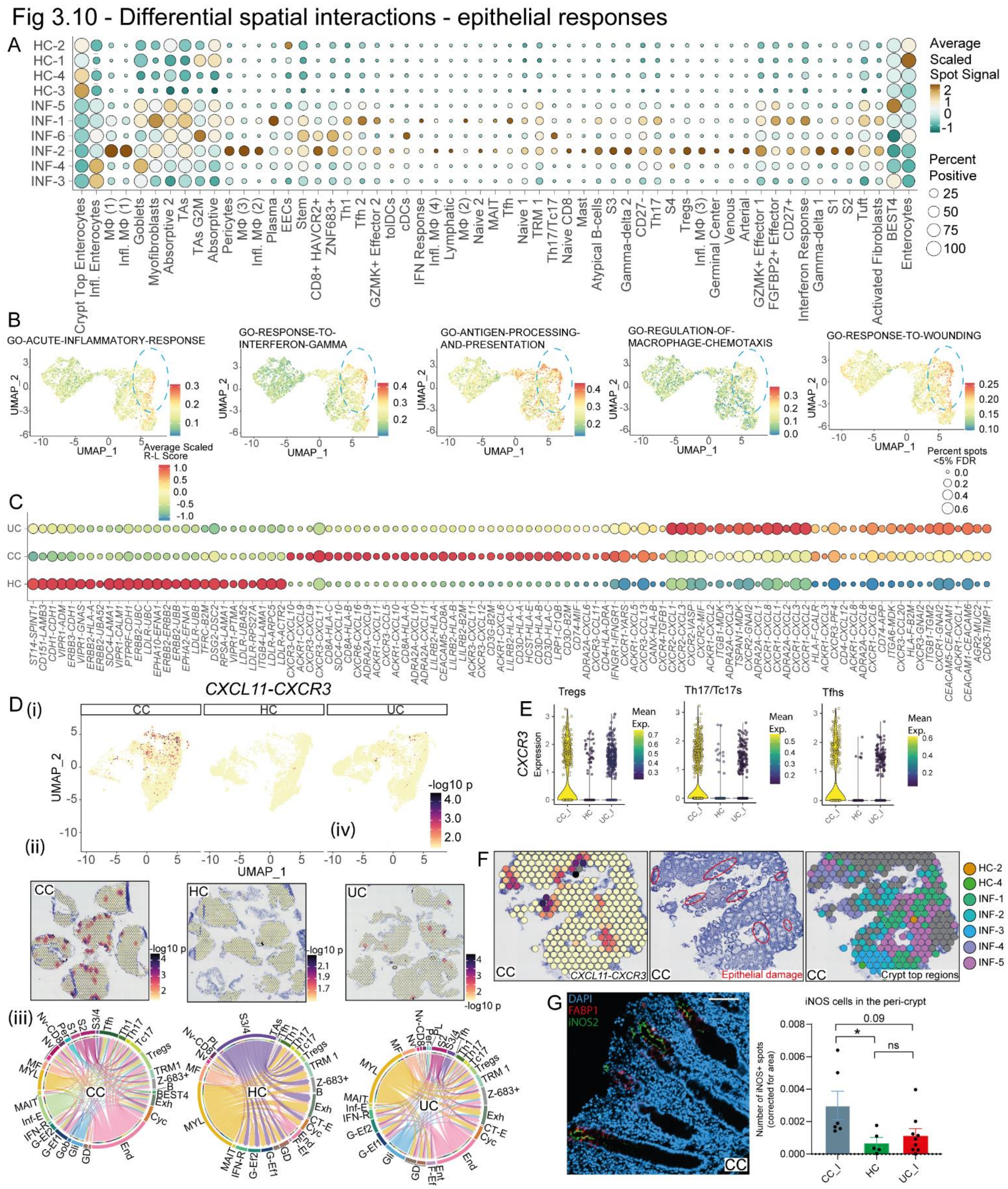


Figure 3.10 (A) Dotplot heatmap visualising cell type signal distribution enrichment in crypt top regions by ST. (B) Selected GO pathways superimposed on UMAP of crypt top spots in all ST slides; intensity corresponds to pathway activity AUC score. Blue dashed line: INF-2 (C) Dotplot heatmap of top condition-specific spatial receptor-ligand interactions in crypt top ST spots across all slides (D) Receptor ligand analysis of spatial and cell type distribution of CXCL11-CXCR3 signalling in CC_I, HC and UC_I. (i) UMAP overlay of spatial distribution of significant co-localisation of CXCL11-CXCR3 spots using ST (ii) Representative ST slides showing CXCL11-CXCR3 spots with significant co-localisation (iii) Circos plots from sc-RNA seq data showing source (ligand) and target (receptor) expressing cell populations in health and disease. Coord width is scaled to expression level. (E) Violin plots of CXCR3 expression in selected T cell populations in disease and health (F) CXCL11-CXCR3 receptor-ligand signalling (Left), paired H&E imaging (Middle) and crypt-top characterization by ST (Right) of representative sections of CC_I. (G) Immunofluorescence staining of induced nitric oxide synthetase 2 (iNOS2), crypt tops (FABP1 high) and cell nuclei (DAPI) for a representative CC section; analysis of iNOS2 expression by IF across a cohort of patients with CC_I, UC_I and health. Scale bar = 100uM. T test, $p<0.05$ (*), ns = not significant. Abbreviations: Gli – Glial cells. GD- gamma-delta cells. F-Ef. – FGFBP2+ effector T-cells, Ent – Enterocytes, End – Endothelial, Cyc – Cycling cells, Exh – CD8+ HAVCR2+ T-cells, B-B-cells, Z-683+ - ZNF683+ TRMs, S3/S4 – Stromal 3 & 4, S2 – Stromal 2, S1 – Stromal 1, Gob – Goblets, Per – pericytes, Nv CD8 – naïve CD8+ T-cells, Nv – naive CD4+ T-cells, MF – myofibroblasts, MYL – myeloid cells, IFN-R – Interferon response B-cells, G-Ef1 – GZMK+ Effector T-cells 1, G-Ef2 – GZMK+ Effector T-cells 2, CT-E – crypt top enterocytes, TAs – transit amplifying epithelium, PL- plasma cells, EEC- enteroendocrine cells, Abs and Abs2 – Absorptive Epithelial Cells 1 and 2. CC_I/CC: Inflamed checkpoint inhibitor-induced colitis, UC_I/UC: Inflamed ulcerative colitis, HC: Healthy control

activated ZNF-683 and Tc17 cells alone (**Fig 3.10A**) , and was associated with a lower regeneration and repair signal (**Fig 3.10B**)

Taken together with the inter-relatedness of these areas (**Fig 3.9E**), we could see that in UC_I, inflammatory INF-1, INF-5 and INF-6 clusters, rich in immune cells, were being supported by INF4 clusters, whereas in CC_I, INF-2 was also present to ameliorate the effects of adjacent inflammation. This may explain why the epithelial damage we observe is classically milder in CC_I than seen in UC_I.

In order to interrogate cellular interactions in these regions in more detail to see if we could derive mechanistic insights, we performed ligand-receptor pair analysis (**Methods**) across disease and health for our data from spatial transcriptomics of tissue. The results of the most significant differences (**Fig 3.10C**) showed a notable increase in interferon gamma and HLA-CD3 signalling in CC_I, accompanied by an increase in signalling via *CXCL9*, *CXCL10* and *CXCL11* chemokine pathways. In contrast, in UC_I, we saw an increase in *CXCL1*, *CXCL2, CXCL3, CXCL8, CXCL13, CCL20* and *CCL21* signalling.

We went on to assess which cellular populations were driving these L-R interactions based on our single-cell data (**Methods**). We superimposed ST derived UMAPS of L-R interactions (**Fig 3.10Di**) with H&E information (**Fig 3.10Dii**) and single-cell derived interactions (**Fig 3.10Diii**). The data for *CXCL9*, *CXCL10* and *CXCL11* showed identical patterns, and we present the exemplar data from *CXCL11-CXCR3* interactions here.

We observed an increase in *CXCL11-CXCR3* signalling in CC_I, across upper crypt, macrophage rich and some mucosal regions (**Fig 3.10Di, ii**). The bulk of this signalling was between myeloid and S3/4 and T cells in health, but this changed dramatically in disease – in UC_I, glial cells played more of a role, whereas in CC_I, inflamed epithelium was a significant producer of *CXCL11*.

Moreover, whilst the populations expressing the cognate ligand, *CXCR3*, remained the same across health and disease (TRMs, Tc17, ZNF683, Th17s), this ligand was significantly upregulated in CC_I (**Fig 3.10E**), likely enhancing the degree to which these cells underwent a greater degree of migration.

Taken together, we could see that areas of epithelial cell damage in crypt tops corresponded to areas of enhanced CXCL11-CXCR3 signalling and corresponded to INF-4 and INF-5 -type spots. By contrast, INF-2 was localised adjacent to these areas of significant damage and contained relatively unaffected tissue (**Fig 3.10F**).

In order to validate our findings we utilized immunofluorescence, using induced nitric oxide synthetase 2 (iNOS2) and FABP1 (increases in a graded expression from crypt base to crypt top, **Fig 3.2B**). We confirmed epithelial expression patterns of iNOS2 protein by IF in health and disease (**Fig 3.10G**) follow the same trends as seen for crypt top *NOS2*-producing epithelium utilizing ST (**Fig 3.9G**).

Epithelial, myeloid and Treg interactions in health and disease

We next looked at how CC_I associated M2 macrophages (*CD163*-high Inflammatory macrophages (1) **Fig 3.4 J,K**) were localizing and interacting in tissue across disease and health. We had established that in CC_I, they were in close association with Tregs and gamma-delta T cells (**Fig 3.9A**) and enriched in INF-2 crypt top areas (**Fig 3.10A**), which in turn were more prevalent in CC_I. We had also established that they expressed higher levels of *CXCL9* in CC_I (**Fig 3.4L**), and CC_I Tregs expressed increased levels of its cognate receptor *CXCR3* (**Fig 3.10E**).

When we looked at *CD163* expression directly, we could see that there was relative enrichment in CC_I as compared to UC_I and health. **Fig 3.11A** shows exemplar ST tissue sections (**Fig 3.11Ai**) and the UMAP for all ST sections (**Fig 3.11Aii**). Moreover, there appeared to be a relative enrichment in peri-crypt areas in CC_I, particularly in the crypt top.

The same pattern was observed for Tregs, although less abundant (**Fig 3.11B**). More specifically in crypt tops, Tregs and macrophages were in close association and enriched in INF-2 (**Fig 3.10A, Fig 3.11C**), which were also areas of high concentrations of *IL10* production. Taken together, we hypothesized that INF-2 (and to a smaller extent, INF-3) was a micro-domain of epithelial regeneration and repair, likely driven by the self-perpetuating association of M2 macrophages and Tregs in these zones.

In order to validate our findings, we again utilized immunofluorescence (**Methods**) for an independent cohort of patients. We validated that CD163-positive cells (macrophages by single-cell analysis, data not shown) were enriched in CC_I as compared to UC_I (**Fig 3.11Di**). Utilizing FABP1 as a marker of crypt depth (low at crypt base, increasing in a gradient to crypt tops) and standardizing across sections (**Methods**), were able to confirm that CD163-positive M2 macrophages were enriched in crypt tops in CC_I as compared to UC_I (**Fig 3.11Dii**).

We went on to repeat a similar analysis for FOXP3-positive cells (**Fig 3.11E**). Although there were no abundance differences (as expected), we could again show that these cells were in more likely to be in association with M2 macrophages (**Methods**) in CC_I than UC_I (**Fig 3.11Eii**). As we had already validated that M2 macrophages were enriched in crypt tops in CC_I, we increased our confidence in the existence of increased INF-2 micro-domains of repair in CC_I.

Prior lines of research had shown that another factor that can skew differentiation of macrophages into an M2 phenotype is cellular apoptosis (as opposed to cellular necrosis, that drives a more M1 phenotype)[34]. It was also known that UC_I had increased numbers of cells dying by necrosis by apoptosis, which could contribute to inflammation[266]. We had also observed that CC_I and UC_I cells were more prone to cell death during tissue processing. We therefore hypothesized that another reason for the skew of macrophages towards an M2 phenotype was an elevated level of apoptosis in CC_I. We opted for immunofluorescence for cleaved caspase 3, which is a common pathway downstream of multiple apoptotic pathways (this effect would not be observable transcriptomically).

We observed a significantly increased level of apoptosis in CC_I as compared to UC_I (**Fig 3.11Fi**), across both epithelial and non-epithelial populations (**Fig 3.11Fii, iii**). We verified that these differences were not due to duration of inflammation in tissue (**Fig 3.11iv**) and instead, disease specific.

Taken together, we concluded that that M2 type macrophages and Tregs were enriched in distinct self-perpetuating clusters in CC_I, particularly towards the crypt top. This was associated with areas

Fig 3.11 - Differential spatial interactions - epithelial, myeloid and Treg responses

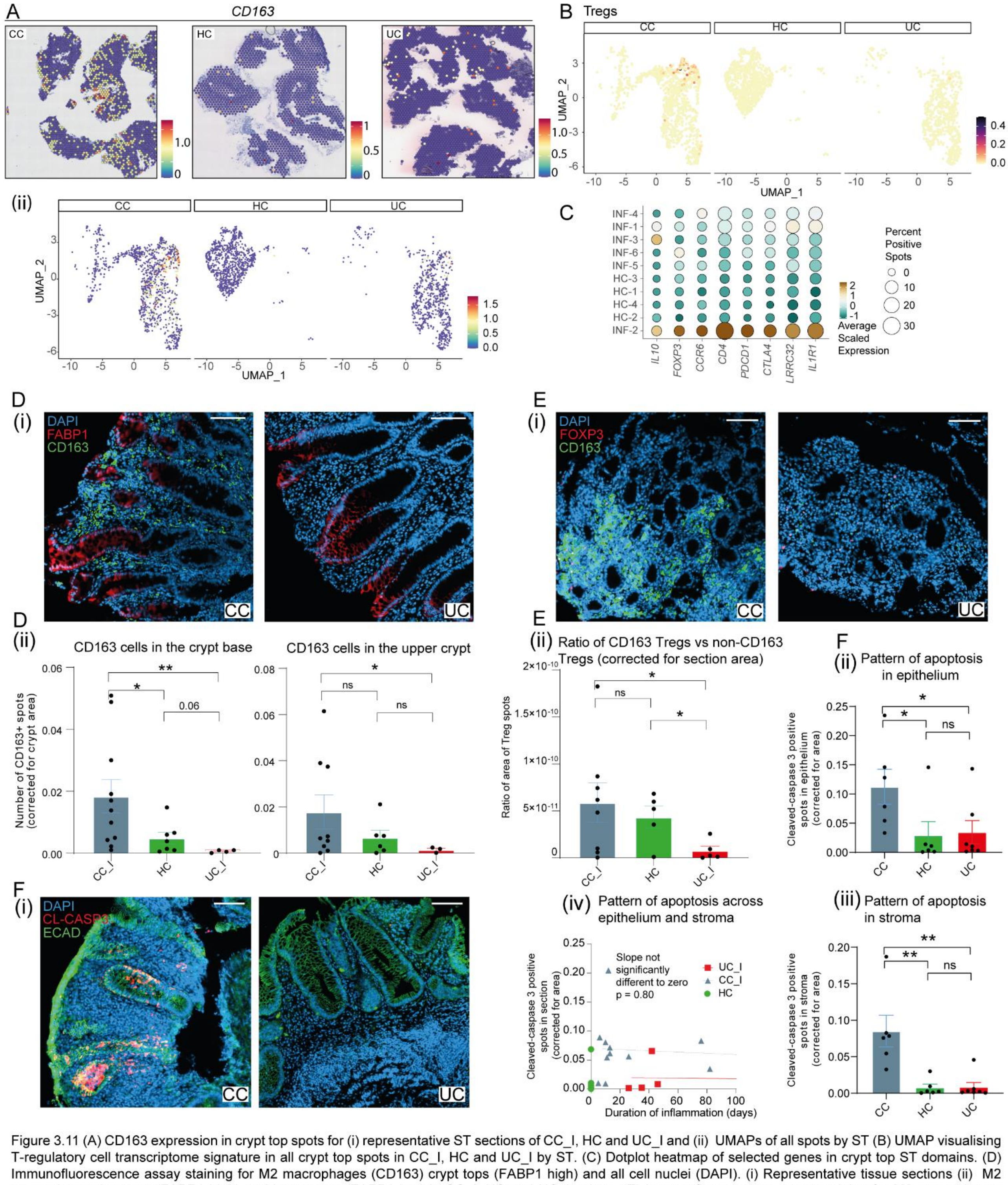


Figure 3.11 (A) CD163 expression in crypt top spots for (i) representative ST sections of CC_I, HC and UC_I and (ii) UMAPs of all spots by ST (B) UMAP visualising T-regulatory cell transcriptome signature in all crypt top spots in CC_I, HC and UC_I by ST. (C) Dotplot heatmap of selected genes in crypt top ST domains. (D) Immunofluorescence assay staining for M2 macrophages (CD163) crypt tops (FABP1 high) and all cell nuclei (DAPI). (i) Representative tissue sections (ii) M2 macrophages in upper (FABP1 high) and lower crypt (FABP1 low) in CC_I, HC and UC_I cohort. (E) Immunofluorescence assay staining for M2 macrophages (CD163), Tregs (FOXP3) and all cell nuclei (DAPI). (i) Representative tissue sections (ii) Tregs in association with M2 macrophages versus non-associated Tregs for CC_I, HC and UC_I cohort. (F) Immunofluorescence assay staining for apoptotic cells (Cleaved caspase-3,CL-CASP3), epithelium (E-cadherin,ECAD) and all cell nuclei (DAPI) (i) Representative tissue sections (ii) epithelial apoptosis (iii) stromal apoptosis (iv) apoptosis across entire section by duration of inflammation for cohort of patients with CC_I, UC_I and HC. CC/CC_I: Checkpoint inhibitor induced colitis, inflamed, UC/UC_I: inflamed ulcerative colitis HC: Healthy control. Statistical analysis for (D) (E) and (F): Unpaired T-test with Mann-Whitney correction; p<0.05(*), p<0.01(**), p<0.001(***), p<0.0001(****), ns = not significant. Scale bar in (D), (E) and (F): 100uM

of relative preservation of tissue architecture, consequent to these zones expressing protective factors such as *IL10*. This patterning was likely due to multiple factors, such as an increase in tryptophan metabolism in CC_I, but in part could also be attributed to an increase in cell death by apoptosis in CC_I, whereas cellular necrosis predominated in UC_I, which would promote the increase in M1 phenotype macrophages we observed there.

Characterizing the interactions of checkpoint-colitis specific HOBIT T cells

In our analysis, we describe an expansion of HOBIT/ZNF683+ T cells by single-cell RNA seq, unique to CC_I (**Fig 3.6B**). We therefore wished to ascertain how these cells were organised and interacting spatially.

As expected, we saw an increase in *ZNF683* expression spatially in CC_I (**Fig 3.12A**), validating our single-cell RNAseq findings. Moreover, this enrichment occurred predominantly across the entire peri-crypt region, which was in keeping with these cells expressing the cognate ligand *CXCR3* (**Fig 3.10Diii**) for the excess chemokines (*CXCL9,10,11*) produced by epithelium in CC_I. There were no specific H&E features visible in these areas of HOBIT cell aggregation (**Fig 3.12Aii**).

From sc-RNA seq data, we could see that these cells, along with cycling, HAVCR2+, GZMK+ effector and FGFBP2+ effector cells, expressed high levels of interferon gamma (**Fig 3.5B**). We had additional evidence that the interferon pathway was highly activated in CC_I, with downstream effects on epithelium and stroma (**Fig 3.2G, 3.3I**). We therefore decided to assess this pathway using spatial transcriptomics.

We saw a marked upregulation of the interferon gamma pathway in CC_I by ST (**Fig 3.12Bi**), validating our findings. This upregulation was most marked in crypt tops in CC_I, but did not manifest obvious features in H&E sections (**Fig 3.12Bii**). TNF followed a similar pattern of expression, being egregiously upregulated in CC_I crypt tops (**Fig 3.12C**).

Fig 3.12 - Differential spatial interactions - epithelial and HOBIT+ T-cell responses

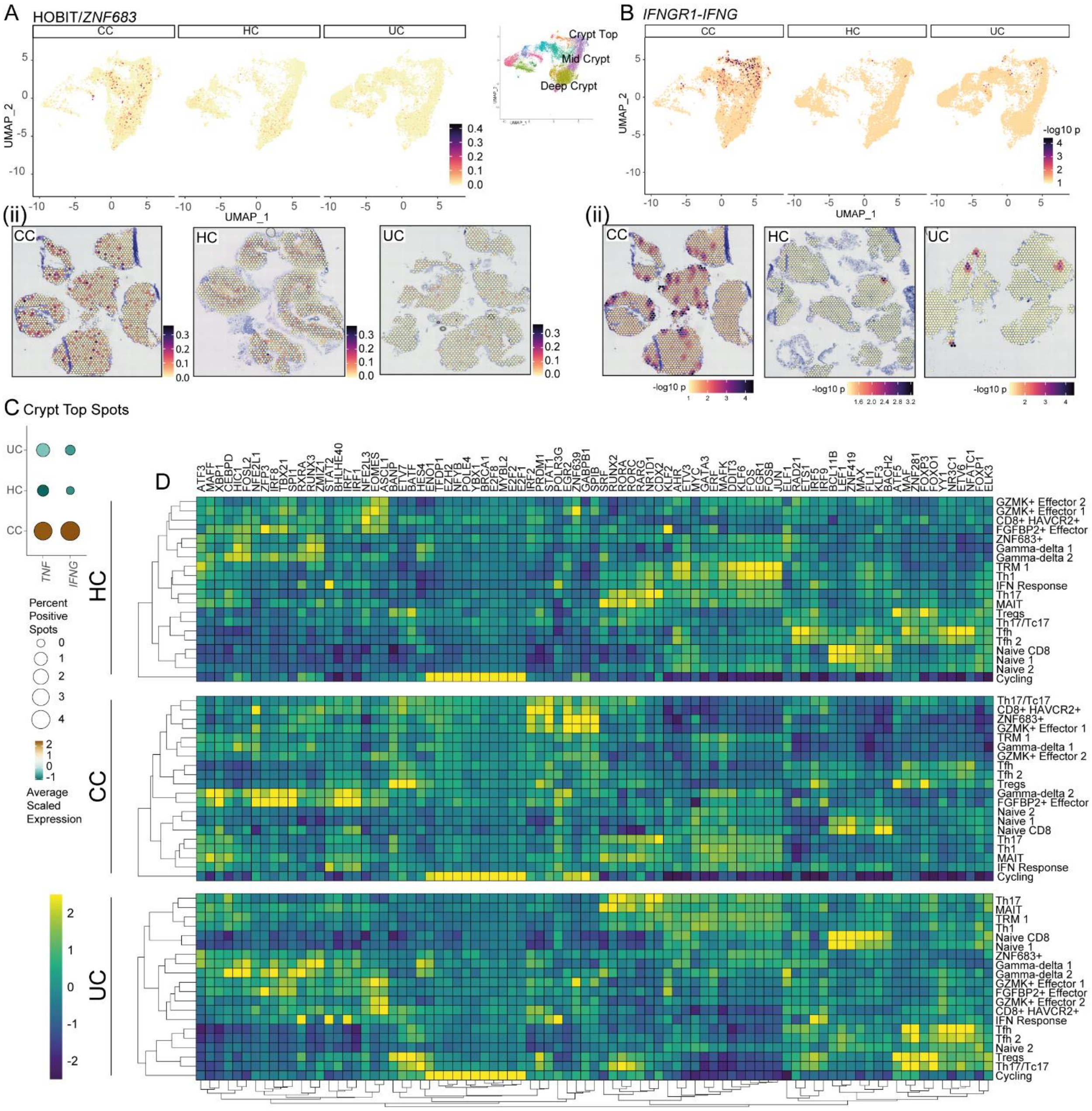


Figure 3.12 (A) (i)Spatial expression of ZNF683+ TRM cell type signature across all sections, visualised over a UMAP embedding and (ii) in representative ST slides for CC_I, HC and UC_I (B) IFNGR1-IFNG receptor-ligand signalling in all ST slides visualized as (i) UMAP overal from all sections and (ii) from representative tissue sections. (C) Relative expression of TNF and IFNG in all crypt top stops in ST in CC_I, UC_I and HC slides (D) Heatmap ofscaled, mean transcription factor network activity AUC value per T-cell subpopulation in cells derived from sc-RNA seq data for health (top), CC_I (middle) and UC_I (bottom).CC/CC_I: Checkpoint inflamed, UC/UC_I: Ulcerative colitis inflamed, HC: Healthy control, Eff: Effector T cell, Mem: Memory T cell. GZMK+: Granzyme K positive, MAIT: Mucosal associated invariant T cells, ZNF683+: HOBIT/ZNF683+ T cells, HAVCR2: Exhausted T cells. TRM: Tissue resident memory.

We performed transcription regulatory network analysis (**Fig 3.12D**) across CC_I, UC_I and health T cells in order to better understand this behaviour. Expression patterns were consistent with established literature (e.g. RORA in Th17 type cells across all states). We also observed other differences for e.g. an upregulation of FOXO1 and MAF in Treg and Tc17 cells that was specific to UC_I, suggesting these cells may be more functionally effective than in CC_I.

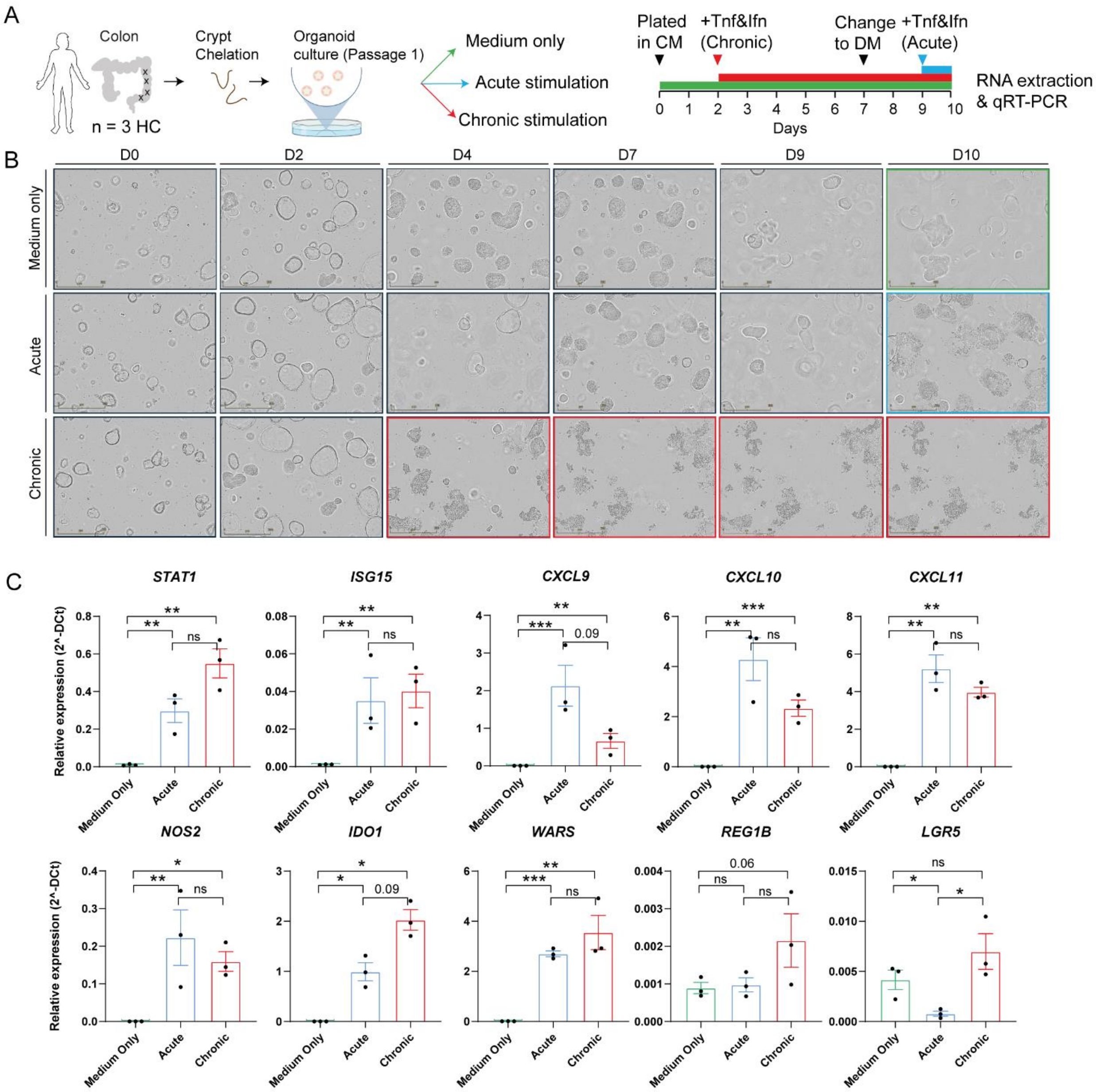
With respect to interferon gamma production, we saw common patterns of activation across diverse cell types, e.g. early response transcription regulator EGR2 in HAVCR2+, ZNF683+, GZMK+ and FGFBP2+ effector and cycling cells in CC_I. In contrast, this was only significantly upregulated in GZMK+ effector populations in UC_I. We saw a marked increase of *STAT1* in these populations in a similar pattern – confirming excess interferon production likely has a positive feedback role in these cells. Taken together, these changes may go some way towards explains the marked increase in interferon gamma signalling in CC_I.

Modelling the effects of TNF alpha and interferon gamma on epithelium.

In order to better understand the functional effects of TNF alpha and interferon gamma on epithelium, we opted for a human colonic organoid model system (**Fig 3.13A, Methods**), which has been extensively characterized and validated, reproducing most of the cellular makeup of homeostasis when allowed to differentiate, with behaviour mirroring that seen in health[15]. As we could see from the spatial data that TNF and IFNg were consistently co-expressed in disease (**Fig 3.12C**), we stimulated an organoid system with a synergistic mixture of both in pre-established concentrations (**Methods**). To better delineate the effects of the duration of injury, we exposed organoids to TNF and interferon both acutely (24hours) and chronically (8 days), recording the results as effects on organoid morphology (**Fig 3.13B**) as well as qRT-PCR outputs (**Fig 3.13C**).

In the medium only and acute conditions, we could see normal organoid development (**Fig 3.13B**), with induction of differentiation (visible as budding) two days after change to differentiation medium (D9), which continued in the medium only condition, in keeping with literature.

Fig 3.13 - Effect of T-cell-produced TNFa and IFNg on epithelium



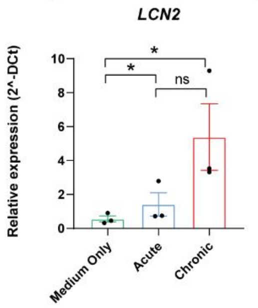


Figure 3.13 (A) Experimental setup - left sided colonic biopsies from 3 healthy patients were processed to form intestinal organoids, passaged once and split into 3 conditions - those exposed to medium only, those chronically exposed to TNFalpha (Tnf, 10ng/ml) and Interferon gamma (Ifn, 10ng/ml, 200IU/ml) or those acutely exposed to both at the same concentration. (B) Representative images taken at 10X zoom of one of those cultures demonstrating changes in organoid morphology over time with conditioning medium and with exposure to TNFalpha and Interferon gamma both chronically and acutely. (C) Quantitative real-time PCR analysis of organoid lysates of a selection of inflammation-associated genes, demonstrating that exposure to these agents is sufficient to induce expression. Housekeeping genes *RPLP0, GAPDH* and *EPCAM* (C): T test (ratio) p<0.05(*), p<0.01(**), p<0.001(***), p<0.0001(****), ns = not significant. T=TNFalpha, I=Interferon gamma. DM = Differentiation medium CM = Conditioning medium HC = Healthy control D2 = Day 2, D7 = Day 7 etc.

Incubation with TNF and interferon gamma, both acutely and chronically, resulted in rapid changes in morphology and an arrest in growth, with no obvious recovery with time or with change to differentiation medium (**Fig 3.13B**).

By qRT-PCR, we saw that TNF and interferon stimulation both acutely or chronically resulted in the upregulation of the interferon response pathway, with an increase in *STAT* and *ISG15*, the (expected) result downstream of interferon receptor engagement, and in keeping with what we saw in UC_I and CC_I tissue. It also resulted in a reproduction of the transcriptomic changes we saw in primary tissue in CC_I, including increases in chemokines (*CXCL9*, *CXCL10* and *CXCL11*), the tryptophan metabolism pathway (*IDO1, WARS*) as well as nitric oxide synthase (*NOS2*) (**Fig 3.13C**).

We also saw a few potentially interesting differences between acute and chronic stimulation (**Fig 3.13C**). Firstly, the transcriptomic marker of cellular regeneration (*REG1B*) increased with chronic, but not acute exposure to an inflammatory environment. In keeping with this, the transcriptomic marker of epithelial stem cells (*LGR5*) decreased markedly with acute exposure to inflammation, but appeared to gradually recover with chronic exposure, possibly indicating epithelial compensatory mechanisms. We also saw a progressive increase in *LCN2* with chronic exposure, a transcript persistently increased despite multiple passages in organoids generated from UC_I colonic tissue[38]. This offered a tantalising possibility that epigenetic changes that we see in UC_I tissue could possibly also be replicated with chronic in vitro exposure to an inflammatory milieu.

The results also offered a potential mechanism for explaining self-perpetuating micro-domains of inflammation in CC_I : activated T cells in CC_I producing TNF and interferon gamma would act on surrounding epithelium to induce production of chemokines which in turn would attract more activated T cells which would produce interferon and TNF (unable to achieve exhaustion because of the presence of checkpoint inhibitors), leading to a positive feedback loop sustaining inflammation.

Interactions between T and B cells in disease and health

From our sc-RNA analysis, we knew that Tfh cells were enriched in UC_I (**Fig 3.6B**) and activated (**Fig 3.12D**). We first confirmed that this effect was not due to the chronicity of inflammation, and a disease specific process (**Fig 3.14A**).

When assessing their functionality in more detail (**Fig 3.14B**), Tfh in UC_I expressed higher levels of chemokine receptors such as *CXCR5*, as well as higher levels of B cell organiser cytokines such as *IL21* and B cell-co-activators such as *POU2AF1*. Taken together, not only are Tfh more numerous in UC_I as compared to CC_I, but also appear to be better adapted to assisting with B cell maturation.

When assessed by ST, we could see this translate into an increased number of lymphoid aggregates in UC_I as compared to CC_I and HC (**Fig 3.14C**), which again, we confirmed were independent of duration of inflammation. We had already observed in our sc-RNA data that germinal centre cycling cells were enriched in UC_I (**Fig 3.4I**), and this helped contextualize how this result came about. In addition, when we superimposed GO term analysis on follicular areas by ST, BCR recombination and B cell maturation was enhanced in UC_I (**Fig 3.14D**), which again would be in keeping with more efficient Tfh and lymphoid aggregate processes in UC_I.

General receptor-ligand analysis (**Methods**) across sections in ST confirmed elevation of diverse pathways associated with lymphoid aggregation and B cell maturation – e.g. stromal 4 interaction (*CCL19/CCL21*), B cell activation (*MIF*) (**Fig 3.14E**).

When combined with sc-RNAseq receptor ligand analysis, we could see aggregation was being driven by chemokines largely secreted by pericytes, stromal 4, Tfh cells and Tc17 cells (**Fig 3.14Gi, 3.14Hi**) across all states, but with the degree of interaction increasing from CC_I, through HC to most in UC_I (**Fig 3.14G, Fig 3.14H**).

We could therefore see how subtle differences in cell prevalence and chemokine secretion (by sc-RNA seq) manifested very differently – epithelial cells being the key attractor in CC_I, leading to a crypt top

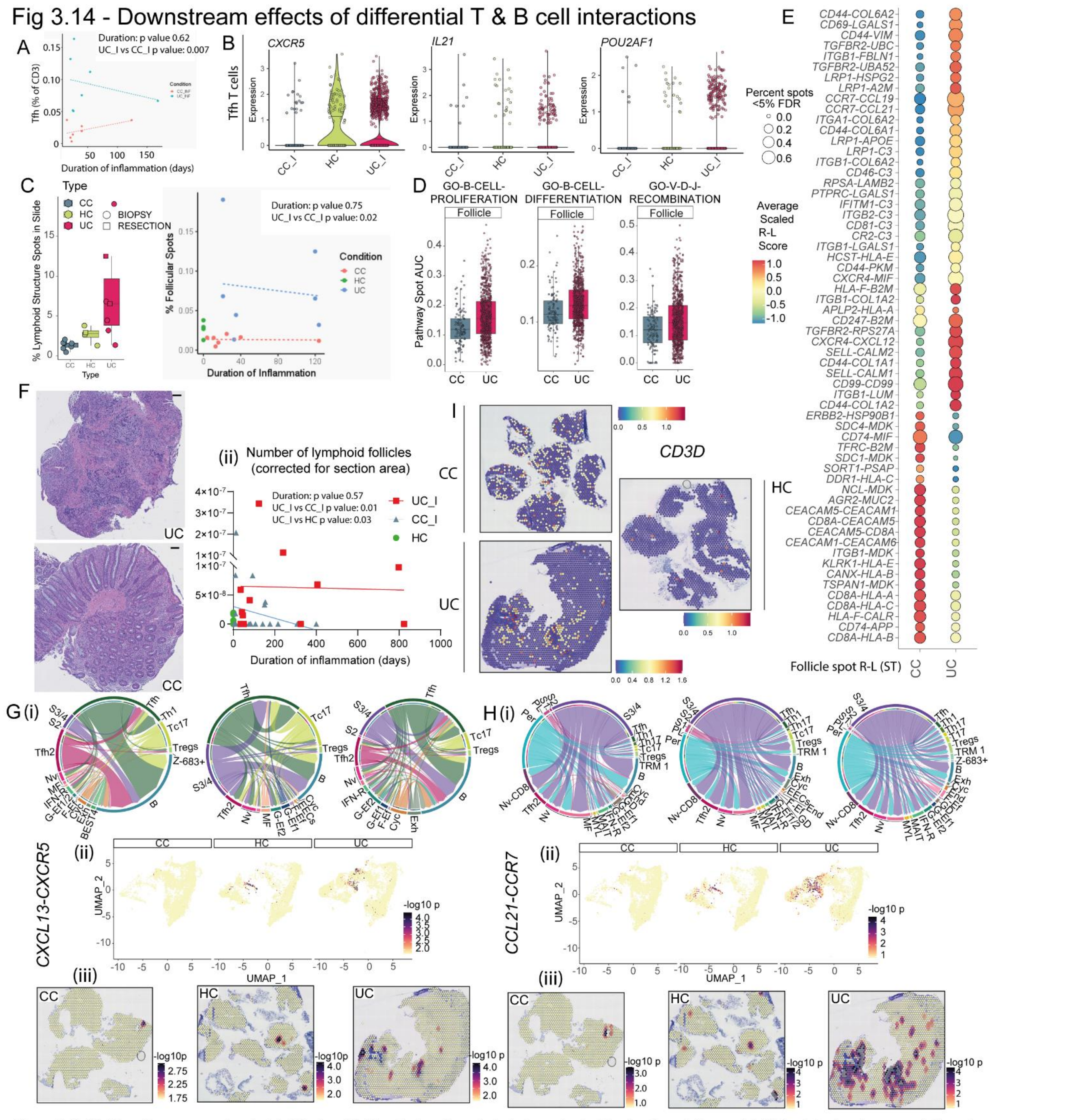


Figure 3.14 (A) Tfh cells are more abundant in UC_I vs CC_I by single cell analysis, independent of the duration of disease (B) Violin plots showing upregulation of expression of CXCR5, IL21 and POU2AF1 in Tfh cells in UC_I as compared to CC_I. (C) Boxplot of the proportion of follicular/lymphoid structure spots for CC_I, UC_I and HC, as a proportion of all ST tissue covered spots passing QC, and how this follicular enrichment in UC_I is independent of duration of disease. (D) Boxplot of GO pathway analysis of significantly differential activity score AUC values across follicular regions by ST in CC_I and UC_I. (E) Dot plot heatmap showing selected receptor-ligand pairs detected as both significantly co-localising within lymphoid structure regions and showing differential signalling strength between UC _I and CC_I slides in lymphoid structure regions by ST. (F) Representative H&E sections from CC_I and UC_I; plot of number of lymphod follicles quantified by H&E for UC_I, CC_I and HC, corrected for tissue area, showing significant enrichment in UC_I which is independent of disease duration. Condition comparison T-test, Linear regression model for duration. (G) CXCR5-CXCL13 pathway: Receptor-ligand analysis (scRNA-seq) paired with ST receptor-ligand analysis and representative sections. (i) : Circos plots showing source (ligand) and target (receptor) expressing cell populations in CC_I (left), HC (middle) and UC_I (right) samples. Coord width is scaled to expression level. (ii): UMAP overlay of spatial distribution of significant co-localisation of CXCR5-CXCL13 spots in CC_I, HC and UC_I by ST. (iii): Representative ST slides showing CXCR5-CXCL13 spots with significant co-localisation in CC_I (left), HC (middle) and UC_I (right). (H) As for (G), but for CCL21-CCR7 pathway. (I): Gene expression overlay for representative ST slides in CC_I, HC and UC_I for T-cell marker gene CD3D. Abbrev: ST: Spatial transcriptomics, CC/CC_I: Checkpoint inflamed, UC/UC_I: Ulcerative colitis inflamed, HC: Healthy control, F-Ef. – FGFBP2+ effector T-cells, Ent – Enterocytes, End – Endothelial, Cyc – Cycling cells, Exh – CD8+ HAVCR2+ T-cells, B- B-cells, Z-683+ - ZNF683+ TRMs, S3/S4 – Stromal 3 & 4, S2 – Stromal 2, S1 – Stromal 1, Gob – Goblets, Per – pericytes, Nv CD8 – naïve CD8+ T-cells, Nv – naive CD4+ T-cells, MF – myofibroblasts, MYL – myeloid cells, IFN-R – Interferon response B-cells, G-Ef1 – GZMK+ Effector T-cells 1, G-Ef2 – GZMK+ Effector T-cells 2, CT-E – crypt top enterocytes, TAs – transit amplifying epithelium, PL- plasma cells, EEC- enteroendocrine cells, Abs and Abs2 – Absorptive Epithelial Cells 1 and 2. High resolution H&E images are available in Supplementary data

aggregation of T cell responses and a relative disorganization of B cell superstructures (**Fig 3.14I**) whereas UC_I was characterized by stromal 4/Tfh cells being key organizers of the inflammatory response, with the T cell response aggregating around distinct lymphoid follicles with enhanced B cell responses. Using this combination of sc-RNA seq and ST, we were able to understand how these processes might self-perpetuate in micro-domains to lead to chronic inflammation.

Identifying cells bound to checkpoint inhibitors in vivo

It is recognised that commercially available anti-PD1 antibody clones (including the CITE-seq clone utilized in this study, EH12.2H7) bind within the PD-1/PD-L1 interaction domain, which is targeted by both Nivolumab and Pembrolizumab[267]. As this antibody also has lower binding affinity for PD1 than both these drugs[268], it would follow that these drugs would likely out-compete the EH12.2H7 clone for binding sites, inhibiting its binding, and this had already been directly demonstrated for Pembrolizumab[269].

Our collaborators in the Davis group characterized the PD1 molecule and the binding sites of Nivolumab and Pembrolizumab (**Fig 3.15A**, unpublished data). They also developed an antibody that bound PD1 at a site distinct to the CITE-seq, Nivolumab or Pembrolizumab binding site, which we designate NonComp PD1 Antibody (manuscript under review).

Using this antibody to verify surface expression of PD1 remained unchanged, we used FACS to demonstrate (**Methods**) that Nivolumab competes with the CITE-seq PD1 clone (EH12.2H7) for PD1 binding (**Fig 3.15Bi**). Nivolumab did not affect the proportion of T cells binding to the non-competitive PD1 binding antibody, but significantly reduced CITE-seq PD1 antibody binding (**Fig 3.15Bii**).

Phase I clinical trial data for Nivolumab suggests that T-cells remain drug-bound for at least 60 days after infusion, and likely up to 200 days[270]. Although phase I trial data was reported for circulating PBMCs, we reasoned that a similar pattern should hold true for T cells in colonic biopsies. We confirmed this by staining for the constant region of Nivolumab and Pembrolizumab (IGG4) and T cells (CD3), confirming co-staining of both in the same cells, a result that could not be expected

Fig 3.15 - Identifying cells bound to checkpoint inhibitors in vivo

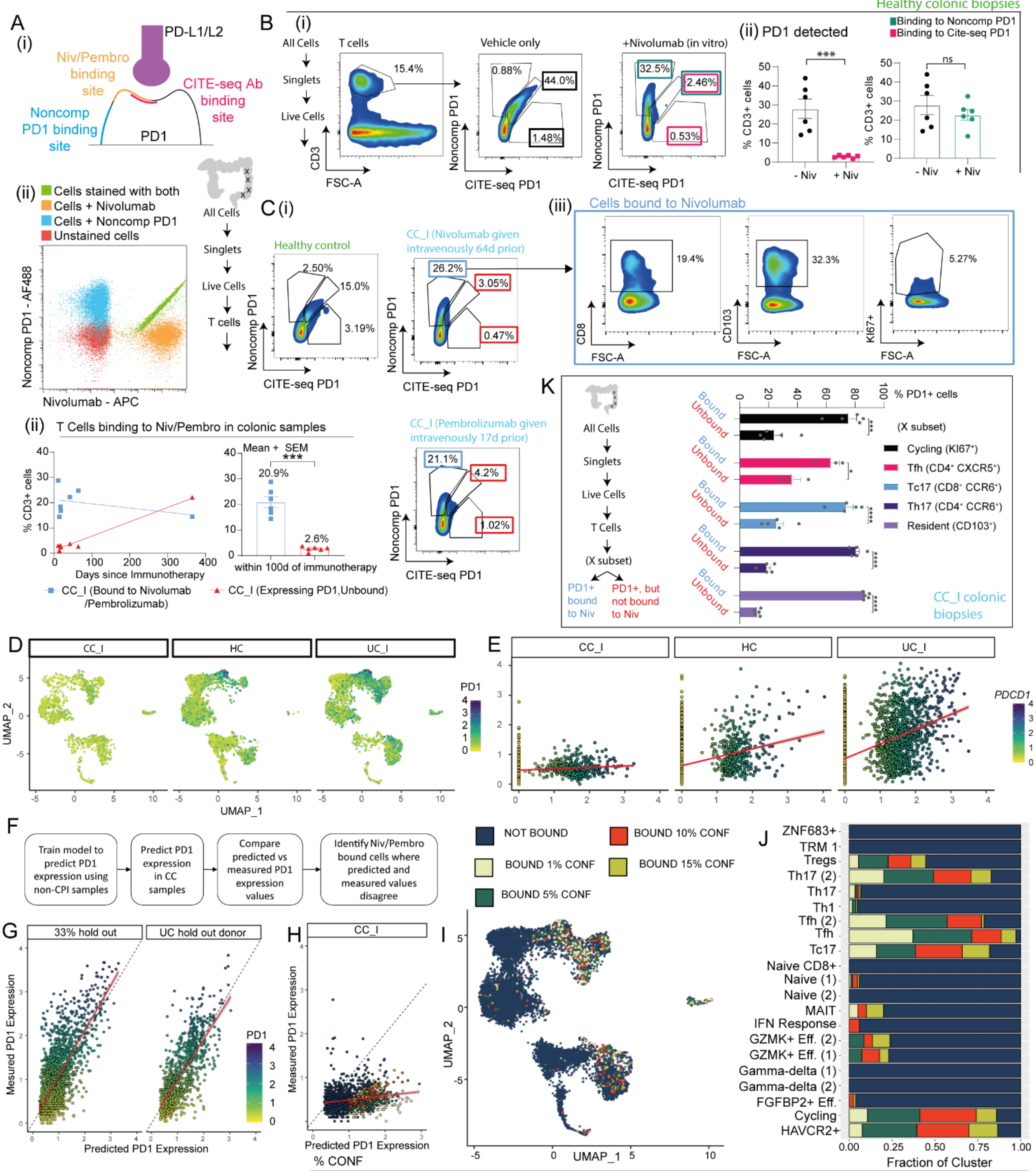


Figure 3.15 (A) (i) Schematic structure of hPD-1 (black) in complex with hPD-L1/L2 (purple), Nivolumab/Pembrolizumab binding site (yellow), CITE-seq EH12.2H7 clone binding site (red) and non-competitive PD1 antibody binding site (blue). (ii) Superimposed FACS plots of engineered JURKAT line over-expressing PD1 stained with non-competitive PD1 antibody conjugated to AF488 (blue), those same cells incubated with Nivolumab conjugated to APC (yellow) or a 1:1 mix (by weight) of both non-comp PD1(AF488) and Nivolumab(APC), coloured green. Cells without any added antibodies (negative controls) are coloured red. On this superimposed plot, double positives (green) lie on a 45-degree angle, demonstrating that Non-comp PD1 and Nivolumab bind the same cells with equivalent affinity. (B)(i) FACS analysis of T cells in healthy colonic biopsy samples in the presence and absence of Nivolumab binding the indicated antibodies (ii) Statistical analysis of PD1 as detected by non-competitive PD1 antibody and CITE-seq antibody clone. (C) FACS analysis of HC and CC_I biopsy samples. (i) Representative plots (ii) Statistical analysis of multiple samples in (Ci). (D): UMAP visualises the protein expression distribution of PD1 (E) Correlation in expression between PDCD1 mRNA and PD1 protein. (F)Computational strategy to identify Nivolumab-bound single cells using CITE-Seq PD1 expression data. (G) PD1 expression regression model validation shows the correlation between predicted and measured PD1 protein expression in non-immunotherapy 33% hold out cells not used for model training (left) and all cells from a single UC hold out donor not used for model training (right). (H) Scatterplot shows the predicted and measured PD1 expression in checkpoint inhibitor treated sample cells. Cells are coloured by the quantile of the divergence between the predicted and measured PD1 values, with the highest confidence bound cells represented as 1% CONF. (I) UMAP overlay visualises the distribution of Nivolumab-bound T-cells. (J) The fraction of each T-cell phenotypic cluster that is predicted to remain bound to Nivolumab in checkpoint colitis samples. (K) FACS analysis to determine proportions of each indicated subtype of T cell that bind Nivolumab. p<0.05(*), p<0.01(**), p<0.001(***), p<0.0001(****) Statistical test (B),(C) and (K): T test. CC_I/CC_NI: Checkpoint inhibitor induced colitis, inflamed/non-inflamed. UC_I: Inflamed ulcerative colitis. HC: Healthy control.

physiologically, clearly indicating we could detect cells bound to anti-PD1 drugs in tissue, distinguishing them from unbound cells (**Supplementary Fig 3.2B**).

We next determined that we could detect cells from colonic biopsies bound to Nivolumab or Pembrolizumab that had been administered to the patient *in vivo* (**Fig 3.15Ci**). PD1+ T cells from healthy colonic tissue freely bound both the CITE-seq antibody and the non-competitive PD1 binding antibody (i.e. were double positive), but PD1+ T cells from patients given Nivolumab or Pembrolizumab were mostly positive only for the non-competitive PD1 antibody (as Nivolumab/Pembrolizumab were occupying the CITE-seq PD1 antibody binding site) (**Fig 3.15Ci**).

For our FACS cohort, all but one sample were collected within 100 days of the last administration of immunotherapy. The outlier, with 400 days between last dose of immunotherapy and sample collection showed a reversion to unbound cells, but remarkably, a significant proportion of those expressing PD1 still remained bound (14.5%) (**Fig 3.15Cii**).

The bound cells in a typical representative CC_I sample included CD8+, CD103+ and KI67+ cells, all of which we knew expressed *PDCD1* in our single cell data, in keeping with what would be expected if our hypothesis were correct (**Fig 3.15Ciii**) .

As all patients within our CITE-Seq RNA-seq dataset received their last immunotherapy (either combination Nivolumab/Ipilimumab or Nivolumab or Pembrolizumab in isolation) within 113 days (Mean 32 days, Median 21 days, range 7-113 days), it was very likely these cells were still bound to Nivolumab or Pembrolizumab, and thus would demonstrate reduced binding to the CITE-seq antibody.

Using CITE-Seq protein expression analysis, we observed that there was an overall reduction in PD1 expression measurements in CC compared to non-checkpoint treated samples (**Fig 3.15D**). Expression of *PDCD1* mRNA showed little correlation between protein and mRNA expression in CC, but not in HC or UC samples (**Fig 3.15E**). As we had demonstrated using non-comp PD1 antibody that this was not due to downregulation or internalisation of PD1 (as these cells still bound this antibody), these cells were likely Nivolumab-bound and the CITE-Seq anti-PD1 antibody was competing for the binding site, and therefore being detected at a lower level.

Next, we leveraged this information to identify individual, Nivolumab and Pembrolizumab-bound cells in our single cell data. Our strategy (**Fig 3.15F, Methods**) trained a quantile regression random forest[271] model to predict PD1 expression in non-checkpoint treated samples using *PDCD1* mRNA expression values, QC-related meta data features and integrated, reduced dimension components. Predicting PD1 expression in our hold-out, non-checkpoint treated sample testing dataset showed good correlation with measured PD1 expression (**Fig 3.15G**). Unsurprisingly, there was little correlation between predicted and measured PD1 expression in checkpoint colitis samples (**Figure 5H**).

Thus, we hypothesized that cells where predicted and measured PD1 expression values disagreed were likely enriched for cells still bound to Nivolumab or Pembrolizumab. To quantify this, we calculated for each cell the probability of where the measured PD1 expression falls within the conditional distribution of model-predicted PD1 values and then thresholded cells at 0.01, 0.05, 0.1 and 0.15 confidence cut-offs as putatively bound T-cells (**Fig 3.15I**).

Bound T-cells across all thresholds were largely consistent, with the majority of high-confidence calls falling into Tfh, Th17, Tc17, cycling and CD8+ HAVCR2+ clusters, as well as a smaller subset of Tregs (**Fig 3.15J**). Effector populations showed putative, low confidence binding at low frequency, while as expected other cell populations, such as naïve cells, IELs and TRM cluster cells were not Nivolumab-bound.

We further confirmed this by FACS analysis, where PD1 expressing cells in CC samples, such as 76% of cycling/KI67+, 64% CD4+ CXCR5+ (Tfh), 74% of CD8+ CCR6+ (Tc17), 81% of CD4+ CCR6+ (Th17s) and 87.5% CD103+ (PD1+ "exhausted" TRMs), were bound to Nivolumab/Pembrolizumab (**Fig 3.15K**).

Clonal analysis of checkpoint inhibitor-bound cells

Nivolumab and Pembrolizumab-bound cells encompassed highly expanded T-cell clonotypes, in particular CD8+ HAVCR2+, cycling cell and Th17 clusters (**Fig 3.16Ai**). However, many of the highly expanded clones in our data, particularly ZNF683+, TRM and FGFBP2+ effector cells, were not predicted to be Nivolumab bound.

Clonal overlap analysis showed that there was also limited sharing of TCR clonotypes between bound cells and unbound cells; however, the most shared clonotypes (4) were between bound cells and ZNF683+ or Treg cells (**Fig 3.16Aii**), followed by GZMK+ effector and Th17 cells (3).

Our scRNA-seq data had indicated that PD1 and CTLA4 were expressed by very diverse T-cell subpopulations exhibiting both population-dependant condition-specific induction (e.g., CD8+ T-cells in CC and UC) and constitutive (e.g., Tregs, Tfh) expression dynamics. Further, some of the strongest CC-associated expanded T-cell populations do not always express *PDCD1*/*CTLA4* (e.g., ZNF683+ TRMs). This apparent contradiction, we speculate, can be resolved if we consider that the T-cells responsible for initiating colitis (i.e., direct drug targets) and those perpetuating colitis due to bystander activation are likely separate populations.

We reasoned that Nivolumab/Pembrolizumab-bound T-cells in our dataset must be present *and* express PD1 at the time of drug administration, prior to the development of colitis. Cells responsible for the initiation of colitis due to direct binding of the drug likely fall within one or more of these cell phenotypes. The "exhausted" CD8+ HAVCR2+ population, cycling cells, as well as PD1-expressing subset of Th17/Tc17s therefore stand out as likely candidates given we demonstrate they bind checkpoint inhibitors and express PD1 in health. This would also mean colitis T-cell dynamics mimic the currently accepted dogma of exhausted T-cell reactivation within the tumour micro-environment[272].

If this hypothesis is correct, existence of pre- "exhausted" T-cell clones would predict which patients would go on to develop colitis. Using our dataset, in CC patients that did not develop colitis, we

Fig 3.16 - Clonal analysis of cells bound to checkpoint inhibitors

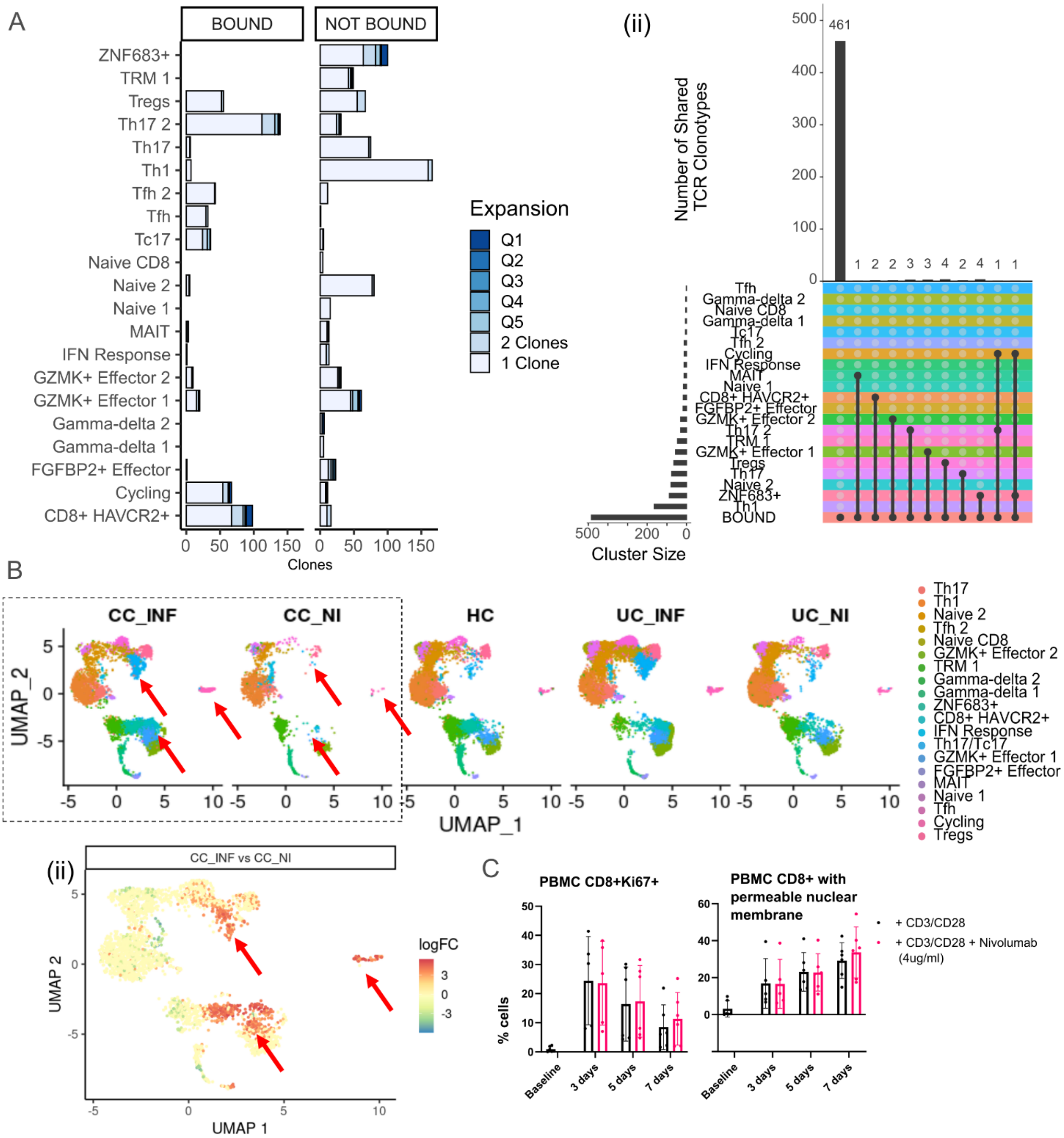


Fig 3.16 (A) (i) Stacked bar plot comparing clonal expansion levels of each T-cell cluster, when split by Nivolumab-bound or unbound T-cells. Q1 indicates highest level of expansion, while 1 Clone indicates only one cell observed with a given TCR (ii) Upset plot showing the TCR clonal overlap between cells predicted to retain Nivolumab binding (BOUND), and cells not bound (all others). (B) UMAP overlay shows T-cell tissue cluster embedding distribution patterns across all clinical sample conditions. Red arrows indicate regions with the highest concentration of Nivolumab-bound T-cells in CC inflamed samples, and the corresponding reduction/absence of these cells in checkpoint treated, but no colitis samples. (ii) UMAP overlay visualising log fold changes of local neighbourhood abundance differences when comparing CC_I and CC_NI. Each point represents a neighbourhood of 10 cells within the single cell k-nearest neighbour graph visualised as a UMAP embedding. Arrows highlight particularly enriched bound populations in CC_I vs CC_NI- Tc17, Tfh2, ZNF683+, GZMK+, exhausted and cycling cells. (C) PBMCs from healthy controls stimulated with CD3/CD28 dynabeads and assessed for active replication (Ki67+) or cell death (high signal from live-dead stain) at FACS in the presence or absence of Nivolumab at a concentration of 4ug/ml. CC_INF/ CC_I: Checkpoint inhibitor induced colitis-inflamed CC_NI: Patients given checkpoint inhibitor with no colitis UC_INF: Inflamed ulcerative colitis UC_NI: Non -inflamed ulcerative colitis; HC: Healthy controls; MAIT: Mucosal-associated invariant T cells

observed a strong reduction of bound-cell enriched populations, such as CD8+ HAVCR2+ and Th17/Tc17, compared to colitis samples (**Fig 3.16Bi**), with significant fold-changes in these populations between CC_I and CC_NI in a direct comparison of the two (**Fig 3.16Bii**). These results suggest that the presence of these cells at the time of checkpoint treatment may predict the future development of colitis.

CC_I specific T-cell populations, such as ZNF683+ TRMs, show both transcriptional signatures consistent with heightened activation and clonal expansion, but we find no evidence of direct binding by Nivolumab. This suggests that while they may strongly contribute to perpetuate colitis, and may derive from bound cells, they are unlikely to be the original, direct initiator cells and were likely re-activated/expanded over the course of colitis.

From a clinical perspective, this analysis approach potentially paves way towards models and assays capable of predicting which checkpoint treatment recipients are at risk of developing colitis or other adverse immune effects. It also implies that factors determining severity and therefore critical to tailoring treatment for CC IRAEs may be different to those responsible for initiation.

Finally, it was not clear whether binding by Nivolumab was inducing cellular death. Although we did not perform detailed cell-tracking experiments across multiple populations, we validated that incubation with Nivolumab in an in-vitro co-culture with PBMCs was insufficient to induce either large-scale cell replication or cell death (**Fig 3.16C,** CD3/CD28 co-stimulation used as control). The same trend was observed for supra-physiological levels of Nivolumab upto 200ug/ml (data not shown).

Developing a model of checkpoint inhibitor-induced colitis

Given the paucity of models of checkpoint colitis[178,257], and our success in modelling the effects of TNF and interferon on epithelium, we attempted to create a more complex epithelial, stromal and T cell model system to allow us to better understand the mechanisms underpinning CC colitis.

Prior attempts to stimulate tissue-derived T cells in isolation in vitro with checkpoint inhibitors had been unsuccessful, with significant T cell death and no observable changes with supra-physiological levels of drug (data not known), so we trialled a system that sought to preserve antigen presenting cells.

We optimized a novel co-culture system (**Fig 3.17A**, **Methods**) that sought to preserve the micro-architecture of native tissue whilst also allowing for epithelial and stromal cell replication. Utilizing FACS with standardized sampling as a measurement system (**Methods**), we established that this co-culture system appeared to preserve T cell and non-immune population numbers, whilst allowing for epithelial and a small amount of stromal propagation (**Fig 3.17B**).

We then wanted to establish whether T cells retained some functionality. Phytohaemagglutinin (PHA) cross-links CD3 and CD28, acting as a non-specific soluble stimulator of T cells[177]. When PHA was added to this system in a concentration series (**Fig 3.17C**), we observed a trend towards increasing *IFNG* expression, along with an increase in the products of IFN gamma receptor-mediated stimulation of epithelium and stroma (*NOS2, STAT1*) and chemokines (*CXCL9*, *CXCL11*) that we had come to associate with inflammation. Notably, incubation with Nivolumab alone had no measurable effect.

We attempted to refine the system further, adding a dual stimulation with PHA and Nivolumab arm (**Fig 3.17D**), assessing organoid morphology as well as RNA outputs. There was no clear visual effect of stimulation by any method over incubation with medium alone (**Fig 3.17E**). Although there was a trend towards increased *IFNG* production by T cells (**Fig 3.17Fi**) as well as the downstream products of epithelial/stromal stimulation (*IDO1*, *NOS2*, *CXCL9,* **Fig 3.17ii**), this effect did not reach significance, primarily due to inter-sample variability. Although there was some suggestion that some interferon-signal driven transcripts (*IFI6*, *ISG15*) were enhanced by the presence of Nivolumab, this was only seen in one sample.

Taken together, although this system represented a novel advance in modelling epithelial, stromal and immune cell interactions in healthy colonic tissue, it was plagued by inter-sample variability, and

Fig 3.17 - Modelling checkpoint inhibitor effects on T cells and epithelium

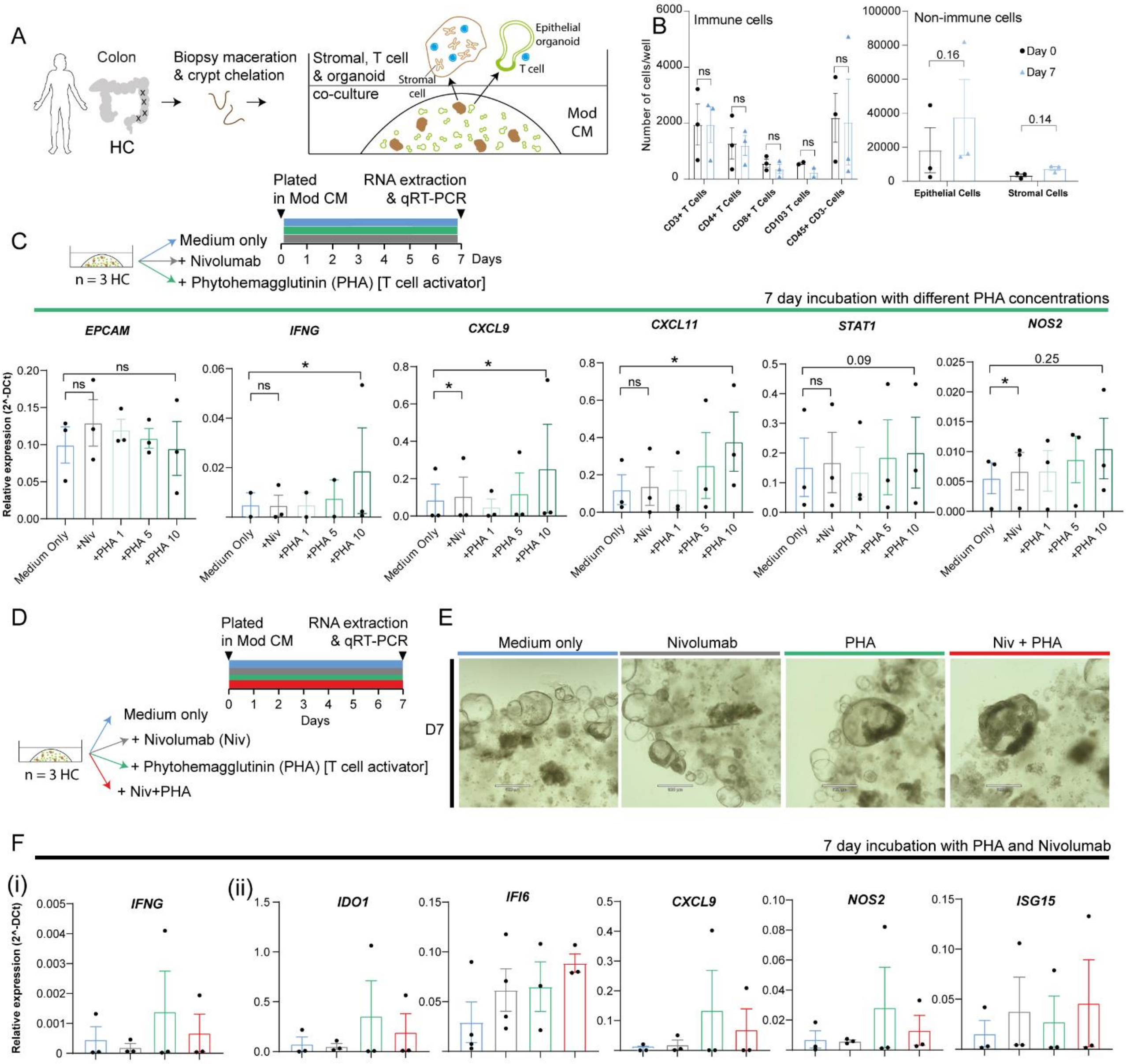


Figure 3.17 (A) Schematic of novel intestinal co-culture model, utilizing left sided colonic biopsies from healthy controls, aimed at understanding epithelial, stromal and immune behaviour. (B) FACS data, survival of different cellular fractions in this model system, comparing baseline and 7day timepoint.(C) Experimental layout of model system exposed to increasing concentrations of phytohaemagglutinin (PHA, values represent concentration in ug/ml) as well as Nivolumab (50 ug/ml) and medium only control, results of quantitative real-time PCR analysis of co-culture model system lysate. Selected comparisons shown, those not shown did not reach significance (D) Experimental layout of model system exposed to PHA (10ug/ml), Nivolumab (50 ug/ml), combined PHA and Nivolumab, with medium only control. (E) Representative images at 10X zoom of each arm in (D) at 7 day timepoint. Scale bar = 520um. (F) Results of experiment (D) at 7 day timepoint by quantitative real-time PCR of co-culture model system lysate. None of the comparisons reached significance, bars not shown for clarity. Housekeeping genes for qRT-PCR *RPLP0, GAPDH* and *EPCAM.* Statistical tests: (B): T test, (C), (F): T test (ratio) p<0.05(*), p<0.01(**), p<0.001(***), p<0.0001(****), ns = not significant. Mod CM = Modified conditioning medium, HC = Healthy control, PHA = Phytohaemagglutinin.

limited to a 7 day interval. Further work, replication and optimization would be required to determine if this was a viable model for understanding colonic or checkpoint-inhibitor induced inflammation.

Discussion

Our study aimed to utilize next generation single-cell RNA technologies combined with novel spatial and functional experiments to understand epithelial, stromal and immune behaviour and interaction in checkpoint inhibitor induced-colitis and ulcerative colitis.

Characterizing immune and non-immune cell behaviour To date, our study represents the only analysis of pan-compartment behaviour in checkpoint inhibitor-induced colitis across blood and tissue. It also represents the only unbiased spatial transcriptomic analysis of both checkpoint inhibitor induced colitis and ulcerative colitis.

Comparing the epithelial and stromal response was very useful – it revealed that despite these diseases having different immune drivers, in addition to a significantly increased burden of interferon gamma in the inflammatory milieu in CC colitis, the essential response of the epithelium and stroma is similar in UC and CC. It raises the possibility that the epithelial barrier defects seen in CC (and ergo, UC) may be the result of epigenetic changes in the context of inflammation rather than an intrinsic defect.

Despite this overall similarity, we detected definite differences between the epithelial response in CC and UC, with significantly higher expression of chemokines *CXCL9, CXCL10* and *CXCL11* in the epithelium. This, in turn, leads to increased migration of resident lymphocytes into the epithelium in CC as compared to UC, which we show both with spatial transcriptomics as well as more conventional immunofluorescence techniques.

The immune response is where the most differences lie between CC and UC, which is perhaps unsurprising because of their clearly different aetiologies. In the non-T cell compartment, we saw significant differences in memory B cell and myeloid populations between UC and CC, the first study to directly demonstrate this. Our study was also the first to study T lymphocyte migration between blood and tissue in UC and CC, demonstrating that effector cell migration played a significantly greater role in UC than in CC or HC. We saw that re-activated HOBIT+ TRM cells are likely responsible for driving the inflammatory response and are significantly enriched and cycling in CC, whereas the IL26+ CD8 population (which lay within the Tc17 cluster in this analysis) was central to UC. Despite this, there were still marked similarities in the immune response in both disease processes, again suggesting a stereotyping of response in inflammation. Identifying these disease-specific processes in closely related but distinct diseases offers the first step towards understanding pathogenesis and tailored treatment strategies.

Understanding cellular interaction Spatial transcriptomics was key to transforming our understanding of cellular organisation and interaction in health and disease. Using this, we firstly were able to identify that distinct subsets of epithelial, stromal and neural cells that manifested similar transcriptional signals were actually spatially co-localized in micro-domains, suggesting that the milieu induces very diverse cell types to behave in similar, stereotyped ways. We went on to establish that in CC, HOBIT+ TRMs formed self-perpetuating micro-domains of inflammation with epithelium and stroma, mediated by cytokines (*TNF*, *IFN*) and chemokines (*CXCL9, 10, 11*). We were also able to determine that Tregs were drawn to these domains, polarising macrophages to M2 phenotype in the crypt top through IL10 secretion, but were still ineffective at controlling inflammation, with high levels of *IFN* and *TNF* in these regions (likely because of checkpoint inhibitors remaining bound to their surface receptors).

It also revealed that the increased *CCL19* and *CCL21* signals we observed in UC were unique to the disease process, likely mediated by an increased/effective Tfh response and in turn, responsible for increased germinal centre B cell and follicular burden in UC, independent of the duration of inflammation.

Finally, we were able to utilize a novel variation of the use of colonic epithelial organoids to demonstrate how increased production of TNF and IFNg by HOBIT+ TRM cells in CC might be sufficient to drive all the transcriptomic changes we observed, as well as set up self-perpetuating microdomains.

Predicting those patients at risk of developing colitis : We developed a wholly novel approach for detecting which cells (at a single-cell level) were bound to anti-PD1 checkpoint inhibitors that had been given in vivo. Using this, we determined, surprisingly, that the cells responsible for initiation of colitis (Tfh, exhausted cells) were distinct to those responsible for its perpetuation (TRM, Cycling cells). It however, did go on to explain why follicles were more prevalent in UC, as Tfh PD1 binding likely prevented effective functioning of these cells, hampering follicle development in CC. It also provided us with an indication that patients with higher burdens of exhausted, Tc17 and follicular cells at baseline may be at higher risk of developing checkpoint inhibitor-induced colitis.

Developing a model system for CC colitis Although not a core aim of the project, we developed a novel epithelial, stromal and T cell co-culture model system, which was able some of the key functional changes we observed in our single cell data in inflammation. Although still requiring optimization, it offers a potential alternative for exploring the effect of different drug therapies that rely on disrupting aberrant cellular interaction.

Altogether, we were successful in our aim of describing how checkpoint inhibitor induced colitis is distinct to UC, as well as putting forward a possible mechanism for its initiation and propagation.

Discussion

Conclusions

Checkpoint inhibitor induced colitis (CC colitis) is an inducible colitis in humans, a novel clinical entity, and its incidence is set to increase. Idiopathic ulcerative colitis (UC colitis) is also increasing in incidence globally and is incompletely understood; both diseases present a therapeutic challenge. Utilizing patient derived data, single-cell transcriptomics, spatial transcriptomics and organoid modelling of disease, we compare both these diseases, deriving insights into the distinction between pathways activated in inflammation universally and those particular to each disease.

Studies published contemporaneously[20,128,187,273], have found similar conclusions to our work. Analysis of immunotherapy induced disease in other organs[79,274] has also highlighted CD8 T cells expressing residency markers as being critical drivers of disease. A different stream of analysis identified polymorphisms in the IL7 gene associated with an increased risk of developing immunotherapy related adverse events[154,155]. IL7 is expressed both centrally in lymphoid tissue and in rapidly replicating epithelial cells [248,275] which supports the hypothesis that interactions between antigen-presenting cells such as B cells and CD8 T cells may be driving CC colitis. Together with our clinical data, it is consistent with mucosal inflammation being the primary driver of inflammation in CC, and being clinically prognostic, whereas systemic inflammation is equally relevant in UC.

Our analysis of CC colitis goes beyond what is known till date, identifying changes in stromal and epithelial populations, trafficking between blood and tissue, and tissue interactions between macrophage, T and B cells that drive disease, as well as differences between CC and UC. We identify aberrant re-activated TRM ZNF683+ CD8+ populations as a key pathogenic driver, and demonstrate how it interacts with epithelium in stroma in novel self-perpetuating micro-domains in inflammation. Such domains are adjacent to anti-inflammatory Treg-M2 macrophage micro-domains and we hypothesize that the relative balance between such concomitant pro-and anti-inflammatory agents may determine prognosis.

We highlight a novel approach to analysing the behaviour of checkpoint inhibitor drugs – utilizing anti-PD1 drug pharmacokinetics to identify drug-bound cells in tissue and blood. We have utilized this to predict which populations may be responsible for triggering colitis, although this would ideally require longitudinal studies for validation.

We, as others, sought to replicate and understand the changes brought about by inflammation in epithelium and stroma utilizing organoid and co-culture experiments in vitro. This appears to be in line with current efforts to produce a better model of human inflammation[168,174,276–280]. We were partially successful in our goal, although this system needs refinement, and all such models are limited by lacking the equivalent of local lymph and peripheral blood compartments. Mouse models of CC colitis are currently lacking, which in of itself is interesting[257,281–284] but efforts are ongoing to improve on the status quo and develop a more physiological model of human disease.

We also investigated patterns of disease in checkpoint inhibitor induced hepatitis[285,286] and checkpoint inhibitor-induced gastritis[287] (analysis ongoing), which we have not presented here for the sake of succinctness and clarity. The conclusions from this work however support our hypothesis that locally resident T cells are drivers of disease in the colon.

Future work

We are currently engaged in analysis of cellular interactions between bound cells and epithelial and stromal populations utilizing novel high resolution spatial transcriptomics (CosMx)[288], which we hope will further resolve the impact that drug binding has on T cell populations and knock-on effects on cellular interactions between these cells and their adjacent stroma. Although we hope it would provide further mechanistic insights into colitis, if successful, it would provide an attractive and novel approach to understanding checkpoint inhibitor behaviour in the context of cancer, which is an area of intense interest.

Such cellular 'labelling' could also be potentially extended to understanding which cellular populations are targeted by current standards of biologic therapy (vedolizumab, infliximab and ustekinumab). Either through spatial or single-cell based analysis, they may help offer finer resolution in advancing our understanding of primary drug resistance.

Our work in **Chapter 2** demonstrates that IL26 may represent a potential new therapeutic target in the treatment of UC. The work would need to be validated in additional mouse models, and may require more direct elucidation of the target receptor and mechanism, for which organoid models may prove useful, prior to considering therapeusis. It may then join the growing armamentarium of drugs that could be utilized for treatment of colitis. Its specificity for UC (as opposed to CC), expression limited to colonic as opposed to PBMC CD8 T cells, incremental increase with disease severity, as well as it being flagged up as a gene of relevance in GWAS based models of predicting the risk of developing UC, may make it a specific biomarker for predicting UC flare recurrence. This may emerge in the near future as the wealth of studies seeking to predictively characterize UC reach fruition.

The work within **Chapter 3** was conceived as occurring within the context of the PRISE study[289], which aimed to prospectively collect blood, stool and tissue from patients prior to, and following on from, initiation of immunotherapy. This was delayed significantly by the Covid-19 pandemic, but is still planned to complete. We hope this will confirm our hypothesis of the cellular populations responsible for initiation of CC colitis, and therefore validate our approach of utilizing drug binding to interrogate disease.

The microbiome, and T cells responsive to the microbiome, have been identified as key to response of cancer to immunotherapy[120,290–294]. Recent work, although yet to be replicated, demonstrates that FMT may be as effective (both in terms of speed and efficacy) as standard immunomodulator therapy in the induction and maintenance of remission in UC[67]. Although only case reports of the efficacy of

FMT in CC colitis exist, our work leads us to hypothesize that the microbiome is also relevant in the development and progression here, particularly given the relevance of the resident T cell population. The PRISE study could help define this in more detail through analysis of stool prior to, and after induction of immunotherapy and/or development of colitis, particularly if a signature emerged that identified patients at higher risk of development, or relapse of colitis. Independently, we are pursuing another line of enquiry to see if we can identify the epitopes that might be binding to the most highly replicating clones in CC colitis[295] and ideally seek to understand what might be driving the disease. Moreover, it may prove interesting to compare any taxonomic or metabolomic signatures that are protective for both UC and CC as it may uncover ground truths for understanding epithelial health in more detail. This might not only improve treatment in UC/CC, but may also provide insights into as yet poorly understood phenomena, such as diversion colitis.

Multi-omic single cell analysis (combining proteome, transcriptome and epigenomic information) has recently been developed[296,297]. It is well recognised that the epithelium undergoes epigenetic changes in inflammation[38,244] that impair its functional replicative capacity long term. We have observed similar patterns in CC colitis derived organoids (unpublished data). It would be of interest to characterize these changes in more detail, utilizing CC colitis (particularly if we can distinguish baseline epithelium from inflammation-induced changes in the context of PRISE) in order to better understand barrier defects in idiopathic UC. Therapeutically targeting epithelial regeneration through modifying epigenomics during an inflammatory episode, if possible, would provide a novel approach to treating inflammation.

As highlighted before, the distinction between inflammation as a physiological response and pathology is persistence or excessiveness. We observe significant overlap in the epithelial and stromal response, that appears to behave in a stereotypic fashion, in two different forms of inflammation, with differences emerging when analysis focused on patterns of cellular interaction, inferred through spatial relationships. It is possible therefore, that the answer to why inflammation does not resolve in

IBD lies in understanding higher-level cellular interaction rather than individual cellular behaviour. High-resolution sub-cellular transcriptomics, if paired with a human tissue sampling strategy that was able to capture resolving versus persistent colonic inflammation, would be key in advancing our understanding into therapeutically targeting epithelial and stromal dysfunction. Combined with improved sophistication in modelling, either in-vitro (e.g. colonoids) or in-silico, forthcoming insights are likely to result from improved understanding of dysfunctional cellular interaction. Consequently, changing cellular interaction may require unforeseen therapeutic strategies, such as transplantation of engineered segments of bowel (e.g. in the case of fistulizing Crohns'), populated by patients' stromal and epithelial cells grown in culture, correcting irretrievably altered tissue relationships that may be prone to recurrent inflammation.

Finally, other forms of inflammatory bowel disease (e.g. microscopic colitis, collagenous colitis, common variable immunodeficiency induced colitis) remain under-explored, largely due to difficulty in prospectively acquiring samples. Comparisons with acute colitis (e.g. CMV-induced colitis), which might advance understanding of epithelial and stromal restitution, are also difficult for similar reasons. Novel FFPE-based technologies such as CosMx and MERSCOPE offer some novel methods for characterization of these rarer forms of disease, advancing us towards a more complete understanding of barrier function in the colon. Longer term, this may change how we classify inflammatory disease in the colon, moving us beyond historical, histological descriptions of disease. The goal is to move beyond understanding dysfunction, towards understanding the individual, complex, pleiotropic yet symbiotic relationship our immune system as a whole holds with our microbiome and the diet we consume.

Appendices

Appendix A

Supplementary figure for chapter 2

Fig Supp 2.1 - Multi-modal profiling increases our understanding of CD8 T cell behaviour

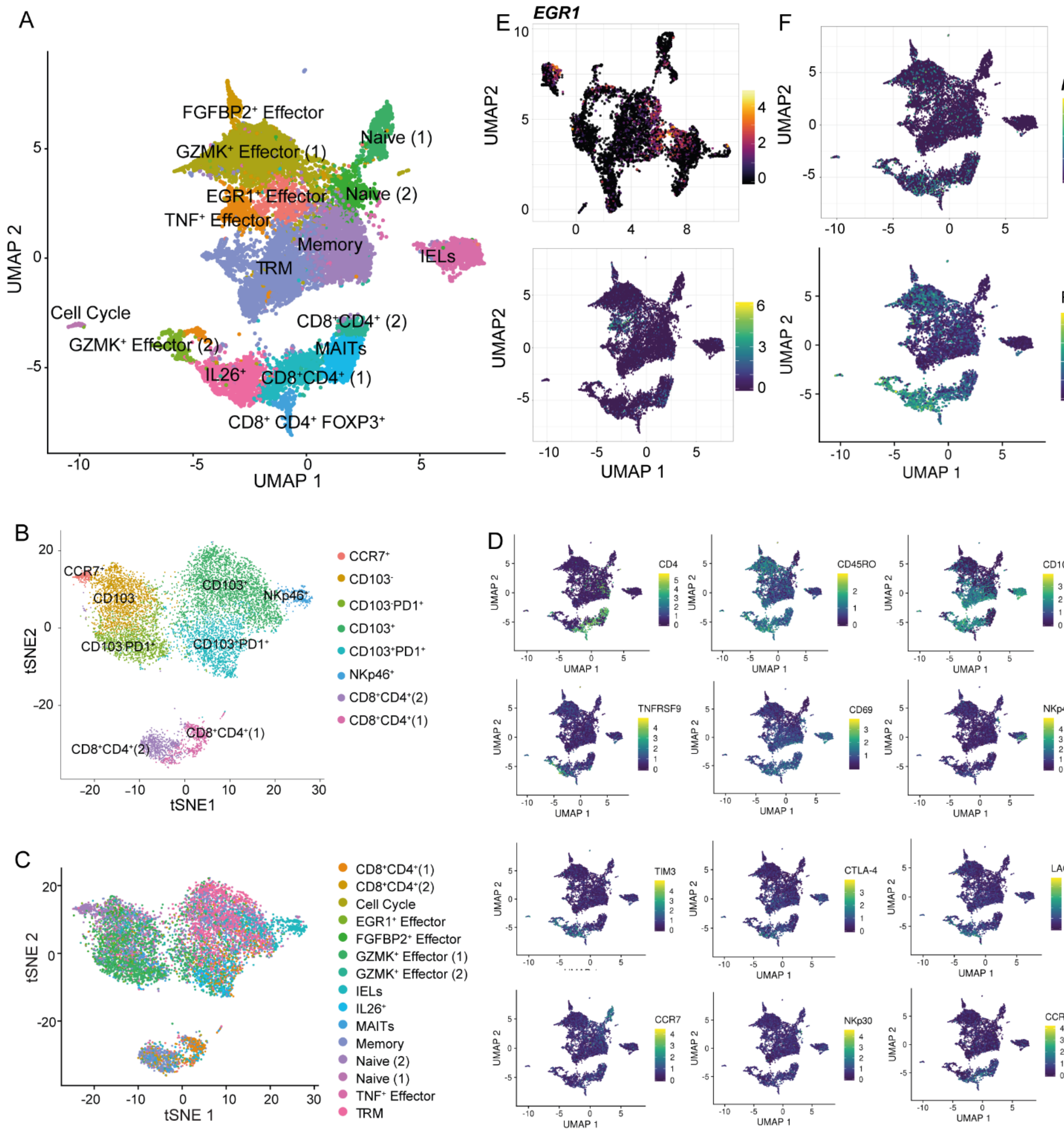


Figure Supp 2.1. (A) UMAP of CD8 T cells in inflamed UC and health derived from a validation cohort that combined single-cell gene-expression and cell surface protein information (CITE-seq). (B) t-sne clustering of cells only utilizing their cell-surface protein data, demonstrating distinct clusters of cells by their co-expression patterns. (C) Clusters of cells as defined by their single-cell gene expression data projected onto the t-sne of clusters as per their single-cell cell surface protein expression (as measured by CITE-seq). (D) Cell surface protein expression projected onto the clusters derived from gene expression profiling of tissue CD8 T cells in inflamed UC and health. Highlighted are the expression of the following proteins: CD4, CD45RO (antigen-experienced cells), CD103 (resident T cells), TNFRSF9 (activation marker), CD69 (activation marker), NKp44 (activation/NK cell activity marker), TIM3, CTLA4 and LAG3 (exhaustion markers), CCR7 (marker of central and naive cells), NKp36 and CCR6. (E) Expression of *EGR1* gene - the effector EGR1 cluster is not clearly delineated in the initial cohort (top) whereas it separates out in the validation cohort (bottom). (F) Expression of the *PDCD1* transcript (top) compared with corresponding protein PD1 expression (bottom) - *PDCD1* is a lowly expressed transcript, and cell surface PD1 allows for easier identification of PD-1 expressing cells

Supplementary table for chapter 2: Cluster summary of tissue CD8+ T cells

Cluster	**Key Marker Genes**	**Key Antibodies**	**Notes**	**Key TFs**	**References**
Naïve (1) and (2)	*LEF1, CCR7, SELL, TCF1*	CD8+		*TBX19, TBX21, LEF1, BACH2, IKZF1, ID1, ZEB1, STAT3*	[298,299]
GZMK+ Effector (1) and (2)	*GZMK,GZMH IFNG, KLRG1, EOMES, TNFRSF9 GZMB*-low,	CD103- CD8+		*EOMES, STAT3, BATF, TBX21*	[2]
TNF+ Effector	*TNF, IFNG, JUN*		Clustered with GZMK+ cells, delineated in CITE-seq analysis	*EGR2*	
EGR1+ Effector	*GZMK, IFNG, KLRG1, EOMES, GZMH, GZMB,*	CD103- CD8+	Clustered with GZMK+ cells, delineated in CITE-seq analysis	*EOMES, EGR1, EGR2, EGR3, EGR4, IRF7, IRF8, IRF2, TBX21*	
Memory	*TCF7 IL7R FXYD2 GZMA*	CD8+ CD45RO	Potentially Central Memory / Resting CD8 T cells / Gradient from Naïve population		[95,188,300]
CD8+CD4+	*TNFRSF4, IL17, IL22, IL1R1, IL1R2,CCR6*			*RORC*	

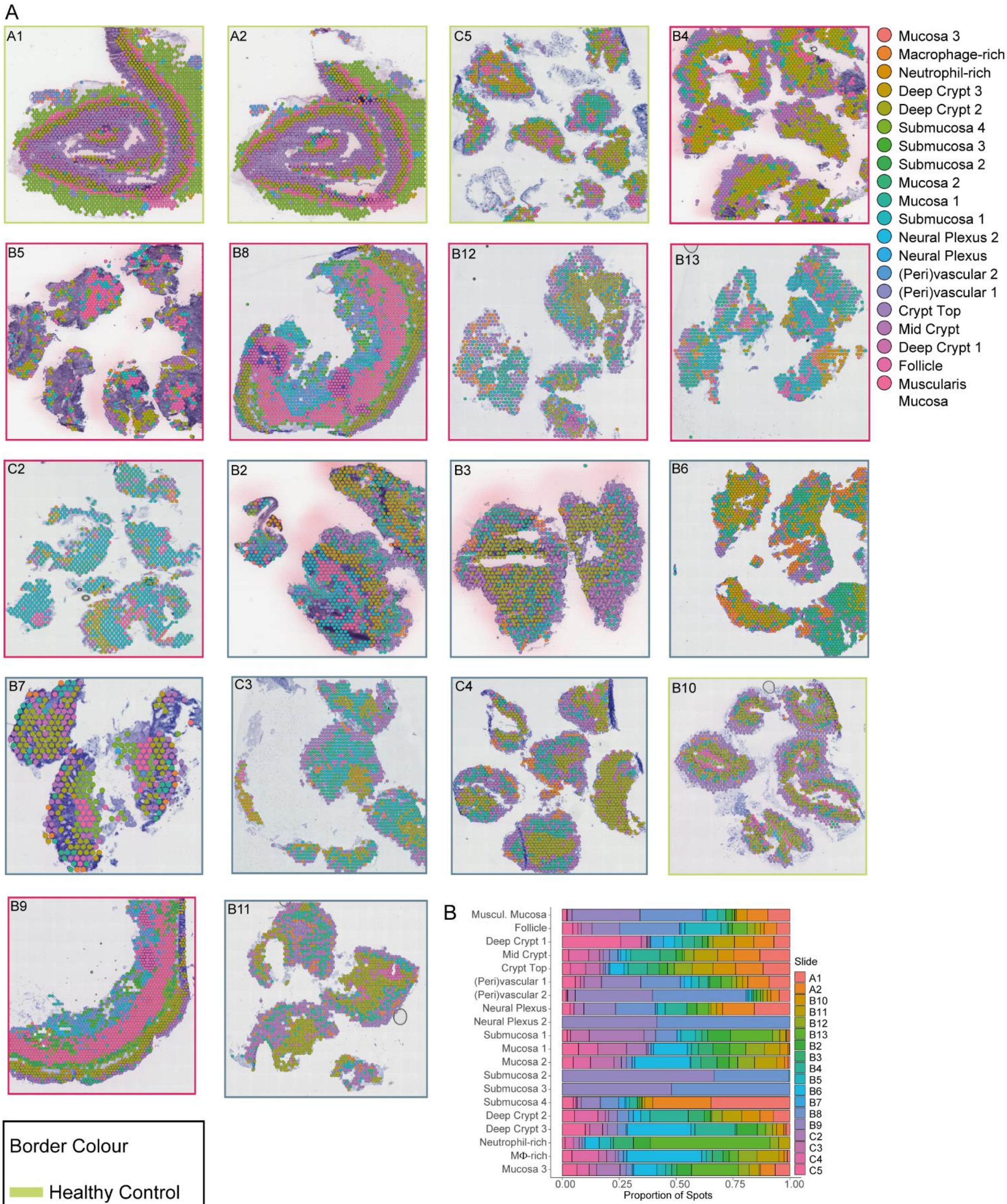
CD8+ CD4+ FOXP3+	*FOXP3, TNFRSF4, IL1R1, IL1R2, CCR6, IL10*	CD4+ CD8+	Regulatory T-cells	*RORC*	[301–303]
IELs	*KIR2DL4, CD160*	CD103	TYROBP+ (natural) and TYROBP- (induced) clusters		
Gamma-Delta	*TRDC*	CD8	Clustered with IEL cells		
TRM	*ITGAE*, *IL7R, TNFRSF9, ZNF683*	CD103+ CD8+	Tissue-resident memory cells	*EGR1*, *EGR2*, *EGR3*, *STAT4*, *RUNX3*	[2,304]
IL26+	*IL26*, *IL23R*, *NCR3*, *IL21*, *IFNG*, *PDCD1*, *CTLA4*, *HAVCR2* *CCR6*, *CCL20, IL22*, *IL17A*,	CD8+ PD1+ CTLA4+	Inflammation-counterpart to Th17 cluster, CD8+, transcriptionally very similar to Th17 2	*STAT3*, *RORC*, *BATF3*, *MAF*, *PRDM1*, *IRF4*, *AHR*, *STAT5*, *GATA3, c-KIT*	[301,305]
FGFBP2+ Effector	*FGFBP2*, *CX3CR1*, *IFNG*, *KLRG1*, *EOMES*, *GZMH*, *GZMB*	CD103- CD8+	Circulating, vessel-confined CD8+ effector cells.	*KLF2*, *KLF3*	[2,306,307]

MAIT	*TRAV1-2*, *TRAJ33*, *ZBTB16*, *KLRG1*, *SLC4A10*, *NCR3*	CD8+	Mucosal-associated invariant T-cells	***ZBTB16*,** *RORC*, *RORA*, *TBX21*, *EOMES*, *IKZF1* (Helios), *EGR1*, *ELK3*, *ZBTB7A*, *ZBTB7B*, *ZBTB17*, *ZBTB14*, *ZBTB33*	[308,309]
Cycling	*MKI67*, *CTLA4*, *STMN1*	CD8+	CD8+ actively proliferating cells in G2M and S-phase	*E2F8*, *E2F2*, *E2F1*, *E2F7*, *E2F4*, *E2F5*	[310,311]

Appendix B

Supplementary figures for chapter 3

Fig 3.Supp 1 - All Visium spatial transcriptome slide sections



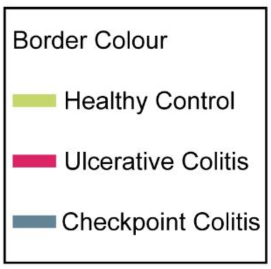


Figure 3.Supplementary1 (A) All spatial transcriptomics sections by disease type, with UMAP overlay of spot types detected. Accompanying high resolution H&E images are in supplementary data. (B) Spot type broken down by contribution from different slides

Fig 3.Supp 2 - CD4 T cells trafficking and clonal dynamics in health and disease

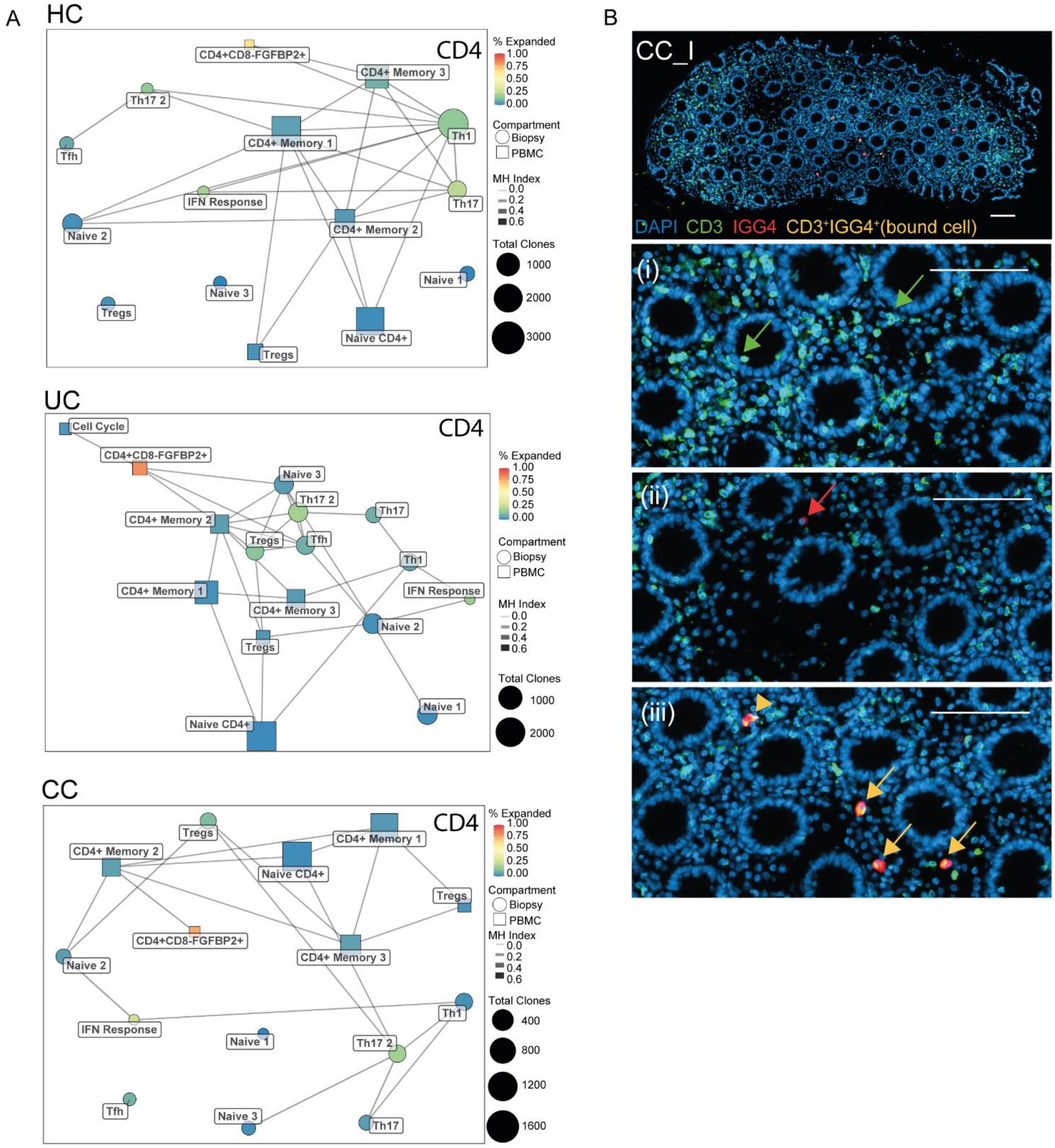


Figure 3.Supplementary2 (A) Analysis of patterns of CD4 T cell trafficking in health and disease. Combined depiction of size of clonotypes, degree of sharing and degree of expansion. (B) Immunofluorescence of representative FFPE section of CC_I. Full image (top), with selected regions (i-iii) below. (i) Single positive CD3+ T cells, green arrow; (ii) Single positive IGG4+ plasma cell, red arrow (iii) Double positive CD3+IGG4+ cells, do not occur physiologically, demonstrate T cells coated with Nivolumab/Pembrolizumab, yellow arrows. Scale bar = 100uM.

Supplementary tables for chapter 3:

3.1 Cluster summary of all cells

	Cluster	Sub-Cluster(s) if present	Key Marker Genes	Abundance	Notes	Refs
Epithelial	Absorptive	Absorptive (1) and (2)	*TIMP3, TAF4B, BAMBI, HIST1H3C, ARC*	Fewer in CC_I vs CC_NI (*)	Two clusters differ because of gradient of gene expression	For all Epithelial sub-clusters below: [20,312]
	BEST4		*BEST4, OTOP2, SPIB, HES4*	Fewer in CC_I vs CC_NI (**)		
	Crypt Top Enterocytes		*AQP8, CEACAM1, CEACAM7*	Fewer in CC_I vs CC_NI (**)		
	EEC (Enteroendocrine)		*CHGA, PYY, CRYBA2, SCGN, FEV*	No changes in Inflammation		
	Enterocytes		*CA1, SLC26A3, SELENBP1*	Fewer in CC_I vs CC_NI (**)		
	Goblets		*MUC2, TFF3, ITLN1, SPINK4, WFDC2, KLK1, REP15, TPSG1,*	CC_I > CC_NI (Tr)		
	Inflammation Enterocytes		*NOS2, WARS, IDO1, DUOX2A, DUOX, CXCL11*	More in CC_I vs CC_NI (**), Same trend in UC		
	Stem Cells		*LGR5, ASCL2, SMOC2, OLFM4*	No changes in inflammation		

	TA (Transit Amplifying) Cells	TA and TA (G2M) clusters	*MKI67, ZWINT, TYMS. GINS2, CENPM, CDC45, MCM10*	No changes in inflammation	TA and TA G2M differ in degree of replication signal expression	
	Tuft		*HTR3E, GNG13, SH2D7, SH2D6,TRPM5,AZ GP1,*	No changes in inflammation		
	Stromal Cells		*CAV1, SOX8, NTRK2, CCL23, MIA, SERPINF2,*	No changes in inflammation	Small residual stromal population in epithelial enriched fraction	
	T-Cells		*CD3E, CD69, CD96, CD7, ZNF331, EVL*	CC_I > CC_NI (Tr)	Intra-epithelial T lymphocytes	
	B-Cells		*CD19,CD79A, CD79B, MS4A1, CD69, EVL*	No changes in Inflammation	Intra-epithelial B cells	
	Mast Cells		*TPSAB1, TPSB2C*	No changes in Inflammation	Intra-epithelial Mast cells	
	Myeloid Cells		*CD14, C1QA, C1QB, C1QC, CTSB,*	CC_I > CC_NI (*), > UC_I (Tr) and HC (Tr)	Intra-epithelial Myeloid cells	
	Plasma Cells		*IGLC2, IGLC3*	Trend towards more in CC_I vs	Intra-epithelial Plasma cells	

				all other conditions		
Stromal Cells (Enriched Fraction)	Myofibroblasts		*ACTG2, MYH11, LUM, HHIP, SOSTDC1, MFAP5, NPNT, FOXL1*	Depleted in CC_I vs CC_NI(**) and HC(**). No differences between UC and HC.	Production of Lumican higher in inflammation – may be secondary to TNFa stimulation. *F3, POSTN, EDNRB, MRVI1, MMP11, FOXL1, PDGFRA* shared with S2	[24,29,59,313,314]
	Myofibroblasts 2		*MYH11, GREM1/2, RSPO3, RSPO2, CHRDL1*	No differences in inflammation	*GREM1/2* as BMP antagonists help maintain a stem-cell niche. Abundant in Submucosa. Deeper Myofibroblast cluster (Seen	[315–317]

					most in resection ST sections B8 and B9) *GREM1* and *GREM2* are a marker of fibroblast activation in lung *RSPO2* shared with S2 *GREM1,GREM2, CHRDL1, RSPO3* shared with S3	
	Pericytes		*COX4I2, RGS5, NOTCH3, C1QTNF1, FAM162B, STEAP4, KCNJ8*	CC_I>CC_NI(****) and HC(**). Trend towards UC_I>UC_NI and HC.		[24,29,59,318,319]
	Endothelial		*PECAM1, CD34, VWF, EGFL7, PLVAP, RAMP2 &3, CLDN5, CYYR1, CALCRL,*	CC_I > CC_NI(***), CC_II > HC(**),	Endothelial cell markers are shared with subcluster 2 of Plasma cells	[24]

			ECSCR, SOX18, ADGRL4, HYAL2,			
	Glial		*S100B, NRXN1, GPM6B, CDH19,PLP1, MPZ, SPP1*	CC_I>UC_I(*), but the same as HC		
	Stromal 1 (Wnt2B+ FosHi and FosLo)	S1 (1) and S1(2) subclusters .	*ADAMDEC1, FABP4/5, CCL8/7 MCTP1*	Depleted in CC_I vs CC_NI(***) and HC(*). Depleted in UC_I vs UC_NI(*) and HC(Tr)	Gradient of *CCL13, SCT, GSTM1,PROM1, POSTN, DMKN, OASL, GCH1, MX2* expression between two subclusters *WNT2B* and *RSPO3* producer, along with S3 *MCTP1* shared production with Activated S1 cells (*CCL2, ABCA8, COL14A1, APOE* that were Historically used were not	[24,29]

					specific S1 markers in this dataset)	
	Stromal 2 (Wnt5B)		*BMP5, Wnt5B, COL4A5, COL4A6 PDGFRA, NSG1, AGT, ENHO, VSTM2A, HAS1,TRPA1, SOX6, NRG1, GLP2R, ALKAL2* Global S2 markers	Depleted in CC_I vs CC_NI(**) and HC(*)	Supports epithelial stem-cell niche through production of Wnt, in conjunction with *GREM1/2* produced by myofibroblast layer that inhibits BMP that would otherwise inhibit the Wnt pathway. *F3, POSTN, EDNRB, MRVI1, MMP11, FOXL1, PDGFRA* shared with Myofibroblasts *FRZB* shared with Pericytes	[20,29,320–324]

					APOD shared with S3 *LIMCH1* shared with Endothelial *HSD17B2* shared with Epithelium	
		S2(1) and S2(2)	*Gradient of SCUBE2, MMP11, FGF9, HAS1,GFOD1 DUSP2 between clusters*	As above	S2(1) and S2(2) likely represent fibroblast expression gradient differences across cluster	
		S2 TLL2+	*RSPO2, SMOC2, CPM, CYP1B1, PTGIS, PDGFRL, TLL2, BLNK, TSPAN7, ADAMTSL3, ASB13*	No changes in inflammation	Function of subset unknown. Primary RSPO2 producing S2 cluster *SMOC* is a intestinal stem cell marker and acts to drive fibrosis through MF -> Fibroblast	

					transition in Kidney & Lung (possibly via TGFB1 pathway) Expresses global S2 markers	
		S2 NPY+	*NPY, LTBP1, AREG,FST, SLITRK6,SLC4A4, CYTL1,FBN2,F2RL 3,SEZ6L2, BMP7,PTX3,GJA1*	CC_I < CC_NI (**) and UC_I(*) and HC (Tr).	Function of subset unknown. Localised around Vascular structures in the deeper stroma by ST Expresses PID pathways for ERBB (Epithelial Growth factor receptor), LIS1 (Neuronal migration), DELTA-NP63 (Hallmark of stemness),A6B1	[325–327]

					-A6B4-INTEGRIN (Hallmark of stemness) NPY only produced by this cluster, does not also produce PYY or PPY which are otherwise produced by EEC *ITGA8, PTGS1, LEF1, PAPLN, PTX3, GJA1, LTBP1, AKR1C1, TBXAS1 , MMP3, SMOX, BMP7, STC1, NRG1, LOX* shared with activated fibroblasts Expresses global S2 markers but	

					does not produce *RSPO1/2* or *NDUFA4L2* like the other S2 subsets	
	Stromal 3		*GREM1, OGN, PRELP, IL33, C1QTNF3, EBF1, FGL2, CD34* *Global S3 markers*	CC_I>CC_NI(***) and HC(**). UC_I>UC_NI (Tr) and HC(Tr)	Clustered around vascular structures by ST	[24,29]
		S3	*SFRP2, THBS2, AKR1C1,PLA2G2A, OSR2,CHRDL2,AOX1,WISP2,SLPI, CHRDL1*		*MGST1, MEDAG, CFB, FIBIN* common with activated fibroblast cluster *RSPO3, SPRP1* shared with S1 *MGST1* common with epithelium	

		S3 PI16+	*PI16,OGN,MFAP5*	CC_NI>CC_I(*)	*CPM, MFAP5, VIT* shared with S3 Belongs to deeper S3 Subset (B8) More RSPO3/less immune focused	[328]
		S3 C7+	*C7,C1QTNF3,CCL19,CYP7B1*	No changes in inflammation	*CHRDL1, GPC3* shared with S3 More superficial S3 Subset (B8) More immune focused	
	Stromal 4 (S4)		*CCL19, CCL21,CTSH,FDCSP*		Localised around lymphoid aggregates/follicles *CTSH,FDCSP* in common with	[29]

					activated fibroblast cluster	
	Activated Fibroblast		*IL13RA2, IL24, FAP* *TWIST1, CXCL8, CXCL10, WNT5A* (but not *WNT5B)* Global Activated Fibroblast markers		Formed predominantly of activated S1, S2, CSF3+ clusters *CD82, IL13RA2, IL24, FAP ,TWIST1 , CHI3L1, NQO1, THRC1, CLU, PDPN* also shared with S3 and S4	59
		Activated S1	*CP* *RAMP1* *CHI3L2*	CC_I>CC_NI(*), UC_I>HC(Tr)	*RAMP1, CHI3L2* also expressed by TNFSFB11 (S1) *CSF3* Shared with Activated TNFSFB11 and activated CSF3	

		Activated S2	*DEFB1, FBH1, KIAA0930, GPX7, IL7R, ITGA2,MMP10, MYH11, IL1RL1, CXCL11*	CC_I>CC_NI (**) and HC(*). UC_I>UC_NI(*)	*SMOC1* shared with activated TNFRSF11 (S1) *ITGA2* shared with S2-NPY *CXCL11* shared with Endothelial cells	
		Activated CSF3+	*LIF,IL6,CXCL8,CSF3,RARRES1,TRAF1*	UC_I>UC_NI (*), HC(Tr), CC_I>CC_NI(Tr) and HC(Tr)	*CSF3, CXCL6, CXCL8, RARRES1, STEAP1, TWIST1, PRRX1, IL24, CHI3L2* shared with other activated S1 clusters *FDCSP* shared with other activated FB clusters and S4 *CCL19* shared with S4 *GREM1* shared with S3/S4 *TRAF1* shared with Activated C3+	

		Activated TNFSFB11 S1	*IL13RA2,RGS5,MGST1,CXCL6,CXCL8,IL11,RAMP1,STEAP1,RARRES1,ASS1,ASL,VIPR2,GK,CEMIP,FIBIN,CHI3L2 PF4V1,CHST15,FAM20A,PI15,SLC16A4,FDCSP,CP*	CC_I>CC_NI (***) and HC(*). UC_I>UC_NI (Tr) and HC (Tr)	*CP* also expressed by S1 close to S3/4 *ASS1* also expressed by S3 *FDCSP* also expressed by S4 *CXCL8* aslo expressed by CSF3+ *SULF2, SNX10, CTHRC1, WISP1* Expressed with other activated fibroblasts RGS5 expressed with pericytes	
		S1 Interferon response	*IFIT1/2, OAS1/2/3/, OASL, CD69*	No differences in inflammation	*CXCL10* (shared with activated Fibroblast cluster) Shares *CCL7/8, ADAMDEC1* with parent S1 cluster	

		Activated C3	*C3, CFB, CD24,TAC3,TNFSF11,TRAF1, CST1*	UC_I> CC_I(**), UC_NI(*) and HC(Tr)	*C3* shared with S3 and S4 *CFB* shared with S1 activated *CD24, TNFSF11* shared with S4. *TRAF1* shared with activated CSF3+ Produces more BMP1 than other activated FB	
	Epithelial		*ELF3, EPCAM, KRT19, LGALS4, CLDN3*		Small number of residual epithelial cells in stromal/CD45 enriched population	
Immune Cells (Stromal Enrichment)	Cycling		*MKI67, ZWINT, CD3D, CD19, PTPRC*	Increased in CC_I and UC_I vs NI and Health.	Conglomerate cluster comprising cycling immune cells, largely T	

					cells, B cells and plasma cells	
	NK&T cells		*CD3D, CD7, KLRB1*	Analysed in more detail utilizing CD3 dataset		
	Mast		*CPA3, TPSAB1, TPSB2, MS4A2, TPSD1, SLC18A2*	No differences in inflammation		
	Myeloid		*CD14, AIF1, SPI1, CD68, C1QA, C1QB, C1QC, IGSF6, LST1*	CC_I>CC_NI (***), CC_I>HC(*), Trend towards CC_I>UC_I		
	B Cells		*C19,CD22, DMS4A1, BANK1, CXCR5,FCRLA, PIKFYVE, TNFRSF13C, FAM129C*	No changes in inflammation		
	Plasma	Plasma 1, 2 and 3 subsets differentiated by immunoglobulin expression	*MZB1, DERL3, IGLL5, CD79A, TNFRSF17, SPAG4, FCRL5, ZBP1, IGLV3-1, IRF4, CCR10, CD38*	CC_I < CC_NI(*), trend towards CC_I < HC.	Co-expressed *KRT19* possibly associated with Epithelium. Co-expressed *DES* – possibly associated with MF	

	cDC		*CD1C, CD1E, FCER1A, CLEC9A, FLT3*	Increased in HC vs CC_I(**) and UC_I(*)		For all Myeloid cell subsets reported below: [59,329,330]
	pDC		*CLEC4C, IL3RA, BCL11A,TCF4*	Trends towards increase in both UC_I and CC_I		
	TolDC (mregDC)		*CCL19,CCR7,LAMP3, FSCN1*	No differences in Inflammation	Tolerogenic Dendritic cells	
	Inflammatory Macrophages	Inflammatory Macrophages (1)	*CCL18, CXCL1, S100A13*	Increased in CC_I vs HC (**) and UC (*). No differences between UC_I and HC		
		Inflammatory Macrophages (2)	*NFKBIA, S100A6, CXCL11, SIRPB1, ETV7,*	Increased in CC_I (*), trend of increase in UC_I vs HC		
Myeloid Cells		Inflammatory Macrophages (3)	*S100A12, SERPINB2, APOBEC3A,NFKBIA, S100A6, IL1B,*	Increased in UC_I (*) and CC_I(**) compared to HC		

		Inflammatory Macrophages (4)	*INHBA, IL1A, SLAMF1*	Trends towards increase in both UC_I and CC_I		
	Macrophages (without pro-inflammatory signals)	Macrophages (1) and (2)	*CD14 CD68*	No differences in inflammation	No pro-inflammatory signal	
		Macrophages (3)	*RNASE1, IGF1, ETV5, MRC1*	Reduced in both UC_I(*) and CC_I(*) compared to HC	Expresses markers of residence (MRC1)	
	Cycling		*MKI67*	Increased in HC vs CC_I(**) and UC_I(*)	Cycling Myeloid cells.	
B and Plasma Cells	Atypical B Cells		*CD19, ITGAX, CD86, EBI3, GPR137B, BATF, SIGLEC6, FCRL4, SOX5, LIMK1, TESC, MPP6, CD58, HMOX1, WEE1, ZBTB32,*	No differences in inflammation		For all B cell subsets described below: [187,329,331–333]
	CD27-		*CD19, TCL1A, IL4R, IGHD, BCL7A, CD200*	No differences in inflammation		

	CD27+		*CD19, CD27, CD70, KLK1, TMEM273, PVT1, TNF*			
	Germinal Centre		*CD19, RGS13, SUGCT, SERPINA9, MYO1E, ELL3, HOPX, GCSAM,*	UC_I> CC_I (Tr) and HC (Tr)		
	Germinal Centre Cycling		*CD19, RGS13, SUGCT, SERPINA9, MYO1E, ELL3, HOPX, GCSAM, MKI67, ZWINT*	UC_I> CC_I (*) and HC (*)		
	Plasma		*ANKRD28, XBP1, PRDM1,TNFRSF17 , TRIB1, RRBP1, DUSP5, , SPAG4, SDC1, AQP3.*		CD19 negative	
	Interferon Response		*IFI6, IFIT3, WARS*	UC_I > HC(*), CC_I > HC (Tr)		

Legend: Tr: Trend, $p<0.05$ (*), $p<0.01$(**), $p<0.001$(***)

3.2 High resolution cluster summary of tissue T cells

Cluster	**Key Marker Genes**	**Key Antibodies**	**Notes**	**Condition Enrichment**	**TCR Analysis**	**Key TFs**	**Ref**
Th17	*IL26, IL17A, IL22, KIT, IL23R, CCR6, CCL20*	CD4+		Enriched in HC, baseline Th17 cluster	BloodShare: 0.0197 % singleton: 0.86	*JUNB, ATF7, BATF3, FOXP3, RORC, STAT3, RORA, RARA, RUNX2*	20,59,188,301,334–336
Th1	*IFNG*	CD4+			BloodShare: 0.0153 % singleton: 0.91	*STAT1, STAT4, TBX19, TBX21, GATA3, FOXP3*	301,337–339
Naïve 2	*LEF1, CCR7, SELL*	CD4+			BloodShare: 0.00531 % singleton: 0.98	*RUNX2, TCF7, KLF2, FOXP1, LEF1, BACH2, IKZF1*	2,340–342

Naïve 3	*LEF1*, *CCR7*, *SELL*	CD4+			BloodShare: 0.00565 % singleton: 0.97	*TCF7*, *FOXP1*, *FOXP3*, *RUNX2*, *KLF3*, *LEF1*, *BACH2*, *IKZF1*, *ARID5A*	[2,340–342]
Naïve CD8	*LEF1*, *CCR7*, *SELL*	CD8+	Naïve CD8+ counterpart of naïve CD4+ cells		BloodShare: 0.0196 % singleton: 0.98	*TBX19*, *TBX21*, *LEF1*, *BACH2*, *IKZF1*, *ID1*, *ZEB1*, *STAT3*	[298,299]
GZMK+ Effector 2	*GZMK*, *IFNG*, *KLRG1*, *EOMES*, *GZMB*-low, *GZMH*-low	CD103- CD8+		Enriched in UC over CC	BloodShare: 0.162 singleton: 0.73	*EOMES*, *STAT3*, *BATF*, *TBX21*	[2]
TRM	*ITGAE*, *IL7R*	CD103+ CD8+	Tissue-resident memory cells	Depleted in UC but not CC	BloodShare: 0.0733 singleton: 0.65	*EGR1*, *EGR2*, *EGR3*, *STAT4*, *RUNX3*	[2,304]

Gamma-delta 2	*TRDC*, *TYROBP*, *NCR3*, *CD160*, *FCER1G*	CD103+	Intra-epithelial, TYROBP+, natural IEL phenotype		BloodShare: 0.148 singleton: 0.52	*RORC*, *MAF*, *SOX4*, *TCF7*, *LEF1*, *TBX21*, *EOMES*	[2,343,344]
Gamma-delta 1	*TRDC*, *NCR3*, *NCR1*, *NCR2*, *CD160*, *FCER1G*	CD103+	Intra-epithelial, TYROBP-, induced IEL phenotype		BloodShare: 0.0417 singleton: 0.67	*RORA*, *RORC*, *RUNX3*, *BATF3*, *TBX21*, *EOMES*, *SOX5*, *STAT3*, *STAT1*, *STAT5*	[2,343,344]
ZNF683+	*ZNF683*, *ITGAE*, *CD160*, *ZEB2*, *GZMB*	CD103+	Re-activated tissue-resident memory cells : May contain second cluster of 'exhausted' cells expressing CD160– not visible because of low cell numbers. Non	Depleted in UC compared to HC Increased in CC	BloodShare: 0.0818 singleton: 0.51	*PRDM1*, *TBX21*, *EOMES*, *ZNF683*, *RUNX3* PMID: **26256443** **PMID: 24839135**	[304,345–361]

			CD160 Have direct anti-neoantigen effects in PD-1 immunotherapy, enriched in neoantigen areas (direct responders to antigen rather than bystanders). May downregulate HOBIT and re-enter circulation if antigen re-challenge. May maintain cytotoxicity. However, patterns in humans and mice are different.			TF regulatory network analysis showed higher activity of T-cell terminal differentiation TFs for CC in *ZNF683* cells *ZEB1*, *GABPA*, *GABPB1*, *SPIB*, *PBX3*, *TAF7*, *PAX5*, *YBX1*, *ZNF639*	
CD8+ HAVCR2+	*HAVCR2*, *LAYN*, *EOMES*,	CD8+ PD1+	Exhausted phenotype CD8+ cells	Enriched in UC and	BloodShare: 0.0663	*EOMES*, *TBX21*, *PRDM1*,	[362–364]

	PDCD1, *CTLA4*, *TNFRSF9*, *GZMH*, *GZMB*		PMID: 31624246	CC over HC Enriched in CC over UC	singleton: 0.70	*BATF*, *STAT1*, *RUNX3*, *TCF3*, *TCF4*, *TCF7*, *TCF21*, *FOXB1*, *BCL3*	
IFN Response	*IFI44L*, *ISG15*, *IFI6*, *MX1*, *ISG20*, *OAS1*		Strong interferon response signature, some sample-specificity		BloodShare: 0.0313 singleton: 0.81	*IRF1*, *IRF2*, *IRF3*, *IRF4*, *IRF5*, *IRF7*, *IRF8*, *IRF9*, *STAT2*, *STAT1*	
Th17 2	*IL22*, *IL17A*, *IL26*, *IL21*, *IL23R*, *IFNG*, *PDCD1*, *CTLA4*, *NCR3*, *CCR6*, *CCL20*	CD4+ PD1+	Inflammation-counterpart to Th17 cluster	Increased in both CC and UC over HC, no CC-UC difference	BloodShare: 0.0142 singleton: 0.87	*JUNB*, *ATF7*, *BATF3*, *FOXP3*, *RORC*, *STAT3*, *RORA*, *RARA*, *RUNX2*	[301,335,336]
Tc17	*IL22*, *IL17A*, *IL26*, *IL21*, *IL23R*, *IFNG*,	CD8+ PD1+	Inflammation-counterpart to Th17 cluster, CD8+, transcriptionally		BloodShare: 0.0361 singleton: 0.69	*STAT3*, *RORC*, *BATF3*, *MAF*, *PRDM1*, *IRF4*, *AHR*,	[301,305]

	PDCD1, *CTLA4*, *NCR3*, *CCR6*, *CCL20*		very similar to Th17 2			*STAT5*, *GATA3*	
GZMK+ Effector 1	*GZMK*, *IFNG*, *KLRG1*, *EOMES*, *GZMH*, *GZMB*,	CD103- CD8+		Increased in both UC and CC over HC, but not significant	BloodShare: 0.259 singleton: 0.56	*EOMES*, *EGR1*, *EGR2*, *EGR3*, *EGR4*, *IRF7*, *IRF8*, *IRF2*, *TBX21*	[2]
FGFBP2+ Effector	*FGFBP2*, *CX3CR1*, *IFNG*, *KLRG1*, *EOMES*, *GZMH*, *GZMB*	CD103- CD8+	Circulating, vessel-confined CD8+ effector cells.	Enriched in CC over HC (and over UC, but not significant)	BloodShare: 0.592 singleton: 0.35	*KLF2*, *KLF3*	[2,306,307]
MAIT	*TRAV1-2*, *TRAJ33*, *ZBTB16*, *KLRG1*, *SLC4A10*, *NCR3*	CD8+	Mucosal-associated invariant T-cells		BloodShare: 0.177 singleton: 0.76 Semi-invariant TCR Alpha Chain	*RORC*, *RORA*, *TBX21*, *EOMES* (Eomes), *IKZF1* (Helios), *EGR1*, *ELK3*, ***ZBTB16,*** *ZBTB7A*,	[308,309]

						ZBTB7B, *ZBTB17*, *ZBTB14*, *ZBTB33*	
Naïve 1	*LEF1*, *CCR7*, *SELL*	CD4+			BloodShare: 0 singleton: 0.99	*KLF2*, *KLF3*, *KLF5*, *FOXP1*, *TCF7*, *LEF1*, *BACH2*, *IKZF1*	[2]
Tfh	*TOX2*, *PDCD1*, *CTLA4*, *CXCR5*, *CXCL13*	CD4+ PD1+	Follicular helper cells	Enriched in UC over HC and CC	BloodShare: 0.00204 singleton: 0.95	*FOXP1*, *FOXP3*, *STAT3*, *BCL6*	[301]
Cycling	*MKI67*, *CTLA4*, *STMN1*	CD8+ and CD4+	A mix of CD4+ and CD8+ actively proliferating cells in G2M and S-phase	Increased in CC over UC and HC, increased in UC over HC	BloodShare: 0.0429 singleton: 0.75	*E2F8*, *E2F2*, *E2F1*, *E2F7*, *E2F4*, *E2F5*	[310,311]
Tregs	*FOXP3*, *IL10*, *CTLA4*, *TNFRSF9*, *CCR6*, *IL1R1*,	CD4+	Regulatory T-cells		BloodShare: 0.00619 singleton: 0.93	*STAT5*, *BATF*, *FOXP3*, *IRF4*, *JUNB*, *FOXO1*, *TCF3*, *LEF1*,	[301–303]

	IL1R2, *LRRC32*					*PRDM1,* *RUNX1*	

3.3 High resolution cluster summary of PBMC T cells

Cluster	Key Marker Genes	Key Antibodies	Notes	TCR Analysis	DEGs	Refs
CD4+ Memory 3	*CCR6* *RORa* *CTSH* *TIMP1* *ANXA1*	CD4+ CD45RO	Potentially Th17 phenotype effector/memory cells	TissueShare: 0.0120 Singleton: 0.94		[187,365–368]
CD4+ Memory 2	*CCR10* *CCR4*	CD4+ CD45RO	Possible skin-homing CD4 Th1-Th2 pleiotropic effector memory phenotype	TissueShare: 0.0150 Singleton : 0.95	↑ HLA-C in UC (vs CC)	[187,369–372]
CD4+ Memory 1	*KRT1* *TCF7* *SELL*	CD4+ CD45RO	T central memory cells	TissueShare: 0.00190 Singleton : 0.97	↑FOS/JUN (AP-1) pathway in CC (vs UC) – induces proliferation but also PD-1 ↑HLA-C, HLA-A and STAT1 in UC (vs CC)	[188]
CD4+ CD8- FGFBP2+	*FGFBP2* *CX3CR1*	CD4+ CD45RO	Terminally differentiated	TissueShare: 0.197		[373–376]

	HOPX *GZMB* *GNLY* *TBX21* *KLRG1*		CD4 effector memory cells. Per literature, in UC, become IELs on trafficking to tissue. However there is considerable donor variation in the population	Singleton : 0.43		
FGFBP2+ Effector 1	*FGFBP2* *CX3CR1* *GZMH* *ADGRG1* *SPON2* *ZEB2*	CD8+ CD45RO	Cytotoxic cells, with adhesion and migration markers, Found to correlate with clinical response in Checkpoint therapy	TissueShare: 0.115 Singleton : 0.54	↑ IFITM3 (activation marker) in CC (vs UC) ↑ HLA-C, KLRG1, IRF1, CCL5 (markers of terminal differentiation) in UC (vs CC). Also however express ↑ RAP1B which is associated with reduced migration into colon.	[306,377,378]

GZMK+ Effector 2	*GZMK* *HLA-DRA* *FABP5* *CD160*	CD8+ CD45RO	Highly active Effector T cell population – because expressing high amounts of HLA type 1, correlated with autoimmune disease in SLE	TissueShare: 0.194 Singleton : 0.54		[187,188,379]
Gamma-delta	*TRDC* *FCGR3A* *TYROBP*			TissueShare: 0.0636 Singleton : 0.74		
GZMK+ Effector 1	*GZMK* *XCL2* *CCL4L2* *CD160*	CD8+ CD45RO	Effector population Expressing some exhaustion markers	TissueShare: 0.0821 Singleton : 0.63	↑ CTSW in CC (vs UC) – marker of cytotoxicity.	[188]
MAITs	*TRAV1-2, TRAJ33, ZBTB16* *SLC4A10*	CD8+ CD45RO	Represent 5-10% of the circulating population	TissueShare: 0.0368 Singleton : 0.76	↑HLA-C in UC (vs CC) ↑ TXNIP (Activation marker) in CC (vs UC)	[380]
Memory CD8+	*TCF7* *IL7R* *FXYD2*	CD8+ CD45RO	Potentially Central Memory / Resting CD8 T	TissueShare: 0.0458 Singleton : 0.87		[187,188,300]

	CCR7		cells / Gradient from Naïve population			
Naïve CD8+	*CCR7* *LEF1*	CD8+		TissueShare: 0 Singleton : 0.99	↑ HLA-C in UC, ↑CD69 in CC vs UC	
Cell Cycle	*MKI67* *STMN1*	CD4+ CD8+ CD45RO		TissueShare: 0.0223 Singleton : 0.95		
Tregs	*FOXP3*	CD4+ CD45RO		TissueShare: 0.00558 Singleton : 0.98		
Naïve CD4+	*CCR7* *LEF1*	CD4+		TissueShare: 0.000416 Singleton : 0.99	↑CD69 and JUNB in CC (vs UC) ↑ HLA-C in UC	

Appendix C - Materials

Chapter 1

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Graphpad Prism v8.1	Graphpad	www.graphpad.com

Chapter 2

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
BV 711- Conjugated CD3 antibody	BD	Cat 740807; RRID:AB_2740470
APC-R700- Conjugated CD8 antibody	BD	Cat 565192; RRID:AB_2739104
FITC- CD45 Antibody	Miltenyi	Cat 130-113-679; RRID:AB_2726220
BV785- EPCAM antibody	Biolegend	Cat 324237; RRID:AB_2632936
PE - Anti IL20RA antibody	RnD	Cat FAB11762P; RRID:AB_663921
AF647 - Anti IL10RB antibody	RnD	Cat FAB874R-100UG;
PE - IL26 antibody	RnD	Cat IC13751P; RRID:AB_10640282
APC anti-human CD3	Biolegend	Cat 317318; RRID:AB_1937212
IgG1,k isotype control (in-vivo experiment, mouse)	Invitrogen	Cat 14-4714-85; RRID:AB_470112
Anti-IL26 antibody (in-vivo experiment)	Juntendo University	Clone 69-10
TotalSeq(TM)-C0045 anti-human CD4	Biolegend	Cat 344651; RRID:AB_2800921

TotalSeq(TM)-C0087 anti-human CD45RO	Biolegend	Cat 304259; RRID:AB_2800766
TotalSeq(TM)-C0088 anti-human CD279 (PD-1)	Biolegend	Cat 329963; RRID:AB_2800862
TotalSeq(TM)-C0101 anti-human CD335 (NKp46)	Biolegend	Cat 331941; RRID:AB_2800874
TotalSeq(TM)-C0143 anti-human CD196 (CCR6)	Biolegend	Cat 353440; RRID:AB_2810563
TotalSeq(TM)-C0145 anti-human CD103 (Integrin alphaE)	Biolegend	Cat 350233; RRID:AB_2800933
TotalSeq(TM)-C0146 anti-human CD69	Biolegend	Cat 310951; RRID:AB_2800810
TotalSeq(TM)-C0148 anti-human CD197 (CCR7)	Biolegend	Cat 353251; RRID:AB_2800943
TotalSeq(TM)-C0151 anti-human CD152 (CTLA-4)	Biolegend	Cat 369621; RRID:AB_2801015
TotalSeq(TM)-C0152 anti-human CD223 (LAG-3)	Biolegend	Cat 369335; RRID:AB_2814327
TotalSeq(TM)-C0169 anti-human CD366 (Tim-3)	Biolegend	Cat 345049; RRID:AB_2800925
TotalSeq(TM)-C0802 anti-human CD336 (NKp44)	Biolegend	Cat 325119; RRID:AB_2810484
TotalSeq(TM)-C0355 anti-human CD137 (4-1BB)	Biolegend	Cat 309839; RRID:AB_2800807
TotalSeq(TM)-C0801 anti-human CD337 (NKp30)	Biolegend	Cat 325219; RRID:AB_2800851
Totalseq C0251 Anti-Human Hashtag 1	Biolegend	Cat#394661, RRID:AB_2801031

Totalseq C-0252 Anti-human Hashtag 2	Biolegend	Cat#394663, RRID:AB_2801032
Totalseq C-0253 Anti-human Hashtag 3	Biolegend	Cat#394665, RRID:AB_2801033
Totalseq C-0254 Anti-human Hashtag 4	Biolegend	Cat#394667, RRID:AB_2801034
Totalseq C-0255 Anti-human Hashtag 5	Biolegend	Cat#394669,RRID:AB_2801035
DAPI (4′,6-diamidino-2-phenylindole) Solution	BD Pharminogen	Cat#564907
Human Trustain FcX (Fc Block)	Biolegend	Cat# 422302, RRID:AB_2818986
Human IL-26 AK155 Antibody	R&D	Cat AF1375-SP
Biological Samples		
Adult human colon resections, biopsies and blood samples	John Radcliffe Hospital NHS Foundation Trust	REC IDs: 18/WM/0237, GI 16/YH/0247 and IBD 09/H1204/30.
Chemicals, Peptides, and Recombinant Proteins		
Collagenase, Type 2, Filtered – CLSS-2	Worthington	LS004204
Cell Staining Buffer	Biolegend	Cat 420201
Lymphoprep	Serumwerk Bernburg	Cat#1858
Normal Goat Serum 2.5%	ImmPRESS Vector	Cat#30023
Recombinant human IL-26 dimer	R&D	Cat 1870 CF
CD14 MicroBeads (human)	Miltenyi	Cat 130-050-201
Critical Commercial Assays		
High sensitivity RNA Screen Tape, Buffer and Reagents (for	Agilent	Cat#5067-5579,5580 and 5581

use with Agilent 2200 TapeStation system)		
High-Capacity cDNA Reverse Transcription Kit with RNAse inhibitor	Thermofisher	Cat no 4374966
TaqMan Fast Advanced Master Mix-1 x 5 mL	Thermofisher	Cat no 4444557
RNAscope™ 2.5 HD Assay - BROWN	ACD Bio-techne	Cat no. 322300
RNEasy Mini kit	Qiagen	Cat no 74106
RNEasy plus Micro kit	Qiagen	Cat no 74034
Deposited Data		
All raw and processed next-generation sequencing data was deposited with GEO under accession nos. GSE148837 and GSE148505	CD3 CD8+ T cells from HC and UC inflamed biopsies	Corridoni, Antanaviciute, Gupta et al, 2020
Epithelial single-cell RNA seq data from Ulcerative colitis (n =3) and Health (n =3)		Parikh et al, 2019
Colonic mesenchymal dataset in health (n = 2 donors) and UC(n = 2 donors)		Kinchen et al, 2018
Experimental Models: Cell Lines		
HDLM-2 cell line	DMSZ GmbH, Germany	ACC 17
SW-480 cell line	ATCC	CCL-228
Experimental Models: Organisms/Strains		

hIL-26Tg mice (Black 6 WT mice with BAC transgene expressing IL26)	T. Aune group	

Chapter 3

REAGENT or RESOURCE	**SOURCE**	**IDENTIFIER**
	Antibodies	
E-cadherin rabbit anti-human antibody	Cell Signalling	Cat#3195, RRID:AB_2291471
CD103 mouse anti-human antibody	Abcam	Cat#ab238010 [ITGAE/2063]
DAPI (4′,6-diamidino-2-phenylindole) Solution	BD Pharminogen	Cat#564907
FITC-Conjugated anti-EPCAM	Miltenyi	Cat#130080301, RRID:AB_244192
APC-conjugated anti-CD45	Miltenyi	Cat#130113676, RRID:AB_2726217
Human Trustain FcX (Fc Block)	Biolegend	Cat# 422302, RRID:AB_2818986
PEDazzle-conjugated anti-CD3	Biolegend	Cat# 300449, RRID:AB_2563617
AF488 Goat Anti-Rabbit Secondary Antibody (Cross Adsorbed)	Thermo Fisher Scientific	Cat# A32731, RRID:AB_2633280
AF647 Goat Anti-Mouse Secondary Antibody (Cross Adsorbed)	Thermo Fisher Scientific	Cat# A-21235, RRID:AB_2535804
Totalseq C0251 Anti-Human Hashtag 1	Biolegend	Cat#394661, RRID:AB_2801031
Totalseq C-0252 Anti-human Hashtag 2	Biolegend	Cat#394663, RRID:AB_2801032
Totalseq C-0253 Anti-human Hashtag 3	Biolegend	Cat#394665, RRID:AB_2801033
Totalseq C-0254 Anti-human Hashtag 4	Biolegend	Cat#394667, RRID:AB_2801034
Totalseq C-0255 Anti-human Hashtag 5	Biolegend	Cat#394669,RRID:AB_2801035

Totalseq C-0046 Anti-human CD8	Biolegend	Cat#344753, RRID:AB_2800922
Totalseq C-0045 Anti-human CD4	Biolegend	Cat#344651, RRID:AB_2800921
Totalseq C-0087 Anti-human CD45RO	Biolegend	Cat#304259, RRID:AB_2800766
Totalseq C-0088 Anti-human PD1	Biolegend	Cat#329963, RRID:AB_2800862
Totalseq C-0101 Anti-human CD335 (NKp46)	Biolegend	Cat#331941, RRID:AB_2800874
Totalseq C-0143 Anti-human CD196 (CCR6)	Biolegend	Cat#353440, RRID:AB_2810563
Totalseq C-0152 Anti-human CD223 (LAG-3)	Biolegend	Cat#369335, RRID:AB_2814327
Totalseq C-0169 Anti-human CD366 (Tim-3)	Biolegend	Cat#345049, RRID:AB_2800925
Totalseq C-0355 Anti-human CD137 (4-1BB)	Biolegend	Cat#309839, RRID:AB_2800807
Totalseq C-0151 Anti-human CD152 (CTLA-4)	Biolegend	Car#369621, RRID:AB_2801015
Totalseq C-0145 Anti-human CD103	Biolegend	Cat#350233, RRID:AB_2800933
Totalseq C-0053 Anti-human CD11c	Biolegend	Cat#371521, RRID:AB_2801018
Totalseq C-0161 Anti-human CD11b	Biolegend	Cat#301359, RRID:AB_2800732
BV421-conjugated anti-KI67	Biolegend	Cat#350505, RRID:AB_10896915
Zombie Aqua fixable viability kit	Biolegend	Cat# 423101
FITC-conjugated anti-PD1	Biolegend	Cat #329904, RRID:AB_940479
BV605-conjugated anti-CD3	Biolegend	Cat #300459,RRID:AB_2564379
BV785-conjugated anti-CD8	Biolegend	Cat #344739, RRID:AB_2566201
FITC-conjugated anti-PD1	Biolegend	Cat# 329904, RRID:AB_940479

FITC-conjugated anti-human mouse IGG1 Isotype control	Thermofisher	Cat#11-4714-42; RRID:AB_10596964
PE-conjugated anti-CCR6	Biolegend	Cat#353409, RRID:AB_10915968
PE-Cy7-conjugated anti-CD103	Biolegend	Cat#350211, RRID:AB_2561598
APC-Cy7- conjugated anti-CXCR5	Biolegend	Cat#356925, RRID:AB_2562592
AF647-conjugated mouse IGG1 isotype control	Biolegend	Cat#400135, RRID:AB_2832978
Non-competitive PD1 binding antibody	Davis Group, WIMM	Not commercially available, validated in-house, data in paper
CD3 rabbit anti-human antibody	Cell Signaling Technology	Cat# 85061, RRID:AB_2721019)
IGG4 mouse anti-human antibody	Bio-Rad	Cat# MCA2098G, RRID:AB_323685
iNOS2 mouse anti-human antibody	R&D Systems	Cat# MAB9502, RRID:AB_2152874
FABP1 rabbit anti-human antibody	Sigma-Aldrich	Cat# HPA028275, RRID:AB_10600909
Neutrophil elastase/ELA2 mouse anti-human antibody	Novus Biologicals	Cat# MAB9167
CD163 mouse anti-human antibody	Novus Biologicals	Cat# NB110-40686, RRID:AB_714951
FOXP3 rabbit anti-human antibody	Atlas Antibodies	Cat# HPA045943, RRID:AB_2679508
E-cadherin mouse anti-human antibody	Cell Signaling Technology	Cat# 14472, RRID:AB_2728770
Cleaved caspase-3 rabbit anti-human antibody	Cell Signaling Technology	Cat# 9664, RRID:AB_2070042
Biological Samples		
Adult human colon resections, biopsies and blood samples	John Radcliffe Hospital NHS Foundation Trust	REC reference(s): PRISE: 18/LO/0412, GI Biobank: 16/YH/0247, IBD Biobank: 09/H1204/30, TIP: 18/WM/0237

		Sample overview detailed Supplementary data table in Mendeley. Available on Publication. Temporary Link: https://data.mendeley.com/datasets/7z8yx644hb/draft?a=e8b9e179-8fb7-448e-a481-7cd4c090f71a
Chemicals, Peptides, and Recombinant Proteins		
OCT Embedding matrix for frozen sections	CellPath	Cat#KMA-0100-00A
Isopentane (2-Methylbutane)	Sigma	Cat#277258-1L
RPMI-1640 medium	Sigma	Cat#R8758-500ml
Dulbecco's Modified Eagle's Medium (DMEM)	Sigma	Cat#D5796-500ML
Penicillin-Streptomycin	Sigma	Cat#P0781-100ML
HEPES Buffer Solution (1M)	Gibco	Cat#15630-056
Fetal Calf Serum / Fetal Bovine Serum	Sigma	Cat#F9665-500ML
Lymphoprep	Serumwerk Bernburg	Cat#1858
Dimethyl Sulfoxide	Sigma	Cat#D8418-100mls
MEM Non-Essential Amino Acids	Sigma	Cat#M7145-100ml
Sodium Pyruvate Solution	Sigma	Cat#S8636-100mls
Vectashield Mounting Medium with DAPI	Vector	Cat #H-1200
Ultrapure 0.5M EDTA, ph8.0	Invitrogen	Cat#15575-038
HBSS medium	Lonza	Cat#10-543F
Pierce DTT (Dithiothretitol)	Thermo Scientific	Cat#A39255
Bovine Serum Albumin	Sigma	Cat#A7906-100G

Tryple Express	Gibco	Cat#12605-028
Lamina Propria Dissociation Kit, Mouse	Miltenyi	Cat#130-097-410
Mayer's Hematoxylin (used for ST)	Dako	Cat#S3309
Dako Bluing Buffer (used for ST)	Dako	Cat#CS702
Eosin Y solution	Sigma	Cat#HT110216-500ml
Normal Goat Serum 2.5%	ImmPRESS Vector	Cat#30023
Trueview Autofluorescence Quenching Kit	Vector	SP-8400
Phosphate Buffered Saline (PBS)	Oxoid Ltd or Sigma (experiment dependent)	Cat#BR0014G / D8537-500ML
CryostorCS10	Sigma	Cat#C2874-100ML
APC conjugation kit (Lightning-Link)	Abcam	Cat#ab201807-300ug
Dynabeads™ Human T-Activator CD3/CD28 for T Cell Expansion and Activation	Thermofisher	Cat#11161D
NIVOLUMAB 40mg in 4ml INJECTION "Opdivo" (Packs of 1 vial)	Oxford University Hospitals Pharmacy	Opdivo
Fixation/Permeabilization Kit for FACS (Cytofix/Cytoperm)	BD biosciences	Cat #554714, AB_2869008
Critical Commercial Assays		
Visium Spatial Tissue Optimization Slide	10x Genomics	Cat#1000191
Visium Spatial Gene Expression Slide	10x Genomics	Cat#1000185
KAPA SyBR FAST qPCR kit	Kapa biosystems	Cat # KK4600
10x Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3	10x Genomics	Cat#1000075

KAPA library quant kit (illumina) uiversal qPCR mix	Kapa biosystems	Cat# KK4824
QuBit dsDNA HS Assay Kit (used with QuBit 3.0)	Invitrogen	Cat#Q32851
High sensitivity RNA Screen Tape, Buffer and Reagents (for use with Agilent 2200 TapeStation system)	Agilent	Cat#5067-5579,5580 and 5581
10x Chromium Single Cell 5' GEM, Library and Gel Bead	10x Genomics	Cat#1000006
RNeasy Plus Micro Kit	Qiagen	Cat#74034
Novaseq 6000 S4 150bp PE reads	Illumina	Cat#20012866
Nextseq 500/550 Hi Output kit v2.5	Illumina	Cat# 20024907
High Sensitivity DNA reagents (Used with Agilent 2100 Bioanalyser system)	Agilent Technologies	Cat#5067-4626
Deposited Data		
Slide A1 and A2 (Spatial Transcriptomics), Raw data	Spatiotemporal analysis of human intestinal development at single-cell resolution (Cell, Fawkner-Corbett et al, DOI: 10.1016/j.cell.2020.12.016)	GEO ID: GSE158328 Mendeley Data (H&E Images): 10.17632/gncg57p5x9.2
Software and Algorithms		
FlowJo v10.7.1	FlowJo	FlowJo.com
Graphpad Prism v9.1.2	Graphpad	www.graphpad.com
Las X Version 3.7.4.23463	Leica Microsystems CMS GmbH	www.leica.com
QuPath v0.2.3	Github (open source)	https://qupath.github.io
Zen Blue Edition v3.3.89.0000 (ZEN lite)	Carl Zeiss Microscopy GmbH	www.zeiss.com

Visiopharm Integrator System (VIS) platform v 2019.07.3	Visiopharm	www.visiopharm.com
Biorender (Graphical Abstract)		Created with BioRender.com
fastQC version 0.11.9	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/	
cellranger version 6.0.1	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger	
spaceranger version 1.2.2	https://support.10xgenomics.com/spatial-gene-expression/software/pipelines/latest/what-is-space-ranger	
bcl2fastq version 2.20.0.422	https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html	
R package DropletUtils version 1.8.0	https://bioconductor.org/packages/release/bioc/html/DropletUtils.html	
R package Seurat version 4.0.1	https://satijalab.org/seurat/	
R package Harmony version 1.0	https://github.com/immunogenomics/harmony	
R package Monocle3 version 0.2.3.0	https://cole-trapnell-lab.github.io/monocle3/	
R package ggplot2 version 3.3.2	https://ggplot2.tidyverse.org/	
R package DESeq2 version 1.28.1	https://bioconductor.org/packages/release/bioc/html/DESeq2.html	

R package ggpubr version 0.4.0	https://cran.r-project.org/web/packages/ggpubr/index.html	
R package miloR version 0.99.19	https://github.com/MarioniLab/miloR	
R package AUCell version 1.10.0	http://bioconductor.org/packages/release/bioc/html/AUCell.html	
pySCENIC version	https://github.com/aertslab/pySCENIC	
R package CellChat version 1.0.0	https://github.com/sqjin/CellChat	
R package SPOTlight version 0.1.0	https://github.com/MarcElosua/SPOTlight	
R package RCTD	https://github.com/dmcable/RCTD	
R package igraph version 1.2.5	https://cran.r-project.org/web/packages/igraph/index.html	
R package ggraph version 2.0.3	https://cran.r-project.org/web/packages/ggraph/index.html	
R package SingleCellSignalR version 1.0.0	http://www.bioconductor.org/packages/release/bioc/html/SingleCellSignalR.html	
R package MAST version 1.14.0	https://www.bioconductor.org/packages/release/bioc/html/MAST.html	
R package clusterProfiler version 3.16.0	https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html	

R package org.Hs.eg.db version 3.11.4	https://bioconductor.org/packages/release/data/annotation/html/org.Hs.eg.db.html	
TRUST4	https://github.com/liulab-dfci/TRUST4	
R package divo version 1.0.1	https://cran.r-project.org/web/packages/divo/index.html	
GLIPH2	https://doi.org/10.1038/s41587-020-0505-4	
R package jcolors version 0.0.4	https://cran.r-project.org/web/packages/jcolors/index.html	
R package immunarch version 0.6.6	https://cran.r-project.org/web/packages/immunarch/index.html	
R package venneuler 1.1-0	https://cran.r-project.org/web/packages/venneuler/index.html	
VDJtools	https://github.com/mikessh/vdjtools	
Other		
Haematoxylin and Eosin images from all Spatial Transcriptomic sections	Zeiss Axioscanner	Supplementary data in Mendeley. Available upon publication
Immunofluorescence Images	Leica Widefield Microscope scanner Zeiss Axioscanner	Supplementary data table in Mendeley. Available upon publication

Appendix D

Patient information

Patient ID Code	Run	Experiment	Chemistry	CITE-Seq	Hashing	Disease	Status	Anatomical compartment	Indications/Past Medical History	UCEIS	Gender	Age	Lag from Immunotherapy (CC)	Duration of Inflammation (UC&CC_I)	Medication
GI 6275	Slide 2	Visium(ST)	N/A	N/A	B6	CC (Dual Therapy)	Inflamed	Descending	Diarrhoea. Melanoma	UCEIS 4/8	F	74	33	21	Prednisolone 60mg od, Flecainide, Warfarin, Atorvastatin,Bisoprolol, Alendronic acid
GI 6278	Slide 3	Visium(ST)	N/A	N/A	B11	CC (Dual Therapy)	Inflamed	Sigmoid	Melanoma, Asthma		M	30		4	IV methylprednisolone x 3, Ramipril, Nystatin. dose
PRISE 31	Slide 1	Visium(ST)	N/A	N/A	B3	CC (Dual Therapy)	Inflamed	Sigmoid	Diarrhoea, Melanoma, Gout, Sciatica	Mild Inflammation	M	47		13	Infliximab x 1 dose, 2x dose methylprednisolone, Carbimazole 40mg OD.
GI 3667 (#1)	Slide 4	Visium(ST)	N/A	N/A	C4	CC (Monotherapy)	Inflamed	Simgoid	Melanoma,	Mild Inflammation	F	46		11	Nil
GI 6274	Slide 4	Visium(ST)	N/A	N/A	C3	CC (Monotherapy)	Inflamed	Sigmoid	Melanoma, HTN, Knee replacement	UCEIS 3/8	F	75	20	41	Prednisolone 5mg (reducing regimen), Amitryptypline, Atenolol, Omeprazole, Timodine, Zapain, Furoseimde, Adcal D3
PRISE 14	Slide 1	Visium(ST)	N/A	N/A	B2	CC (Monotherapy)	Inflamed	Descending and Sigmoid	Diarrhoea, Melanoma	UCEIS 2/8	M	72	64	40	Prednisolone 30mg OD, Omeprazole
PRISE 23	Slide 2	Visium(ST)	N/A	N/A	B7	CC (Monotherapy)	Inflamed	Sigmoid	Diarrhoea. Bladder Cancer	UCEIS 2/8	M	74		25	Lansoprazole, Naproxen, Mometasone, Morphine
GI 6925	Slide 3	Visium(ST)	N/A	N/A	B10	Healthy	Not inflamed	Sigmoid pairs	FIT test positive. Osteoporosis, Stroke, Hypertension (HTN), Lumbar spine fixation.		F	66		0	Clopidogrel, Sertraline, Cetirizine, Amlodipine, Atorvastatin, Amitryptyline, Morphine, Adcal D3
TIP 346	Slide 5	Visium(ST)	N/A	N/A	A1&A2	Healthy	Not inflamed	Descending	Anterior resection for Rectal cancer (Mid rectum),		F	53		0	Nil

TIP 517	Slide 4	Visium(ST)	N/A	N/A	C5	Healthy	Non inflamed	Sigmoid	Previous Colonic Polyp		F	41		0	Nil
BB4652	Slide 4	Visium(ST)	N/A	N/A	C2	UC	Inflamed	Sigmoid	Ulcerative Colitis, Left Ventricular Failure, Implantable Cardiac Defib, Atrial Fibrillation (AF)	UCEIS 5/8	M	65		35	Vedolizumab, Apixaban, Atorvastatin, Ramipril, Bisoprolol
GI 4980	Slide 3	Visium(ST)	N/A	N/A	B13	UC	Inflamed	Sigmoid -> Rectum	Acute severe colitis	UCEIS 5/8	F	60		21	1x Dose hydrocortisone , Ciprofloxacin
GI 6658	Slide 3	Visium(ST)	N/A	N/A	B12	UC	Inflamed	Sigmoid	Ulcerative Colitis (UC)	UCEIS 5/8	M	55		124	Nil
GI 6967	Slide 2 x 2 sections	Visium(ST)	N/A	N/A	B8 & B9	UC	Inflamed	Left sided	UC.		F	30		120	Azathioprine, Mesalazine, Ferric Carboxymaltose
GI 6264	Slide 1	Visium(ST)	N/A	N/A	B4	UC	Inflamed	Sigmoid/Descending	Ulcerative Colitis	UCEIS 3/8	F	53		33	Pentasa.
GI 6266	Slide 1	Visium(ST)	N/A	N/A	B5	UC	Inflamed	Sigmoid/Descending	Ulcerative Colitis	UCEIS 3/8	M	33		21	Seretide.
GI 3031 (SA)	CD45(2)	Biopsy CD45+	3' 10X scRNA	No	N/A	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Melanoma	Mild Inflammation	M	72		25	Prednisolone 60mg, Finasteride, Tamsulosin, Morphine, Paracetamol.
GI 3126 (SC)	CD45(2)	Biopsy CD45+	3' 10X scRNA	No	N/A	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Renal Cancer, AF, Nephrectomy	Mild Inflammation	M	75		40	Warfarin, Verapamil, Levothyroxine
GI 3060 (SB)	CD45(2)	Biopsy CD45+	3' 10X scRNA	No	N/A	CC (Monotherapy)	Not inflamed	Sigmoid/Descending	Melanoma		M	72		0	Nil
GI 3236 (SD)	CD45(2)	Biopsy CD45+	3' 10X scRNA	No	N/A	CC (Monotherapy)	Not Inflamed	Sigmoid/Descending	Renal Cancer		M	61		0	Prednisolone 60mg for 24 hrs
GI 3369 (GI 3726)	CD45(1)	Biopsy CD45+	3' 10X scRNA	No	N/A	Healthy	Not inflamed	Pan-colonic	PR Bleed		M	23		0	Nil
GI 3409	CD45(1)	Biopsy CD45+	3' 10X scRNA	No	N/A	Healthy	Not inflamed	Sigmoid	Change in bowel habits		F	48		0	Nil
GI 3410	CD45(1)	Biopsy CD45+	3' 10X scRNA	No	N/A	Healthy	Not inflamed	Descending	Surveillance polyps, Hypertension		F	58		0	Amitriptyline 10mg Bendroflumethiazide 2.5mg Bisoprolol 2.5mg Simvastatin 40mg Omeprazole 40mg

IBD 2218	CD45(1)	Biopsy CD45+	3' 10X scRNA	No	N/A	UC	Inflamed and Non inflamed	Sigmoid/Rectum	UC, Anxiety	UCEIS 5/8	M	64		40	Pentasa 1gr mercaptopurine 50mg Cetrizine 10mg Prednisolone 20mg Sertraline
IBD 2368	CD45(1)	Biopsy CD45+	3' 10X scRNA	No	N/A	UC	Inflamed and Non inflamed	Sigmoid	Colitis, PR bleeding	UCEIS 4/8	M	46		Unclear	Omeprazole, Olsalazine, Mirtazapine, Sertraline, Salofalk Suppositories
IBD 3552	CD45(1)	Biopsy CD45+	3' 10X scRNA	No	N/A	UC	Inflamed and Non inflamed	Sigmoid/Rectum	UC	UCEIS 5/8 (Proctitis)	F	31		27	Salofalk Laxido Rigevidon
IBD 3736	CD45(1)	Biopsy CD45+	3' 10X scRNA	No	N/A	UC	Inflamed and Non inflamed	Sigmoid/Rectum	UC flare, COPD	UCEIS 5/8	M	67		25	Pentasa Prednisolone Ventolin Laxido Amoxicillin
GI 6275	CD3 - Run 1&3	Biopsy&Blood CD3+	5' 10X scRNA	Yes	HTO_2	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma, Ischaemic Heart Disease, HTN, AF	UCEIS 4/8	F	74	33	21	1 week of Prednisolone 60mg, Flecainide, Warfarin, Atorvastatin,Bisoprolol, Alendronic acid
PRISE 14	CD3 - Run 2&3	Biopsy&Blood CD3+	5' 10X scRNA	Yes	HTO_3	CC (Dual Therapy)	Inflamed	Sigmoid	Melanoma	UCEIS 5/8	M	72	64	40	Steroids and Infliximab
GI 6274	CD3 - Run 1&3	Biopsy&Blood CD3+	5' 10X scRNA	Yes	HTO_1	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Melanoma, Arthritis, HTN	UCEIS 3/8	F	75	20	41	Amitryptyline, Atenlol, Zapain, Furosemide, Omeprazole.
GI 6286	CD3 - Run 2&3	Biopsy&Blood CD3+	5' 10X scRNA	Yes	HTO_5	Healthy	Non inflamed	Sigmoid->Rectum	Epilepsy, Colonic Polyps		M	58		0	Epilim, Clonazepam
TIP 376	CD3 - Run 4&5	Biopsy&Blood CD3+	5' 10X scRNA	Yes	HTO_1	Healthy	Non inflamed	Descending	PR Bleeding, Loose stools. Mild Asthma		F	65		0	Nil
TIP 489	CD3 - Run 4&5	Biopsy&Blood CD3+	5' 10X scRNA	Yes	HTO_3	Healthy	Non Inflamed	Sigmoid	FIT test positive. Chronic Obstructive Pulmonary Disease (COPD), Diabetes, High Cholesterol		M	71		0	Salbutamol, Metformin, Atorvastatin,Gliclazide, Apixaban,Alogliptin
GI 4980	CD3 - Run 1&3	Biopsy&Blood CD3+	5' 10X scRNA	Yes	HTO_3	UC	Inflamed	Sigmoid	Acute Severe colitis	UCEIS 5/8	F	60		21	1x Dose hydrocortisone , Ciprofloxacin

GI 6929	CD3 - Run 2&3	Biopsy&Blood CD3+	5' 10X scRNA	Yes	HTO_4	UC	Inflamed	Distal transverse, Rectum	Ulcerative Colitis	UCEIS 4/8	M	42		26	Steroids reducing regimen stopped 1 week prior, Mesalazine
IBD1464	CD3 - Run 4&5	Biopsy&Blood CD3+	5' 10X scRNA	Yes	HTO_2	UC	Inflamed	Sigmoid->Rectum	Flare of established UC	UCEIS 5/8	M	58		168	Prednisolone & Mesalazine, reducing regimen
GI 4051	S58(CD45/Str)	Epi+CD45/Str	5' 10X scRNA	Yes	HTO_5	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma, Checkpoint induced dermatitis	UCEIS 4/8	F	63		66	Infliximab, Prednisolone 60mg OD, Omeprazole, Co-Trimoxazole, Levothyroxine, Dermovate.
GI 4442	S32(CD45/Str)	Epi+CD45/Str	5' 10X scRNA	Yes	HTO_4	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Renal Cell, AF, prior Pulmonary Embolus, Previous Falls, Seminoma, Prostate Cancer (Treated)	UCEIS 2/8	M	82	23	11	Prednisolone 10mg OD, Dalteparin, Risedronate, Bisoprolo, Paracetamol.
GI 3662	S31(Epi), S32(CD45/Str)	Epi+CD45/Str	5' 10X scRNA	Yes	HTO_5	CC (Dual Therapy)	Not Inflamed	Sigmoid/Descending	Melanoma		F	29	28	0	Prednisolone 30mg OD, Ranitidine, Propranolol PRN
GI 4449	S58(CD45/Str)	Epi+CD45/Str	5' 10X scRNA	Yes	HTO_4	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Lung Cancer	UCEIS 5/8	M	61	71	65	Infliximab, Prednisolone (>2 weeks), Methylprednisolone, Omeprazole
TIP 332	S31(Epi), S32(CD45/Str)	Epi+CD45/Str	5' 10X scRNA	Yes	HTO_1	Healthy	Not Inflamed	Descending	Rectal Bleeding, Irritable bowel syndrome (IBS)		F	36		0	Paracetamol PRN
TIP 309	S58(CD45/Str)	Epi+CD45/Str	5' 10X scRNA	Yes	HTO_1	Infectious Colitis	Inflamed	Pan-colonic	Anal Fissure		M	53		Unclear	Laxido
TIP 299	S58(CD45/Str)	Epi+CD45/Str	5' 10X scRNA	Yes	HTO_2	Microscopic Colitis	Inflamed	Descending	Asthma, B12 deficiency, Anxiety/Depression		F	28		Unclear	Amitryptyline, Combined Contraceptive Pill
GI 4379	S58(CD45/Str)	Epi+CD45/Str	5' 10X scRNA	Yes	HTO_3	UC	Inflamed	Sigmoid	UC, Primary Biliary Cholangitis (Antibody negative), Meniere's disease	UCEIS 3/8	F	66		Unclear	Ursofalk

IBD 3778	S31(Epi), S32(CD45/Str)	Epi+CD45/Str	5' 10X scRNA	Yes	HTO_2	UC	Inflamed	Transverse (4 pairs)	UC	UCEIS 6/8	M	38		34	Omeprazole, azathioprine
IBD 3778	S31(Epi), S32(CD45/Str)	Epi+CD45/Str	5' 10X scRNA	Yes	HTO_3	UC	Not Inflamed	Sigmoid (4 pairs)	UC		M	38		0	Omeprazole, azathioprine
GI 3152	S13(CD45/Str)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_4	CC (Dual Therapy)	Inflamed	Sigmoid/Rectu m	Melanoma, Hypertension, Osteoarthritis, Hypercholeste rolaemia	UCEIS 6/8	M	68	14	8	Prednisolone 60mgOD, Amlodipine, Ramipril, Simvastatin, Omeprazole
GI 4059	S2(Epi), S3(CD45/Str), S1(PBMC)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_4	CC (Dual Therapy)	Inflamed	Sigmoid/Desc ending	Melanoma, Atrial Flutter, Hypertension	UCEIS 5/8	M	70		29	IV MethylPrednis olone (2 days), Prednisolone (>2 weeks), Ramipril, Warfarin
GI 4069	S37(Epi), S38(CD45/Str) , S43(PBMC)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_3	CC (Dual Therapy)	Inflamed	Sigmoid/Rectu m	Melanoma, Asthma, Hypercholeste rolaemia	UCEIS 4/8	M	71	11	10	IV Methylprednis olone, Prednisolone 60mg OD <1 week, Atorvastatin, Salbutamol and Steroid Inhaler.
GI 4158	S22(Epi), S23(CD45/Str) , S21(PBMC)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_4	CC (Dual Therapy)	Inflamed	Sigmoid/Desc ending	Melanoma, Epilepsy	UCEIS 6/8	M	68	59	23	Methylprednis olone, Prednisolone (>2 weeks), Mycophenolat e Mofetil 1g BD, Omeprazole, Paracetamol, Codeine, Phenytoin, Levetiracetam,
GI 4680	S51(Epi), S52(CD45/Str) , S44(PBMC)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_3	CC (Dual Therapy)	Not Inflamed	Sigmoid/Desc ending	Melanoma, Hypertension		M	68	113	0	Infliximab, Prednisolone (10mg OD), Amlodipine, B12 injections 3 monthly
PRISE 1	S3(CD45/Str), S1(PBMC)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_5	CC (Dual Therapy)	Not Inflamed	Sigmoid/Desc ending	Melanoma, Checkpoint inhibitor induced Myositis and Thyroiditis		F	55	39	0	Prednisolone 60mg, Paracetamol, Codeine, Levothyroxine, Omeprazole
GI 4379	S58(CD45/Str)	Epi+CD45/Str	5' 10X scRNA	Yes	HTO_3	UC	Inflamed	Sigmoid	Cholangitis (Antibody negative), Meniere's disease	UCEIS 3/8	F	66		Unclear	Ursofalk

PRISE 11	S22(Epi), S23(CD45/Str), S21(PBMC)	Epi+CD45/Str+PBMC	5' 10X scRNA	Yes	HTO_5	CC (Dual Therapy)	Not Inflamed	Sigmoid/Descending	Melanoma, Rheumatoid Arthritis, OA Knees, Previous Basal Carcinoma of the Skin		F	74	15	0	Prednisolone (20mg), Statin, Omeprazole
PRISE 12	S13(CD45/Str)	Epi+CD45/Str+PBMC	5' 10X scRNA	Yes	HTO_5	CC (Dual Therapy)	Not Inflamed	Sigmoid/Descending	Melanoma, Transient Ischaemic Attack, Benign Prostatic Hypertrophy, Hypertension		M	60	14	0	Citalopram, Aspirin, Amlodipine, Atorvastatin, Tamsulosin.
GI 2935	S51(Epi), S52(CD45/Str), S44(PBMC)	Epi+CD45/Str+PBMC	5' 10X scRNA	Yes	HTO_2	CC (Monotherapy)	Inflamed	Rectum	Lung (Squamous), Prev Colon cancer (Treated with Sigmoid colectomy)	UCEIS 1/8	M	73	17	10	Amitryptiline, Aspirin, Clenil Modulite, Finasteride, Simvastatin, Tamsulosin, Inhalers
GI 3052	S37(Epi), S38(CD45/Str), S43(PBMC)	Epi+CD45/Str+PBMC	5' 10X scRNA	Yes	HTO_1	CC (Monotherapy)	Inflamed	Sigmoid/Rectum	Melanoma	UCEIS 1/8	F	44	8	15	Nil
GI 4738	S51(Epi), S52(CD45/Str), S44(PBMC)	Epi+CD45/Str+PBMC	5' 10X scRNA	Yes	HTO_1	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Lung Ca (Adenocarcinoma)	UCEIS 4/8	F	67	21	27	Loperamide only
PRISE 5	S38(CD45/Str), S43(PBMC)	Epi+CD45/Str+PBMC	5' 10X scRNA	Yes	HTO_2	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Renal Cancer, Checkpoint Thyroiditis, Hypothyroidism, Cervical Laminoplasty	UCEIS 4/8	M	49	28	<52 days	Amitryptyline, Gabapentin, Levothyroxine, Morphine, Venlafaxine,
GI 3128	S51(Epi), S52(CD45/Str), S44(PBMC)	Epi+CD45/Str+PBMC	5' 10X scRNA	Yes	HTO_4	CC (Monotherapy)	Not Inflamed	Sigmoid/Descending	Melanoma		M	72	19	0	Mometasone ointment
GI 4448	S37(Epi), S38(CD45/Str), S43(PBMC)	Epi+CD45/Str+PBMC	5' 10X scRNA	Yes	HTO_5	CC (Monotherapy)	Not Inflamed	Sigmoid/Descending	Melanoma, Rheumatoid Arthritis		F	67	11	0	Prednisolone (60mg), Abatacept, Dosulepin, Inhalers
PRISE 13	S51(Epi), S52(CD45/Str), S44(PBMC)	Epi+CD45/Str+PBMC	5' 10X scRNA	Yes	HTO_5	CC (Monotherapy)	Not inflamed	Sigmoid/Descending	Melanoma, Type 2 Diabetes (Type 2 DM), Osteoarthitis		M	76	11	0	Atorvastatin, co-codamol, gliclazide, metformin, mariosea XL, tamsulosin, quinine, omeprazole
PRISE 4	S37(Epi), S38(CD45/Str), S43(PBMC)	Epi+CD45/Str+PBMC	5' 10X scRNA	Yes	HTO_4	CC (Monotherapy)	Not inflamed	Sigmoid/Descending	Melanoma, Type 2 DM, HTN, AF		M	74	7	0	Rivaroxaban, Lisinopril, Omeprazole, Bisoprolo, Bendroflumethiazide

TIP 200	S22(Epi), S23(CD45/Str), S21(PBMC)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_1	Healthy	Not inflamed	Pan-colonic	Rectal bleed		F	30		0	Nil
TIP 246	S12(Epi), S13(CD45/Str)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_1	Healthy	Not Inflamed	Pan-colonic	Polyp surveillance, Vitamin D deficiency		F	63		0	Vit D, Omeprazole, lymecycline
TIP 278	S2(Epi), S3(CD45/Str), S1(PBMC)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_1	Healthy	Not inflamed	Descending (4pairs)	Rectal bleeding, Diarrhoea		F	70		0	Risedronate, atorvastatin, cholecalciferol, rabeprazole
GI 4381	S22(Epi), S23(CD45/Str), S21(PBMC)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_2	UC	Inflamed	Sigmoid (4 pairs)	UC, Polyp surveillance	UCEIS 4/8	F	74		54	Azathioprine, prednisolone supp, calci-D, laxido, prednisolone
IBD 3350	S12(Epi), S13(CD45/Str)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_2	UC	Inflamed	Caecum (4 pairs)	UC, Asthma	UCEIS 4/8	M	28		146	Olsalazine. Pentasa, Ventolin, Pelia
IBD 3947	S2(Epi), S3(CD45/Str), S1(PBMC)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_2	UC	Inflamed	Descending (4 pairs)	Assessment of IBD (asthma, hypertension)	UCEIS 3/8	M	68		Unclear	Losartan, folic acid, pentasa, spiriva, allopurinol, statin
GI 4381	S22(Epi), S23(CD45/Str)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_3	UC	Not Inflamed	Descending (4 pairs)	UC, Polyp surveillance		F	74		54	Azathioprine, prednisolone supp, calci-D, laxido, prednisolone
IBD 3350	S12(Epi), S13(CD45/Str)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_3	UC	Not Inflamed	Transverse (4 pairs)	UC, Asthma		M	28		146	Olsalazine. Pentasa, Ventolin, Pelia
IBD 3947	S2(Epi), S3(CD45/Str)	Epi+CD45/Str +PBMC	5' 10X scRNA	Yes	HTO_3	UC	Not Inflamed	Pan-colonic	Asthma, HTN		M	68		Unclear	Losartan, folic acid, pentasa, spiriva, allopurinol, statin
4055(23)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						38	
4055(26)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						76	
PRISE14(34)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						10	
3692 (6)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						3	
4449(28)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						151	
4310 (10)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						82	
4310 (30)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						10	
4059 (15)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						29	
4054(24)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						12	
4054(38)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						107	

4051(3)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						216	
4051(14)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						58	
3700(20)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						308	
3700(5)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						6	
4449(5)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						26	
PRISE14(1)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						400	
4059 (28)	NA	H&E	N/A	N/A	N/A	CC	Inflamed	Sigmoid/Descending						93	
3593	NA	H&E	N/A	N/A	N/A	Healthy	Not inflamed	Sigmoid/Descending						0	
3599	NA	H&E	N/A	N/A	N/A	Healthy	Not inflamed	Sigmoid/Descending						0	
TIP 289	NA	H&E	N/A	N/A	N/A	Healthy	Not inflamed	Sigmoid/Descending						0	
GICor2	NA	H&E	N/A	N/A	N/A	Healthy	Not inflamed	Sigmoid/Descending						0	
GICor3	NA	H&E	N/A	N/A	N/A	Healthy	Not inflamed	Sigmoid/Descending						0	
GICor10	NA	H&E	N/A	N/A	N/A	Healthy	Not inflamed	Sigmoid/Descending						0	
GICor12	NA	H&E	N/A	N/A	N/A	Healthy	Not inflamed	Sigmoid/Descending						0	
2861	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending						46	
3349	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending						36	
2174	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending						798	
1497	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending						824	
3643	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending						50	
TC65	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending						79	
46570/16	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending							
3328	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending						325	
28473/16	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending							
4652	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending						35	
TYNE14	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending							
3592	NA	H&E	N/A	N/A	N/A	UC	Inflamed	Sigmoid/Descending						42	
GI 2836	N/A	Immunofluorescence	CD103 Ecad DAPI	N/A	N/A	Healthy	Not inflamed	Sigmoid/Descending	Colonic Polyps		F	62			Omeprazole, Oxybutynin, Rantidine
DNDT 133	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending	Colonic Polyps		F	54			Nil

TIP 275	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending	Type 2 Diabetes, Asthma, Hypertension, Diverticulosis		F	60			Atorvastatin, Fluoxetine, Fostair, Pregabalin, Levemir insulin, Lisinopril, Loratadine, Metformin
TIP 289	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending	Previous Colonic Polyp		M	33			Nil
IBD 1497	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	UC	Inflamed	Sigmoid/Descending	UC, Barretts' Oesophagus	UCEIS 3/8	M	60			Mesalazine, Omeprazole
GI 2353	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	UC	Inflamed	Sigmoid/Descending	UC	UCEIS 5/8	M	65			Asacol
IBD 2825	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	UC	Inflamed	Sigmoid/Descending	UC	UCEIS 5/8	F	59			Cortiment
GI 3349	N/A	Immunofluorescence	CD103 Ecad DAPI; CD163 FOXP3 DAPI	NA	NA	UC	Inflamed	Sigmoid/Descending	UC	UCEIS 5/8	F	70			Nil
DNDT 110	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	UC	Inflamed	Rectum	UC Proctitis		F	39			Nil
TC 65	N/A	Immunofluorescence	CD103 Ecad DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	UC	Inflamed	Rectum	UC, Endometriosis, Enteric Arthropathy	UCEIS 1/8	F	43			Sulfasalazine, Mesalazine, Gabapentin, Contraceptive Pill
TC 62	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	UC	Inflamed	Rectum	UC Proctitis	UCEIS 6/8	F	43			Pentasa, Predsol enema
GI 3692	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Lung Cancer	UCEIS 5/8	M	68			Nil
GI 4051	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 4/8	F	63			Co-trimoxazole, Prednisolone, Levothyroxine, Omeprazole
GI 4059	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 5/8	M	70			Methylprednisolone, Warfarin
GI 3700	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 1/8	M	48			Nil
GI 3691	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 1/8	F	68			Prednisolone, Sando-K
PRISE 14	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 6/8	M	71			Prednisolone, Methylprednisolone
PRISE 2	N/A	Immunofluorescence	CD103 Ecad DAPI	NA	NA	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 1/8	M	74			Nil
GI 3691 (17)	N/A	Immunofluorescence	CD163 FABP1 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 1/8	F	68		5	Prednisolone, Sando-K
GI 4310(10)	N/A	Immunofluorescence	CD163 FABP1 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 0/8				82	

GI 4449(5)	N/A	Immunofluorescence	CD163 FABP1 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Lung Ca (Adenocarcinoma)	UCEIS 5/8				26	Prednisolone, Infliximab
GI 4449(28)	N/A	Immunofluorescence	CD163 FABP1 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Lung Ca (Adenocarcinoma)	UCEIS 1/8				151	Prednisolone, Vedolizumab, Tofacitinib
GI PRISE14(34)	N/A	Immunofluorescence	CD163 FABP1 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 6/8	M	72		10	Prednisolone 30mg OD, Omeprazole
GI PRISE14(1)	N/A	Immunofluorescence	CD163 FABP1 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	Severely inflamed	M	74		400	Steroids and Infliximab
GI 4310(30)	N/A	Immunofluorescence	CD163 FABP1 DAPI, Ecad CleavedCaspase3 DAPI	NA	NA	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 0/8				13	Prednisolone
GI 4055(23)	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 3/8				38	
GI 4055(26)	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 3/8				76	
GI 4054(38)	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Rectum	Melanoma	Mild Inflammation				107	
GI 4051(14)	N/A	Immunofluorescence	CD163 FABP1 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 4/8				58	Prednisolone, Infliximab.
GI 3700(5)	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 1/8	M	48		6	Nil
GI Cor10	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending							
GI 2863	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending							
GI 3560	N/A	Immunofluorescence	CD163 FABP1 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending							
GI COR12	N/A	Immunofluorescence	CD163 FABP1 DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending							

GI COR3	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending							
GI COR2	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending							
GI COR12	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending							
BB4652	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	UC	Inflamed	Sigmoid	Ulcerative Colitis, Left Ventricular Failure, Implantable Cardiac Defib, Atrial Fibrillation (AF)	UCEIS 5/8	M	65		35	Vedolizumab, Apixaban, Atorvastatin, Ramipril, Bisoprolol
IBD 3592	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	UC	Inflamed	Sigmoid/Descending	Ulcerative Colitis	UCEIS 5/8				36	
BB 2837	N/A	Immunofluorescence	CD163 FABP1 DAPI	NA	NA	UC	Inflamed	Rectum	Ulcerative Colitis					26	
IBD 2174	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	UC	Inflamed	Sigmoid/Descending	Ulcerative Colitis	UCEIS 5/8				798	
TIP TYNE 14	N/A	Immunofluorescence	CD163 FABP1 DAPI; CD163 FOXP3 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	UC	Inflamed	Sigmoid/Descending	Ulcerative Colitis						
GI 4051(3)	N/A	Immunofluorescence	CD163 FOXP3 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	Very mild				216	Prednisolone, Vedolizumab, FMT.
GI 3700(20)	N/A	Immunofluorescence	CD163 FOXP3 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 0/8	M	49		308	
GI 4054(24)	N/A	Immunofluorescence	CD163 FOXP3 DAPI; Ecad CleavedCaspase3 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 5/8				12	Prednisolone
GI 4059(15)	N/A	Immunofluorescence	CD163 FOXP3 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 5/8				29	Prednisolone, Infliximab
GI 4059(28)	N/A	Immunofluorescence	CD163 FOXP3 DAPI	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma	UCEIS 5/8				93	Prednisolone, Infliximab
PRISE7 (7)	N/A	Immunofluorescence	Ecad CleavedCaspase3 DAPI	NA	NA	CC	Inflamed	Sigmoid/Descending						11	

PRISE14 (27)	N/A	Immunofluorescence	Ecad CleavedCaspase3 DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending						0	
PRISE15 (25)	N/A	Immunofluorescence	Ecad CleavedCaspase3 DAPI	NA	NA	Healthy	Not inflamed	Sigmoid/Descending						0	
GI2861	N/A	Immunofluorescence	Ecad CleavedCaspase3 DAPI	NA	NA	UC	Inflamed	Sigmoid/Descending						46	
GI2837	N/A	Immunofluorescence	Ecad CleavedCaspase3 DAPI	NA	NA	UC	Inflamed	Rectum						26	
GI 2935	N/A	FACS	Bound cell analysis	NA	NA	CC (Monotherapy)	Inflamed	Rectum	Lung (Squamous), Prev Colon cancer (Treated with Sigmoid colectomy)	UCEIS 1/8	M	73	17	10	Amitryptiline, Aspirin, Clenil Modulite, Finasteride, Simvastatin, Tamsulosin, Inhalers
GI 3031	N/A	FACS	Bound cell analysis	NA	NA	CC (Monotherapy)	Inflamed	Sigmoid/Descending	Melanoma	Mild Inflammation	M	72	11		Prednisolone 60mg, Finasteride, Tamsulosin, Morphine, Paracetamol.
GI 4680	N/A	FACS	Bound cell analysis	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid/Descending	Melanoma, Hypertension	UCEIS 5/8	M	68	42		Infliximab, Prednisolone (10mg OD), Amlodipine, B12 injections 3 monthly
GI 3032	N/A	FACS	Bound cell analysis	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid	Melanoma	UCEIS 2/8	M	70	64		Infliximab, Prednisolone
GI 4859	N/A	FACS	Bound cell analysis	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid	Melanoma	UCEIS 2/8	F	37	18		Prednisolone, Omeprazole, Naproxen
GI 3061	N/A	FACS	Bound cell analysis,not subset analysis	NA	NA	CC (Dual Therapy)	Inflamed	Sigmoid	Melanoma	UCEIS 1/8			13		None
PRISE 23	N/A	FACS	Duration analysis only	NA	NA	CC (Monotherapy)	Inflamed	Sigmoid	Diarrhoea. Bladder Cancer	UCEIS 2/8	M	74	40	21	Lansoprazole, Naproxen, Mometasone, Morphine
GI Cor21	N/A	FACS	Duration analysis only	NA	NA	CC (Monotherapy)	Not inflamed	Sigmoid	Melanoma				456		
TIP 301	N/A	FACS	Nivolumab binding in vitro	NA	NA	Healthy	Not inflamed	Sigmoid							
TIP 381	N/A	FACS	Nivolumab binding in vitro	NA	NA	Healthy	Not inflamed	Sigmoid							
GI 6878	N/A	FACS	Nivolumab binding in vitro	NA	NA	Healthy	Not inflamed	Sigmoid							
TIP 599	N/A	FACS	Nivolumab binding in vitro	NA	NA	Healthy	Not inflamed	Sigmoid							
TIP 617	N/A	FACS	Nivolumab binding in vitro	NA	NA	Healthy	Not inflamed	Sigmoid							
GI 7936	N/A	FACS	Nivolumab binding in vitro	NA	NA	Healthy	Not inflamed	Sigmoid							

S1	CITE-seq 24/10/19 Hashing #2	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_2	UC	Inflamed	Sigmoid/Descending		UCEIS 3/8	F	62			Calcium-Vitamin D3, Mesalazine, Octasa
S2	CITE-seq 24/10/19 Hashing #3	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_3	Healthy	Not inflamed	Sigmoid/Descending	Osteoporosis		F	64			Risendronate 35mg
S3	CITE-seq 24/10/19 Hashing #4	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_4	UC	Inflamed	Sigmoid/Descending		Severe	M	33			Octasa, Beclometasone
S4	CITE-seq 24/10/19 Hashing #1	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_1	UC	Inflamed	Sigmoid/Descending		Severe	M	56			Methotrexate, Infliximab
S5	CITE-seq 24/10/19 Hashing #5	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_5	Healthy	Not inflamed	Sigmoid/Descending	Hypertension		M	88			Apixaban, Co-codamol, Bendroflumethiazide, Beclometasone, Amlodipine
S6	CITE-seq 23/10/19 Hashing #3	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_3	UC	Inflamed	Pan-colonic	Asthma	UCEIS 4/8	F	24			Salbutamol, Clenil Modulate
S7	CITE-seq 23/10/19 Hashing #5	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_5	Healthy	Not inflamed	Sigmoid/Descending	Ischaemic heart disease, Heart valve replacement, Hypertension		M	72			Aspirin, Atorvastatin, Clopidogrel, Ramipril, Tamsulosin, Alendronic acid
S8	CITE-seq 23/10/19 Hashing #1	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_1	UC	Inflamed	Sigmoid/Descending		UCEIS 4/8	F	30			Adalimumab, Asacol
S9	CITE-seq 23/10/19 Hashing #2	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_2	Healthy	Not inflamed	Sigmoid/Descending	Osteoarthritis		M	50			
S10	CITE-seq 23/10/19 Hashing #4	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_4	UC	Inflamed	Sigmoid/Descending	Asthma, Bronchiectasis	UCEIS 5/8	M	38			Salbutamol
S11	CITE-seq 2/10/19 Hashing #5	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_5	Healthy	Not inflamed		None		M	45			
S12	CITE-seq, 2/10/19 Hashing #3	Biopsy CD3+CD8+	5' 10X scRNA	Yes	HTO_3	UC	Inflamed	Sigmoid/Descending	Asthma, Parkinsons, Osteopenia	UCEIS 3/8	F	65			Mesalazine, Salazopyrin, Rasagiline,Madopar
S21	scRNA-seq	Biopsy CD3+CD8+	5' 10X scRNA	No	N/A	Healthy	Not inflamed	Pan-colonic	Hypertension		F	54			Ramipril

S22	scRNA-seq	Biopsy CD3+CD8+	5' 10X scRNA	No	N/A	Healthy	Not inflamed	Sigmoid/Descending	Hypopituitarism (Pituitary Tumour Treated 1987), High Cholesterol, Hypertension.		M	82			Hydrocortisone, Omeprazole, Levothyroxine, Testosterone sachet, Somatropin, Aspirin, Levothyroxine, Quinine, Amlodipine, Atorvastatin
S23	scRNA-seq	Biopsy CD3+CD8+	5' 10X scRNA	No	N/A	Healthy	Not inflamed	Pan-colonic	Irritable Bowel Syndrome, Asthma		F	38			Seretide, Sertraline
S24	scRNA-seq	Biopsy CD3+CD8+	5' 10X scRNA	No	N/A	UC	Inflamed	Sigmoid/Descending		UCEIS 5/8	F	67			Salofalk Suppositories
S33	scRNA-seq	Biopsy CD3+CD8+	5' 10X scRNA	No	N/A	UC	Inflamed	Sigmoid/Descending		UCEIS 4/8	F	29			Pentas,a Sertraline
S34	scRNA-seq	Biopsy CD3+CD8+	5' 10X scRNA	No	N/A	UC	Inflamed	Sigmoid/Descending	Asthma	UCEIS 5/8	F	40			Mercaptopurine, Olsalazine, Symbicort
S_3480		qPCR	*IL26* detection from whole tissue	N/A	N/A	Healthy	Not inflamed	Left sided	Parkinsons, Hypothyroidism	N/A	M	80			Sinemet, Mirtazapine, Levothyroxine
S_2327		qPCR	*IL26* detection from whole tissue	N/A	N/A	Healthy	Not inflamed	Left sided	Nil	N/A	F	24			
S_2957		qPCR	*IL26* detection from whole tissue	N/A	N/A	Healthy	Not inflamed	Left sided	Corneal Scar	N/A	M	52			
S_3264		qPCR	*IL26* detection from whole tissue	N/A	N/A	Healthy	Not inflamed	Left sided	Asthma, Eczema	N/A	M	56			Ventolin, Steroid ointment
S_3579		qPCR	*IL26* detection from whole tissue	N/A	N/A	Healthy	Not inflamed	Left sided	Hypertension	N/A	F	68			Amlodipine, Risperidone
P_26		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 8/8	F	29			Asacol
P_26		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	F	29			Asacol
P_43		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 2/8	M	56			Octasa, Mercaptopurine
P_43		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	M	56			Octasa, Mercaptopurine
P_45		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 1/8	F	17			Vedolizumab, Mesavant
P_45		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	F	17			Vedolizumab, Mesavant

P_49		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 1/8	M	52			Azathioprine, Mesalazine
P_49		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	M	52			Azathioprine, Mesalazine
P_52		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 4/8	M	41			Asacol
P_52		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	M	41			Asacol
P_56		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 4/8	F	35			Asacol, Mycophenolate
P_56		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	F	35			Asacol, Mycophenolate
P_59		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 4/8	F	48			Vedolizumab, Mesalazine, Pred
P_59		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	F	48			Vedolizumab, Mesalazine, Pred
P_61		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 5/8	F	41			Pentasa
P_61		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	F	41			Pentasa
P_62		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 2/8	F	38			Pentasa
P_62		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	F	38			Pentasa
P_63 Timepoint 1		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 5/8	M	54			Pentasa
P_63 Timepoint 1		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	M	54			Pentasa
P_63 Timepoint 2		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 6/8	M	55			Vedolizumab
P_63 Timepoint 2		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	M	55			Vedolizumab
P_68		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 3/8	F	38			Mesalazine
P_68		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	F	38			Mesalazine
P_71		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 1/8	F	48			Vedolizumab

P_71		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	F	48			Vedolizumab
P_79 Timepoint 1		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 1/8	M	19			Infliximab, Methotrexate
P_79 Timepoint 1		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	M	19			Infliximab, Methotrexate
P_79 Timepoint 2		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 1/8	M	19			Infliximab, Methotrexate
P_79 Timepoint 2		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	M	19			Infliximab, Methotrexate
P_80		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 2/8	M	50			Azathioprine
P_80		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	M	50			Azathioprine
P_85		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 4/8	F	34			Pentasa, Mesalazine
P_85		qPCR	*IL26* detection from whole tissue	N/A	N/A	UC NI	Not inflamed	Left sided		UCEIS 0/8	F	34			Pentasa, Mesalazine
S_1901		ELISA	IL26 detection	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 2/8	F	59			Vedolizumab
H_11		ELISA	IL26 detection	N/A	N/A	UC INF	Inflamed	Left sided		Mild-Moderate	M	41			Azathioprine
H_16		ELISA	IL26 detection	N/A	N/A	UC INF	Inflamed	Left sided	Cholecystectomy	Mild	M	36			Mesalazine
H_20		ELISA	IL26 detection	N/A	N/A	UC INF	Inflamed	Left sided		Mild	F	57			Salofalk
H_21		ELISA	IL26 detection	N/A	N/A	UC INF	Inflamed	Left sided		Mild	F	30			Pentasa, Mesalazine
S_3741		ELISA	IL26 detection	N/A	N/A	UC INF	Inflamed	Left sided		UCEIS 2/8	M	30			Infliximab, Mesalazine
S_132		ELISA	IL26 detection	N/A	N/A	Healthy	Not inflamed	Left sided	Hypertension, Atrial Flutter, Prostatism	N/A	M	59			Bisoprolol, Finasteride, Losartan, Simvastatin, Buscopan
S_105		ELISA	IL26 detection	N/A	N/A	Healthy	Not inflamed	Left sided	Multiple Sclerosis	N/A	F	46			
S_067		ELISA	IL26 detection	N/A	N/A	Healthy	Not inflamed	Left sided		N/A	M	70			Lansoprazole
S_169		ELISA	IL26 detection	N/A	N/A	Healthy	Not inflamed	Left sided	Hypertension, COPD	N/A	/				Lisinopril, Simvastatin, PRN Salbutamol
H_3829		ELISA	IL26 detection	N/A	N/A	Healthy	Not inflamed	Left sided	Diabetes	N/A	M	72			Metformin
H_956		ELISA	IL26 detection	N/A	N/A	Healthy	Not inflamed	Left sided		N/A	/				
TIP_279		FACS	IL26 receptor detection	N/A	N/A	Healthy	Not inflamed	Left sided	Migraines	N/A	F	27			Folic Acid, Sumatriptan
TIP_106		FACS	IL26 receptor detection	N/A	N/A	Healthy	Not inflamed	Left sided	Dyspepsia, Depression	N/A	F	60			Lansoprazole, Fluoxetine
TIP_23		FACS	IL26 receptor detection	N/A	N/A	Healthy	Not inflamed	Left sided	Hypertension Hypothyroidism	N/A	M	81			Lisinopril, Thyroxine, Amlodpine.

EAST_13		FACS	IL26 receptor detection	N/A	N/A	UC INF	Inflamed	Left sided		Mild-Moderate	F	34			Salofalk
EAST_21		FACS	IL26 receptor detection	N/A	N/A	UC INF	Inflamed	Left sided	Dyspepsia	Mild	F	31			Prednisolone, Infliximab

References

1. Cheung, V. T. F. *et al.* Immune checkpoint inhibitor-related colitis assessment and prognosis: can IBD scoring point the way? *British Journal of Cancer 2020 123:2* **123**, 207–215 (2020).

2. Corridoni, D. *et al.* Single-cell atlas of colonic CD8 + T cells in ulcerative colitis. *Nat Med* **26**, 1480–1490 (2020).

3. Lee, K. A. *et al.* The gut microbiome: what the oncologist ought to know. *Br J Cancer* **125**, 1197–1209 (2021).

4. Trøseid, M., Andersen, G. Ø., Broch, K. & Hov, J. R. The gut microbiome in coronary artery disease and heart failure: Current knowledge and future directions. *EBioMedicine* **52**, (2020).

5. Witkowski, M., Weeks, T. L. & Hazen, S. L. Gut Microbiota and Cardiovascular Disease. *Circ Res* **127**, 553–570 (2020).

6. Fetissov, S. O. Role of the gut microbiota in host appetite control: bacterial growth to animal feeding behaviour. *Nat Rev Endocrinol* **13**, 11–25 (2017).

7. Ananthakrishnan, A. N. *et al.* Environmental triggers in IBD: a review of progress and evidence. *Nature Reviews Gastroenterology & Hepatology 2017 15:1* **15**, 39–49 (2017).

8. Ni, J., Wu, G. D., Albenberg, L. & Tomov, V. T. Gut microbiota and IBD: causation or correlation? *Nature Reviews Gastroenterology & Hepatology 2017 14:10* **14**, 573–584 (2017).

9. Hindryckx, P., Jairath, V. & D'Haens, G. Acute severe ulcerative colitis: from pathophysiology to clinical management. *Nature Reviews Gastroenterology & Hepatology 2016 13:11* **13**, 654–664 (2016).

10. Roda, G., Jharap, B., Neeraj, N. & Colombel, J. F. Loss of Response to Anti-TNFs: Definition, Epidemiology, and Management. *Clin Transl Gastroenterol* **7**, e135 (2016).

11. Glick, L. R., Cifu, A. S. & Feld, L. Ulcerative Colitis in Adults. *JAMA* **324**, 1205–1206 (2020).

12. Lamb, C. A. *et al.* British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. *Gut* **68**, s1–s106 (2019).

13. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).

14. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature 2009 459:7244* **459**, 262–265 (2009).

15. Sato, T. *et al.* Long-term Expansion of Epithelial Organoids From Human Colon, Adenoma, Adenocarcinoma, and Barrett's Epithelium. (2011) doi:10.1053/j.gastro.2011.07.050.

16. Rubin, S. J. S. *et al.* Mass cytometry reveals systemic and local immune signatures that distinguish inflammatory bowel diseases. *Nature Communications 2019 10:1* **10**, 1–14 (2019).

17. Kondo, A. *et al.* Highly Multiplexed Image Analysis of Intestinal Tissue Sections in Patients With Inflammatory Bowel Disease. *Gastroenterology* **0**, (2021).

18. Rao, A., Barkley, D., França, G. S. & Yanai, I. Exploring tissue architecture using spatial transcriptomics. *Nature* **596**, 211–220 (2021).

19. Parikh, K. *et al.* Colonic epithelial cell diversity in health and inflammatory bowel disease. *Nature 2019 567:7746* **567**, 49–55 (2019).

20. Smillie, C. S. *et al.* Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714-730.e22 (2019).

21. Koo, B.-K. & Clevers, H. Stem Cells Marked by the R-Spondin Receptor LGR5. *Reviews in basic and clinical gastroenterology and hepatology* (2014) doi:10.1053/j.gastro.2014.05.007.

22. Zhu, G., Hu, J. & Xi, R. The cellular niche for intestinal stem cells: a team effort. *Cell Regeneration 2021 10:1* **10**, 1–16 (2021).

23. Biton, M. *et al.* T Helper Cell Cytokines Modulate Intestinal Stem Cell Renewal and Differentiation. *Cell* **175**, 1307-1320.e22 (2018).

24. Fawkner-Corbett, D. *et al.* Spatiotemporal analysis of human intestinal development at single-cell resolution. *Cell* **184**, 810-826.e23 (2021).

25. Elmentaite, R. *et al.* Cells of the human intestinal tract mapped across space and time. *Nature 2021 597:7875* **597**, 250–255 (2021).

26. Wang, K. *et al.* Butyrate induces development-dependent necrotizing enterocolitis-like intestinal epithelial injury via necroptosis. *Pediatric Research 2022* 1–9 (2022) doi:10.1038/s41390-022-02333-z.

27. Grouls, M. *et al.* Differential gene expression in iPSC-derived human intestinal epithelial cell layers following exposure to two concentrations of butyrate, propionate and acetate. *Scientific Reports 2022 12:1* **12**, 1–15 (2022).

28. Venegas, D. P. *et al.* Short chain fatty acids (SCFAs)mediated gut epithelial and immune regulation and its relevance for inflammatory bowel diseases. *Front Immunol* **10**, 277 (2019).

29. Kinchen, J. *et al.* Structural Remodeling of the Human Colonic Mesenchyme in Inflammatory Bowel Disease. *Cell* **175**, 372-386.e17 (2018).

30. Yan, K. S. *et al.* Non-equivalence of Wnt and R-spondin ligands during Lgr5+ intestinal stem-cell self-renewal. *Nature 2017 545:7653* **545**, 238–242 (2017).

31. Castro-Dopico, T. *et al.* Anti-commensal IgG Drives Intestinal Inflammation and Type 17 Immunity in Ulcerative Colitis. *Immunity* **50**, 1099-1114.e10 (2019).

32. Yin, Z., Zhou, Y., Turnquist, H. R. & Liu, Q. Neuro–epithelial–ILC2 crosstalk in barrier tissues. *Trends Immunol* **43**, 901–916 (2022).

33. Martin, J. C. *et al.* Single-Cell Analysis of Crohn's Disease Lesions Identifies a Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy. *Cell* **178**, 1493-1508.e20 (2019).

34. Cummings, R. J. *et al.* Different tissue phagocytes sample apoptotic cells to direct distinct homeostasis programs. *Nature 2016 539:7630* **539**, 565–569 (2016).

35. Parikh, K. *et al.* Colonic epithelial cell diversity in health and inflammatory bowel disease. *Nature* **567**, 49–55 (2019).

36. Singh, V. *et al.* Chronic Inflammation in Ulcerative Colitis Causes Long-Term Changes in Goblet Cell Function. *Cell Mol Gastroenterol Hepatol* **13**, 219–232 (2022).

37. Dotti, I. *et al.* Alterations in the epithelial stem cell compartment could contribute to permanent changes in the mucosa of patients with ulcerative colitis. *Gut* **66**, 2069–2079 (2017).

38. Sarvestani, S. K. *et al.* Induced organoids derived from patients with ulcerative colitis recapitulate colitic reactivity. *Nature Communications 2021 12:1* **12**, 1–18 (2021).

39. Actis, G. C., Pellicano, R., Fagoonee, S. & Ribaldone, D. G. History of Inflammatory Bowel Diseases. *J Clin Med* **8**, (2019).

40. Poggi, A. *et al.* Human gut-associated natural killer cells in health and disease. *Frontiers in Immunology* vol. 10 Preprint at https://doi.org/10.3389/fimmu.2019.00961 (2019).

41. Smids, C. *et al.* Intestinal T Cell Profiling in Inflammatory Bowel Disease: Linking T Cell Subsets to Disease Activity and Disease Course. *J Crohns Colitis* **12**, 465–475 (2018).

42. Roosenboom, B. *et al.* Intestinal CD103+CD4+ and CD103+CD8+ T-Cell Subsets in the Gut of Inflammatory Bowel Disease Patients at Diagnosis and During Follow-up. *Inflamm Bowel Dis* **25**, 1497–1509 (2019).

43. Yamada, A. *et al.* Role of regulatory T cell in the pathogenesis of inflammatory bowel disease. *World J Gastroenterol* **22**, 2195 (2016).

44. Voskens, C. *et al.* Autologous regulatory T-cell transfer in refractory ulcerative colitis with concomitant primary sclerosing cholangitis. *Gut* **72**, 49–53 (2023).

45. Noviello, D. *et al.* The IL23-IL17 Immune Axis in the Treatment of Ulcerative Colitis: Successes, Defeats, and Ongoing Challenges. *Front Immunol* **12**, 1735 (2021).

46. Sands, B. E. *et al.* Ustekinumab as Induction and Maintenance Therapy for Ulcerative Colitis. *New England Journal of Medicine* **381**, 1201–1214 (2019).

47. Parkes, M. *et al.* PRedicting Outcomes For Crohn's dIsease using a moLecular biomarkEr (PROFILE): protocol for a multicentre, randomised, biomarker-stratified trial. *BMJ Open* **8**, e026767 (2018).

48. Fina, D. *et al.* Regulation of Gut Inflammation and Th17 Cell Response by Interleukin-21. *Gastroenterology* **134**, 1038-1048.e2 (2008).

49. de Nitto, D., Sarra, M., Pallone, F. & Monteleone, G. Interleukin-21 triggers effector cell responses in the gut. *World Journal of Gastroenterology : WJG* **16**, 3638 (2010).

50. Wang, Y. *et al.* IL-21/IL-21R signaling suppresses intestinal inflammation induced by DSS through regulation of Th responses in lamina propria in mice. *Scientific Reports 2016 6:1* **6**, 1–14 (2016).

51. Long, Y. *et al.* The Imbalance of Circulating Follicular Helper T Cells and Follicular Regulatory T Cells Is Associated With Disease Activity in Patients With Ulcerative Colitis. *Front Immunol* **11**, (2020).

52. Sun, L., Kong, R., Li, H. & Wang, D. The Role of T Follicular Helper Cells and Interleukin-21 in the Pathogenesis of Inflammatory Bowel Disease. *Gastroenterol Res Pract* **2021**, (2021).

53. Paap, E. M., Müller, T. M., Sommer, K., Neurath, M. F. & Zundler, S. Total Recall: Intestinal TRM Cells in Health and Disease. *Front Immunol* **11**, 3529 (2021).

54. Lamb, C. A. *et al.* αEβ7 Integrin Identifies Subsets of Pro-Inflammatory Colonic CD4+ T Lymphocytes in Ulcerative Colitis. *J Crohns Colitis* **11**, 610–620 (2017).

55. Peyrin-Biroulet, L. *et al.* Etrolizumab as induction and maintenance therapy for ulcerative colitis in patients previously treated with tumour necrosis factor inhibitors (HICKORY): a phase 3, randomised, controlled trial. *Lancet Gastroenterol Hepatol* **7**, 128–140 (2022).

56. Agrawal, M. & Verstockt, B. Etrolizumab for ulcerative colitis: beyond what meets the eye. *Lancet Gastroenterol Hepatol* **7**, 2–4 (2022).

57. Lee, J. C. *et al.* Gene expression profiling of CD8+ T cells predicts prognosis in patients with Crohn disease and ulcerative colitis. *J Clin Invest* **121**, 4170–4179 (2011).

58. West, N. R. *et al.* Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor–neutralizing therapy in patients with inflammatory bowel disease. *Nature Medicine 2017 23:5* **23**, 579–589 (2017).

59. Martin, J. C. *et al.* Single-Cell Analysis of Crohn’s Disease Lesions Identifies a Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy. *Cell* **178**, 1493-1508.e20 (2019).

60. Hou, K. *et al.* Microbiota in health and diseases. *Signal Transduction and Targeted Therapy 2022 7:1* **7**, 1–28 (2022).

61. Den Besten, G. *et al.* The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res* **54**, 2325 (2013).

62. Schmidt, T. S. B. *et al.* Drivers and determinants of strain dynamics following fecal microbiota transplantation. *Nature Medicine 2022 28:9* **28**, 1902–1912 (2022).

63. Ianiro, G. *et al.* Variability of strain engraftment and predictability of microbiome composition after fecal microbiota transplantation across different diseases. *Nature Medicine 2022 28:9* **28**, 1913–1923 (2022).

64. Imdad, A. *et al.* Fecal transplantation for treatment of inflammatory bowel disease. *Cochrane Database of Systematic Reviews* **2023**, (2023).

65. Tan, P., Li, X., Shen, J. & Feng, Q. Fecal Microbiota Transplantation for the Treatment of Inflammatory Bowel Disease: An Update. *Front Pharmacol* **11**, 574533 (2020).

66. Gupta, S., Allen-Vercoe, E. & Petrof, E. O. Fecal microbiota transplantation: in perspective. *Therap Adv Gastroenterol* **9**, 229 (2016).

67. Kedia, S. *et al.* Faecal microbiota transplantation with anti-inflammatory diet (FMT-AID) followed by anti-inflammatory diet alone is effective in inducing and maintaining remission over 1 year in mild to moderate ulcerative colitis: a randomised controlled trial. *Gut* **71**, 2401–2413 (2022).

68. Lavelle, A. & Sokol, H. Understanding and predicting the efficacy of FMT. *Nature Medicine 2022 28:9* **28**, 1759–1760 (2022).

69. Trapecar, M. *et al.* Gut-Liver Physiomimetics Reveal Paradoxical Modulation of IBD-Related Inflammation by Short-Chain Fatty Acids. *Cell Syst* **10**, 223-239.e9 (2020).

70. Salvi, P. S. & Cowles, R. A. Butyrate and the Intestinal Epithelium: Modulation of Proliferation and Inflammation in Homeostasis and Disease. *Cells* **10**, (2021).

71. Nanki, K. *et al.* Somatic inflammatory gene mutations in human ulcerative colitis epithelium. *Nature 2019 577:7789* **577**, 254–259 (2019).

72. Okazaki, T. & Honjo, T. PD-1 and PD-1 ligands: from discovery to clinical application. *Int Immunol* **19**, 813–824 (2007).

73. Seidel, J. A., Otsuka, A. & Kabashima, K. Anti-PD-1 and Anti-CTLA-4 Therapies in Cancer: Mechanisms of Action, Efficacy, and Limitations. *Front Oncol* **8**, 86 (2018).

74. de Sousa Linhares, A., Leitner, J., Grabmeier-Pfistershammer, K. & Steinberger, P. Not All Immune Checkpoints Are Created Equal. *Front Immunol* **9**, 1909 (2018).

75. Tokuhiro, S. *et al.* An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. *Nat Genet* **35**, 341–348 (2003).

76. Helms, C. *et al.* A putative RUNX1 binding site variant between SLC9A3R1 and NAT9 is associated with susceptibility to psoriasis. *Nat Genet* **35**, 349–356 (2003).

77. Kroner, A. *et al.* A PD-1 polymorphism is associated with disease progression in multiple sclerosis. *Ann Neurol* **58**, 50–57 (2005).

78. Maier, L. M. & Wicker, L. S. Genetic susceptibility to type 1 diabetes. *Curr Opin Immunol* **17**, 601–608 (2005).

79. Axelrod, M. L. *et al.* T cells specific for α-myosin drive immunotherapy-related myocarditis. *Nature 2022 611:7937* **611**, 818–826 (2022).

80. Zhang, Y. H., Tian, M., Tang, M. X., Liu, Z. Z. & Liao, A. H. Recent Insight into the Role of the PD-1/PD-L1 Pathway in Feto-Maternal Tolerance and Pregnancy. *Am J Reprod Immunol* **74**, 201–208 (2015).

81. McCarthy, E. F. The Toxins of William B. Coley and the Treatment of Bone and Soft-Tissue Sarcomas. *Iowa Orthop J* **26**, 154 (2006).

82. Wolchok, J. D. *et al.* Overall Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. *N Engl J Med* **377**, 1345–1356 (2017).

83. Larkin, J. *et al.* Five-Year Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. *New England Journal of Medicine* **381**, 1535–1546 (2019).

84. Garon, E. B. *et al.* Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med* **372**, 2018–2028 (2015).

85. Armand, P. *et al.* Nivolumab for Relapsed/Refractory Classic Hodgkin Lymphoma After Failure of Autologous Hematopoietic Cell Transplantation: Extended Follow-Up of the Multicohort Single-Arm Phase II CheckMate 205 Trial. *J Clin Oncol* **36**, 1428–1439 (2018).

86. Motzer, R. J. *et al.* Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *New England Journal of Medicine* **373**, 1803–1813 (2015).

87. Yau, T. *et al.* Nivolumab versus sorafenib in advanced hepatocellular carcinoma (CheckMate 459): a randomised, multicentre, open-label, phase 3 trial. *Lancet Oncol* **23**, 77–90 (2022).

88. Robert, C. A decade of immune-checkpoint inhibitors in cancer therapy. *Nature Communications 2020 11:1* **11**, 1–3 (2020).

89. Haslam, A. & Prasad, V. Estimation of the Percentage of US Patients With Cancer Who Are Eligible for and Respond to Checkpoint Inhibitor Immunotherapy Drugs. *JAMA Netw Open* **2**, e192535–e192535 (2019).

90. Probst, H. C., McCoy, K., Okazaki, T., Honjo, T. & van den Broek, M. Resting dendritic cells induce peripheral CD8+ T cell tolerance through PD-1 and CTLA-4. *Nat Immunol* **6**, 280–286 (2005).

91. Okazaki, T. & Honjo, T. Rejuvenating exhausted T cells during chronic viral infection. *Cell* **124**, 459–461 (2006).

92. Tawbi, H. A. *et al.* Relatlimab and Nivolumab versus Nivolumab in Untreated Advanced Melanoma. *New England Journal of Medicine* **386**, 24–34 (2022).

93. Haslam, A. & Prasad, V. Estimation of the Percentage of US Patients With Cancer Who Are Eligible for and Respond to Checkpoint Inhibitor Immunotherapy Drugs. *JAMA Netw Open* **2**, e192535 (2019).

94. Ji, M. *et al.* PD-1/PD-L1 pathway in non-small-cell lung cancer and its relation with EGFR mutation. *J Transl Med* **13**, 1–6 (2015).

95. Boland, P. M. & Ma, W. W. Immunotherapy for Colorectal Cancer. *Cancers (Basel)* **9**, (2017).

96. Sade-Feldman, M. *et al.* Defining T cell states associated with response to checkpoint immunotherapy in melanoma. *Cell* **175**, 998 (2018).

97. Thommen, D. S. The First Shall (Be) Last: Understanding Durable T Cell Responses in Immunotherapy. *Immunity* **50**, 6–8 (2019).

98. Kurtulus, S. *et al.* Checkpoint Blockade Immunotherapy Induces Dynamic Changes in PD-1−CD8+ Tumor-Infiltrating T Cells. *Immunity* **50**, 181-194.e6 (2019).

99. Thommen, D. S. *et al.* A transcriptionally and functionally distinct PD-1+ CD8+ T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. *Nat Med* **24**, 994–1004 (2018).

100. Sade-Feldman, M. *et al.* Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. *Cell* **175**, 998-1013.e20 (2018).

101. Blank, C. U. *et al.* Defining 'T cell exhaustion'. *Nat Rev Immunol* **19**, 665–674 (2019).

102. Keir, M. E., Butte, M. J., Freeman, G. J. & Sharpe, A. H. PD-1 and Its Ligands in Tolerance and Immunity. **26**, 677–704 (2008).

103. Wykes, M. N. & Lewin, S. R. Immune checkpoint blockade in infectious diseases. *Nature Reviews Immunology* vol. 18 91–104 Preprint at https://doi.org/10.1038/nri.2017.112 (2018).

104. Gambichler, T., Reuther, J., Scheel, C. H. & Becker, J. C. On the use of immune checkpoint inhibitors in patients with viral infections including COVID-19. *J Immunother Cancer* **8**, 1145 (2020).

105. Michot, J. M. *et al.* Immune-related adverse events with immune checkpoint blockade: a comprehensive review. *Eur J Cancer* **54**, 139–148 (2016).

106. Tan, M. H. *et al.* Spectrum of immune checkpoint inhibitors-induced endocrinopathies in cancer patients: a scoping review of case reports. *Clin Diabetes Endocrinol* **5**, (2019).

107. Schneider, S., Potthast, S., Komminoth, P., Schwegler, G. & Böhm, S. PD-1 Checkpoint Inhibitor Associated Autoimmune Encephalitis. *Case Rep Oncol* **10**, 473–478 (2017).

108. Min Lee, C. K. *et al.* Characterization of dermatitis after PD-1/PD-L1 inhibitor therapy and association with multiple oncologic outcomes: A retrospective case-control study. *J Am Acad Dermatol* **79**, 1047–1052 (2018).

109. Safa, H. *et al.* Immune checkpoint inhibitor related myasthenia gravis: single center experience and systematic review of the literature. *J Immunother Cancer* **7**, 319 (2019).

110. Tahir, S. A. *et al.* Autoimmune antibodies correlate with immune checkpoint therapy-induced toxicities. *Proc Natl Acad Sci U S A* **116**, 22246–22251 (2019).

111. Larkin, J. *et al.* Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *New England Journal of Medicine* **373**, 23–34 (2015).

112. Hodi, F. S. *et al.* Combined nivolumab and ipilimumab versus ipilimumab alone in patients with advanced melanoma: 2-year overall survival outcomes in a multicentre, randomised, controlled, phase 2 trial. *Lancet Oncol* **17**, 1558–1568 (2016).

113. Marthey, L. *et al.* Cancer Immunotherapy with Anti-CTLA-4 Monoclonal Antibodies Induces an Inflammatory Bowel Disease. *J Crohns Colitis* **10**, 395–401 (2016).

114. Bamias, G. *et al.* Immunological Characteristics of Colitis Associated with Anti-CTLA-4 Antibody Therapy. *Cancer Invest* **35**, 443–455 (2017).

115. Chen, J. H., Pezhouh, M. K., Lauwers, G. Y. & Masia, R. Histopathologic Features of Colitis Due to Immunotherapy With Anti-PD-1 Antibodies. *Am J Surg Pathol* **41**, 643–654 (2017).

116. Coutzac, C. *et al.* Colon Immune-Related Adverse Events: Anti-CTLA-4 and Anti-PD-1 Blockade Induce Distinct Immunopathological Entities. *J Crohns Colitis* 1238–1246 (2017) doi:10.1093/ecco-jcc/jjx081.

117. Shahabi, V. *et al.* Gene expression profiling of whole blood in ipilimumab-treated patients for identification of potential biomarkers of immune-related gastrointestinal adverse events. *J Transl Med* **11**, 75 (2013).

118. Menzies, A. M. *et al.* Anti-PD-1 therapy in patients with advanced melanoma and preexisting autoimmune disorders or major toxicity with ipilimumab. *Ann Oncol* **28**, 368–376 (2017).

119. Vetizou, M. *et al.* Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science (1979)* **350**, 1079–1084 (2015).

120. Routy, B. *et al.* Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359**, 91–97 (2018).

121. Gopalakrishnan, V. *et al.* Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **359**, 97–103 (2018).

122. Chaput, N. *et al.* Baseline gut microbiota predicts clinical response and colitis in metastatic melanoma patients treated with ipilimumab. *Ann Oncol* **28**, 1368–1379 (2017).

123. Wu, J. *et al.* Modulation of Gut Microbiota to Enhance Effect of Checkpoint Inhibitor Immunotherapy. *Front Immunol* **12**, 2554 (2021).

124. Berman, D. *et al.* Blockade of cytotoxic T-lymphocyte antigen-4 by ipilimumab results in dysregulation of gastrointestinal immunity in patients with advanced melanoma. *Cancer Immun* **10**, 11 (2010).

125. Dubin, K. *et al.* Intestinal microbiome analyses identify melanoma patients at risk for checkpoint-blockade-induced colitis. *Nat Commun* **7**, (2016).

126. Wang, Y. *et al.* Fecal microbiota transplantation for refractory immune checkpoint inhibitor-associated colitis. *Nature Medicine 2018 24:12* **24**, 1804–1808 (2018).

127. Dougan, M., Luoma, A. M., Dougan, S. K. & Wucherpfennig, K. W. Understanding and treating the inflammatory adverse events of cancer immunotherapy. *Cell* **184**, 1–14 (2021).

128. Sasson, S. C. *et al.* Interferon-Gamma–Producing CD8+ Tissue Resident Memory T Cells Are a Targetable Hallmark of Immune Checkpoint Inhibitor–Colitis. *Gastroenterology* **161**, 1229-1244.e9 (2021).

129. Som, A. *et al.* Immune checkpoint inhibitor-induced colitis: A comprehensive review. *World J Clin Cases* **7**, 405 (2019).

130. Hashash, J. G., Francis, F. F. & Farraye, F. A. Diagnosis and Management of Immune Checkpoint Inhibitor Colitis. *Gastroenterol Hepatol (N Y)* **17**, 358 (2021).

131. Bishu, S. *et al.* Efficacy and Outcome of Tofacitinib in Immune checkpoint Inhibitor Colitis. *Gastroenterology* **160**, 932-934.e3 (2021).

132. Zou, F. *et al.* Efficacy and safety of vedolizumab and infliximab treatment for immune-mediated diarrhea and colitis in patients with cancer: a two-center observational study. *J Immunother Cancer* **9**, e003277 (2021).

133. Raffals, L. E. *et al.* The Development and Initial Findings of A Study of a Prospective Adult Research Cohort with Inflammatory Bowel Disease (SPARC IBD). *Inflamm Bowel Dis* **28**, 192–199 (2022).

134. Venkateswaran, S. *et al.* Longitudinal DNA methylation profiling of the rectal mucosa identifies cell-specific signatures of disease status, severity and clinical outcomes in ulcerative colitis cell-specific DNA methylation signatures of UC. *Clin Epigenetics* **15**, 50 (2023).

135. Pascal, V. *et al.* A microbial signature for Crohn's disease. *Gut* **66**, 813–822 (2017).

136. Mehta, R. S. *et al.* Gut microbial metabolism of 5-ASA diminishes its clinical efficacy in inflammatory bowel disease. *Nature Medicine 2023 29:3* **29**, 700–709 (2023).

137. Ho, Y. T. *et al.* Longitudinal Single-Cell Transcriptomics Reveals a Role for Serpina3n-Mediated Resolution of Inflammation in a Mouse Colitis Model. *Cell Mol Gastroenterol Hepatol* **12**, 547–566 (2021).

138. Park, H., Yeo, S., Kang, S. & Huh, C. S. Longitudinal Microbiome Analysis in a Dextran Sulfate Sodium-Induced Colitis Mouse Model. *Microorganisms* **9**, 1–18 (2021).

139. Cadinu, P. *et al.* Charting the cellular biogeography in colitis reveals fibroblast trajectories and coordinated spatial remodeling. *bioRxiv* (2023) doi:10.1101/2023.05.08.539701.

140. Ho, Y. T. *et al.* Longitudinal Single-Cell Transcriptomics Reveals a Role for Serpina3n-Mediated Resolution of Inflammation in a Mouse Colitis Model. *Cell Mol Gastroenterol Hepatol* **12**, 547–566 (2021).

141. Langer, V. *et al.* IFN-γ drives inflammatory bowel disease pathogenesis through VE-cadherin–directed vascular barrier disruption. *J Clin Invest* **129**, 4691 (2019).

142. Ma, S., Zhang, J., Liu, H., Li, S. & Wang, Q. The Role of Tissue-Resident Macrophages in the Development and Treatment of Inflammatory Bowel Disease. *Front Cell Dev Biol* **10**, (2022).

143. Ghosh, S., Chaudhary, R., Carpani, M. & Playford, R. Interfering with interferons in inflammatory bowel disease. *Gut* **55**, 1071 (2006).

144. Reinisch, W. *et al.* Fontolizumab in moderate to severe Crohn's disease: a phase 2, randomized, double-blind, placebo-controlled, multiple-dose study. *Inflamm Bowel Dis* **16**, 233–242 (2010).

145. Thelemann, C. *et al.* Interferon-γ Induces Expression of MHC Class II on Intestinal Epithelial Cells and Protects Mice from Colitis. *PLoS One* **9**, (2014).

146. Fauny, M. *et al.* Paradoxical gastrointestinal effects of interleukin-17 blockers. *Ann Rheum Dis* **79**, 1132–1138 (2020).

147. Iriarte, A., Zaera, C., Bachiller-Corral, J. & López-Sanromán, A. Inflammatory bowel disease as a paradoxical effect of anti-TNF alpha therapy. *Gastroenterol Hepatol* **40**, 117–121 (2017).

148. Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119 (2012).

149. Sazonovs, A. *et al.* Large-scale sequencing identifies multiple genes and rare variants associated with Crohn's disease susceptibility. *Nature Genetics 2022 54:9* **54**, 1275–1283 (2022).

150. Liu, T. C. & Stappenbeck, T. S. Genetics and Pathogenesis of Inflammatory Bowel Disease. *Annu Rev Pathol* **11**, 127–148 (2016).

151. Uhlig, H. H. & Muise, A. M. Clinical Genomics in Inflammatory Bowel Disease. *Trends Genet* **33**, 629–641 (2017).

152. JBAG, H. *et al.* Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* **28**, iv119–iv142 (2017).

153. Saunders, M. P. *et al.* CXD101 and nivolumab in patients with metastatic microsatellite-stable colorectal cancer (CAROSELL): a multicentre, open-label, single-arm, phase II trial. *ESMO Open* **7**, (2022).

154. Taylor, C. A. *et al.* IL7 genetic variation and toxicity to immune checkpoint blockade in patients with melanoma. *Nature Medicine 2022 28:12* **28**, 2592–2600 (2022).

155. Groha, S. *et al.* Germline variants associated with toxicity to immune checkpoint blockade. *Nature Medicine 2022 28:12* **28**, 2584–2591 (2022).

156. He, Z. *et al.* Lineage recording in human cerebral organoids. *Nature Methods 2021 19:1* **19**, 90–99 (2021).

157. Truelove, S. C. & Witts, L. J. Cortisone in Ulcerative Colitis. *Br Med J* **2**, 1048 (1955).

158. Ho, G. T. *et al.* Predicting the outcome of severe ulcerative colitis: development of a novel risk score to aid early selection of patients for second-line medical therapy or surgery. *Aliment Pharmacol Ther* **19**, 1079–1087 (2004).

159. National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE) | Protocol Development | CTEP. https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm#ctc_50.

160. Hör, S. *et al.* The T-cell lymphokine interleukin-26 targets epithelial cells through the interleukin-20 receptor 1 and interleukin-10 receptor 2 chains. *J Biol Chem* **279**, 33343–33351 (2004).

161. Khare, V. *et al.* IL10R2 Overexpression Promotes IL22/STAT3 Signaling in Colorectal Carcinogenesis. *Cancer Immunol Res* **3**, 1227–1235 (2015).

162. Schoenborn, J. R. *et al.* Comprehensive epigenetic profiling identifies multiple distal regulatory elements directing transcription of the gene encoding interferon-gamma. *Nat Immunol* **8**, 732–742 (2007).

163. Collins, P. L., Henderson, M. A. & Aune, T. M. Lineage-specific adjacent IFNG and IL26 genes share a common distal enhancer element. *Genes Immun* **13**, 481–488 (2012).

164. Hatano, R. *et al.* Characterization of novel anti-IL-26 neutralizing monoclonal antibodies for the treatment of inflammatory diseases including psoriasis. *MAbs* **11**, 1428–1442 (2019).

165. Hatano, R. *et al.* Humanized anti-IL-26 monoclonal antibody as a novel targeted therapy for chronic graft-versus-host disease. *Am J Transplant* **22**, 2804–2820 (2022).

166. Chassaing, B., Aitken, J. D., Malleshappa, M. & Vijay-Kumar, M. Dextran sulfate sodium (DSS)-induced colitis in mice. *Curr Protoc Immunol* **104**, (2014).

167. Generation of Mo-DCs | All protocols | Applications | Miltenyi Biotec | Great Britain. https://www.miltenyibiotec.com/GB-en/applications/all-protocols/generation-of-mo-dcs.html.

168. Constant, D. A. *et al.* Transcriptional and Cytotoxic Responses of Human Intestinal Organoids to IFN Types I, II, and III. *Immunohorizons* **6**, 416–429 (2022).

169. Woznicki, J. A. *et al.* TNF-α synergises with IFN-γ to induce caspase-8-JAK1/2-STAT1-dependent death of intestinal epithelial cells. *Cell Death Dis* **12**, (2021).

170. Volk, V. *et al.* PD-1 Blockade Aggravates Epstein–Barr Virus+ Post-Transplant Lymphoproliferative Disorder in Humanized Mice Resulting in Central Nervous System Involvement and CD4+ T Cell Dysregulations. *Front Oncol* **10**, 1 (2021).

171. Zhou, J. *et al.* Potentiating Functional Antigen-specific CD8+ T Cell Immunity by a Novel PD1 Isoform-based Fusion DNA Vaccine. *Molecular Therapy* **21**, 1445–1455 (2013).

172. Fung, K. *et al.* Steric effects in PD1 binding of diagnostic anti-PD-1 antibodies with therapeutic anti-PD1 drugs, Pembrolizumab and Nivolumab. https://www.researchgate.net/publication/333433570_Steric_effects_in_PD1_binding_of_diagnostic_anti-PD-1_antibodies_with_therapeutic_anti-PD1_drugs_Pembrolizumab_and_Nivolumab.

173. Baker, L. A., Tiriac, H., Clevers, H. & Tuveson, D. A. Modeling Pancreatic Cancer with Organoids. *Trends in Cancer* vol. 2 176–190 Preprint at https://doi.org/10.1016/j.trecan.2016.03.004 (2016).

174. Neal, J. T. *et al.* Organoid Modeling of the Tumor Immune Microenvironment. *Cell* **175**, 1972-1988.e16 (2018).

175. Fujii, M. *et al.* Human Intestinal Organoids Maintain Self-Renewal Capacity and Cellular Diversity in Niche-Inspired Culture Condition. *Cell Stem Cell* **23**, 787-793.e6 (2018).

176. Neal, J. T. *et al.* Organoid Modeling of the Tumor Immune Microenvironment. *Cell* **175**, 1972-1988.e16 (2018).

177. PHA (Phytohaemagglutinin P) for T cells activation | InvivoGen. https://www.invivogen.com/phap.

178. Brahmer, J. R. *et al.* Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. *J Clin Oncol* **28**, 3167–75 (2010).

179. Andrews, S. FastQC: a quality control tool for high throughput sequence data. *https://www.bioinformatics.babraham.ac.uk/projects/fastqc/* (2010).

180. Downloads -Software -Single Cell Gene Expression -Official 10x Genomics Support. https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest.

181. Stoeckius, M. *et al.* Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol* **19**, 1–12 (2018).

182. Lun, A. T. L. *et al.* EmptyDrops: Distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol* **20**, 1–9 (2019).

183. Griffiths, J. A., Richard, A. C., Bach, K., Lun, A. T. L. & Marioni, J. C. Detection and removal of barcode swapping in single-cell RNA-seq data. *Nat Commun* **9**, (2018).

184. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411–420 (2018).

185. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. (2018) doi:10.48550/arxiv.1802.03426.

186. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289–1296 (2019).

187. Boland, B. S. *et al.* Heterogeneity and clonal relationships of adaptive immune cells in ulcerative colitis revealed by single-cell analyses. *Sci Immunol* **5**, 4432 (2020).

188. Szabo, P. A. *et al.* Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nature Communications 2019 10:1* **10**, 1–16 (2019).

189. Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* **14**, 979–982 (2017).

190. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, (2014).

191. Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D. & Marioni, J. C. Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat Biotechnol* **40**, 245–253 (2022).

192. Huynh-Thu, V. A., Irrthum, A., Wehenkel, L. & Geurts, P. Inferring regulatory networks from expression data using tree-based methods. *PLoS One* **5**, (2010).

193. Aibar, S. *et al.* SCENIC: Single-cell regulatory network inference and clustering. *Nat Methods* **14**, 1083 (2017).

194. Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**, (2021).

195. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25–29 (2000).

196. Fabregat, A. *et al.* The Reactome pathway Knowledgebase. *Nucleic Acids Res* **44**, D481–D487 (2016).

197. Song, L. *et al.* TRUST4: immune repertoire reconstruction from bulk and single-cell RNA-seq data. *Nature Methods 2021 18:6* **18**, 627–630 (2021).

198. Shugay, M. *et al.* VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. *PLoS Comput Biol* **11**, (2015).

199. Bagaev, D. v. *et al.* VDJdb in 2019: database extension, new analysis infrastructure and a T-cell receptor motif compendium. *Nucleic Acids Res* **48**, D1057–D1062 (2020).

200. Tickotsky, N., Sagiv, T., Prilusky, J., Shifrut, E. & Friedman, N. McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics* **33**, 2924–2929 (2017).

201. Huang, H., Wang, C., Rubelt, F., Scriba, T. J. & Davis, M. M. Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat Biotechnol* **38**, 1194–1202 (2020).

202. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20**, (2019).

203. Elosua-Bayes, M., Nieto, P., Mereu, E., Gut, I. & Heyn, H. SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res* **49**, E50 (2021).

204. Cable, D. M. *et al.* Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat Biotechnol* **40**, 517–526 (2022).

205. Csárdi, G. & Nepusz, T. The igraph software package for complex network research. *Corpus ID: 16923281* (2006).

206. Cabello-Aguilar, S. *et al.* SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. *Nucleic Acids Res* **48**, e55 (2020).

207. Jin, S. *et al.* Inference and analysis of cell-cell communication using CellChat. *Nature Communications 2021 12:1* **12**, 1–20 (2021).

208. Keir, M. E., Butte, M. J., Freeman, G. J. & Sharpe, A. H. PD-1 and Its Ligands in Tolerance and Immunity. *Annu Rev Immunol* **26**, 677–704 (2008).

209. Webster, R. M. The immune checkpoint inhibitors: where are we now? *Nature Reviews Drug Discovery 2014 13:12* **13**, 883–884 (2014).

210. Amaria, R. N. *et al.* Neoadjuvant relatlimab and nivolumab in resectable melanoma. *Nature 2022 611:7934* **611**, 155–160 (2022).

211. Amaria, R. N. *et al.* Neoadjuvant immune checkpoint blockade in high-risk resectable melanoma. *Nature Medicine 2018 24:11* **24**, 1649–1654 (2018).

212. Brahmer, J. R. *et al.* Management of Immune-Related Adverse Events in Patients Treated With Immune Checkpoint Inhibitor Therapy: American Society of Clinical Oncology Clinical Practice Guideline. *J Clin Oncol* **36**, 1714–1768 (2018).

213. Puzanov, I. *et al.* Managing toxicities associated with immune checkpoint inhibitors: consensus recommendations from the Society for Immunotherapy of Cancer (SITC) Toxicity Management Working Group. *J Immunother Cancer* **5**, (2017).

214. Travis, S. P. L. *et al.* Reliability and initial validation of the ulcerative colitis endoscopic index of severity. *Gastroenterology* **145**, 987–995 (2013).

215. Marchal-Bressenot, A. *et al.* Development and validation of the Nancy histological index for UC. *Gut* **66**, 43–49 (2017).

216. Corte, C. *et al.* Association between the ulcerative colitis endoscopic index of severity (UCEIS) and outcomes in acute severe ulcerative colitis. *J Crohns Colitis* **9**, 376–381 (2015).

217. Travis, S. P. L. *et al.* Predicting outcome in severe ulcerative colitis. *Gut* **38**, 905–910 (1996).

218. Wang, Y. *et al.* Endoscopic and Histologic Features of Immune Checkpoint Inhibitor-Related Colitis. *Inflamm Bowel Dis* **24**, 1695–1705 (2018).

219. Abu-Sbeih, H. *et al.* Importance of endoscopic and histological evaluation in the management of immune checkpoint inhibitor-induced colitis. *J Immunother Cancer* **6**, (2018).

220. Lees, C. W., Ironside, J., Wallace, W. A. H. & Satsangi, J. Resolution of non-small-cell lung cancer after withdrawal of anti-TNF therapy. *N Engl J Med* **359**, 320–321 (2008).

221. Bergqvist, V. *et al.* Vedolizumab treatment for immune checkpoint inhibitor-induced enterocolitis. *Cancer Immunol Immunother* **66**, 581–592 (2017).

222. Berkowitz, L. *et al.* Impact of cigarette smoking on the gastrointestinal tract inflammation: Opposing effects in Crohn's disease and ulcerative colitis. *Front Immunol* **9**, 74 (2018).

223. Ma, C. *et al.* CORE-IBD: A Multidisciplinary International Consensus Initiative to Develop a Core Outcome Set for Randomized Controlled Trials in Inflammatory Bowel Disease. *Gastroenterology* **163**, 950–964 (2022).

224. Powell, N. *et al.* British Society of Gastroenterology endorsed guidance for the management of immune checkpoint inhibitor-induced enterocolitis. *Lancet Gastroenterol Hepatol* **5**, 679–697 (2020).

225. Zou, F. *et al.* Fecal calprotectin concentration to assess endoscopic and histologic remission in patients with cancer with immune-mediated diarrhea and colitis. *J Immunother Cancer* **9**, (2021).

226. Sarlos, P. *et al.* Genetic update on inflammatory factors in ulcerative colitis: Review of the current literature. *World J Gastrointest Pathophysiol* **5**, 304 (2014).

227. Martini, E., Krug, S. M., Siegmund, B., Neurath, M. F. & Becker, C. Mend Your Fences: The Epithelial Barrier and its Relationship With Mucosal Immunity in Inflammatory Bowel Disease. *Cell Mol Gastroenterol Hepatol* **4**, 33–46 (2017).

228. Ni, J., Wu, G. D., Albenberg, L. & Tomov, V. T. Gut microbiota and IBD: causation or correlation? *Nat Rev Gastroenterol Hepatol* **14**, 573–584 (2017).

229. Bjerrum, J. T. *et al.* Transcriptional analysis of left-sided colitis, pancolitis, and ulcerative colitis-associated dysplasia. *Inflamm Bowel Dis* **20**, 2340–2352 (2014).

230. Arijs, I. *et al.* Mucosal gene signatures to predict response to infliximab in patients with ulcerative colitis. *Gut* **58**, 1612–1619 (2009).

231. Kiesler, P., Fuss, I. J. & Strober, W. Experimental Models of Inflammatory Bowel Diseases. *Cell Mol Gastroenterol Hepatol* **1**, 241–248 (2015).

232. Darch, K. M., Holland, T. L. & Spelman, L. J. Secukinumab-Induced Inflammatory Bowel Disease in a Patient Treated for Chronic Plaque Psoriasis and Psoriatic Arthritis: A Case Report and Review of the Role of Novel Biologic Agents Targeting the p19 Subunit of IL-23. *Case Rep Med* **2020**, (2020).

233. Miragaia, R. J. *et al.* Single-Cell Transcriptomics of Regulatory T Cells Reveals Trajectories of Tissue Adaptation. *Immunity* **50**, 493-504.e7 (2019).

234. Villani, A. C. *et al.* Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science (1979)* **356**, (2017).

235. Krensky, A. M. & Clayberger, C. Biology and clinical relevance of granulysin. *Tissue Antigens* **73**, 193 (2009).

236. Ward-Kavanagh, L. K., Lin, W. W., Sedy´, J. R., Sedy´, S. & Ware, C. F. The TNF Receptor Superfamily in Co-stimulating and Co-inhibitory Responses. *Immunity* **44**, 1005–1019 (2016).

237. Konduri, V. *et al.* CD8+CD161+ T-Cells: Cytotoxic Memory Cells With High Therapeutic Potential. *Front Immunol* **11**, 3621 (2021).

238. Sebzda, E., Zou, Z., Lee, J. S., Wang, T. & Kahn, M. L. Transcription factor KLF2 regulates the migration of naive T cells by restricting chemokine receptor expression patterns. *Nat Immunol* **9**, 292–300 (2008).

239. Gnanaprakasam, J. N. R. & Wang, R. MYC in Regulating Immunity: Metabolism and Beyond. *Genes 2017, Vol. 8, Page 88* **8**, 88 (2017).

240. Pearce, E. L. *et al.* Control of effector CD8+ T cell function by the transcription factor Eomesodermin. *Science* **302**, 1041–1043 (2003).

241. Huber, M. *et al.* IRF9 Prevents CD8+ T Cell Exhaustion in an Extrinsic Manner during Acute Lymphocytic Choriomeningitis Virus Infection. *J Virol* **91**, (2017).

242. Bahrami, S. & Drabløs, F. Gene regulation in the immediate-early response process. *Adv Biol Regul* **62**, 37–49 (2016).

243. Powell, M. D., Read, K. A., Sreekumar, B. K. & Oestreich, K. J. Ikaros zinc finger transcription factors: Regulators of cytokine signaling pathways and CD4+ T helper cell differentiation. *Front Immunol* **10**, 1299 (2019).

244. d'Aldebert, E. *et al.* Characterization of Human Colon Organoids From Inflammatory Bowel Disease Patients. *Front Cell Dev Biol* **8**, 363 (2020).

245. Petrie, E. J. *et al.* CD94-NKG2A recognition of human leukocyte antigen (HLA)-E bound to an HLA class I leader sequence. *J Exp Med* **205**, 725–735 (2008).

246. Dambacher, J. *et al.* The role of the novel Th17 cytokine IL-26 in intestinal inflammation. *Gut* **58**, 1207–1217 (2009).

247. Sheikh, F. *et al.* Cutting Edge: IL-26 Signals through a Novel Receptor Complex Composed of IL-20 Receptor 1 and IL-10 Receptor 2. *The Journal of Immunology* **172**, 2006–2010 (2004).

248. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science (1979)* **347**, (2015).

249. Larochette, V. *et al.* IL-26, a Cytokine With Roles in Extracellular DNA-Induced Inflammation and Microbial Defense. *Front Immunol* **10**, (2019).

250. Yu, D. *et al.* Super-Enhancer Induced IL-20RA Promotes Proliferation/Metastasis and Immune Evasion in Colorectal Cancer. *Front Oncol* **11**, (2021).

251. Khare, V. *et al.* IL10R2 Overexpression Promotes IL22/STAT3 Signaling in Colorectal Carcinogenesis. *Cancer Immunol Res* **3**, 1227–1235 (2015).

252. Itoh, T. *et al.* Biological Effects of IL-26 on T Cell–Mediated Skin Inflammation, Including Psoriasis. *Journal of Investigative Dermatology* **139**, 878–889 (2019).

253. Broux, B. *et al.* Interleukin-26, preferentially produced by TH17 lymphocytes, regulates CNS barrier function. *Neurology - Neuroimmunology Neuroinflammation* **7**, (2020).

254. Perez-Ruiz, E. *et al.* Prophylactic TNF blockade uncouples efficacy and toxicity in dual CTLA-4 and PD-1 immunotherapy. *Nature* 1 (2019) doi:10.1038/s41586-019-1162-y.

255. 042058 - Ctla-4[-] Strain Details. https://www.jax.org/strain/029905.

256. 028276 - PD-1[-] Strain Details. https://www.jax.org/strain/028276.

257. Perez-Ruiz, E. *et al.* Prophylactic TNF blockade uncouples efficacy and toxicity in dual CTLA-4 and PD-1 immunotherapy. *Nature 2019 569:7756* **569**, 428–432 (2019).

258. Guilliams, M. *et al.* Dendritic cells, monocytes and macrophages: A unified nomenclature based on ontogeny. *Nature Reviews Immunology* vol. 14 571–578 Preprint at https://doi.org/10.1038/nri3712 (2014).

259. Smillie, C. S. *et al.* Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714-730.e22 (2019).

260. Kinchen, J. *et al.* Structural Remodeling of the Human Colonic Mesenchyme in Inflammatory Bowel Disease. *Cell* **175**, 372-386.e17 (2018).

261. Waeytens, A., de Vos, M. & Laukens, D. Evidence for a potential role of metallothioneins in inflammatory bowel diseases. *Mediators Inflamm* **2009**, (2009).

262. Bain, C. C. & Schridde, A. Origin, Differentiation, and Function of Intestinal Macrophages. *Front Immunol* **0**, 2733 (2018).

263. Ruytinx, P., Proost, P., van Damme, J. & Struyf, S. Chemokine-Induced Macrophage Polarization in Inflammatory Conditions. *Front Immunol* **9**, (2018).

264. Mantovani, A. *et al.* The chemokine system in diverse forms of macrophage activation and polarization. *Trends Immunol* **25**, 677–686 (2004).

265. Hoffman, J. M., Rankin, C. R. & Pothoulakis, C. Here to Heal: Mucosal CD74 Signaling in Colitis. *CMGH* vol. 10 197–198 Preprint at https://doi.org/10.1016/j.jcmgh.2020.03.002 (2020).

266. Boyapati, R. K., Rossi, A. G., Satsangi, J. & Ho, G. T. Gut mucosal DAMPs in IBD: from mechanisms to therapeutic implications. *Mucosal Immunol* **9**, 567–582 (2016).

267. Cheng, L. *et al.* Monoclonal antibodies specific to human Δ42PD1: A novel immunoregulator potentially involved in HIV-1 and tumor pathogenesis. *MAbs* **7**, 620–629 (2015).

268. Zelba, H. *et al.* Accurate quantification of T-cells expressing PD-1 in patients on anti-PD-1 immunotherapy. *Cancer Immunol Immunother* **67**, 1845–1851 (2018).

269. Clouthier, D. L. *et al.* An interim report on the investigator-initiated phase 2 study of pembrolizumab immunological response evaluation (INSPIRE). *J Immunother Cancer* **7**, 72 (2019).

270. Brahmer, J. R. *et al.* Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. *J Clin Oncol* **28**, 3167–3175 (2010).

271. Meinshausen, N. Quantile Regression Forests. *Journal of Machine Learning Research* **7**, 983–999 (2006).

272. Liu, J. *et al.* PD-1/PD-L1 Checkpoint Inhibitors in Tumor Immunotherapy. *Front Pharmacol* **12**, (2021).

273. Luoma, A. M. *et al.* Molecular Pathways of Colon Inflammation Induced by Cancer Immunotherapy. *Cell* **182**, 655-671.e22 (2020).

274. Franken, A. *et al.* Single-cell transcriptomics identifies pathogenic T-helper 17.1 cells and pro-inflammatory monocytes in immune checkpoint inhibitor-related pneumonitis. *J Immunother Cancer* **10**, e005323 (2022).

275. Kim, G., Hong, C. & Park, J. Seeing Is Believing: Illuminating the Source of In Vivo Interleukin-7. *Immune Netw* **11**, 1–10 (2011).

276. Hammoudi, N. *et al.* Autologous organoid co-culture model reveals T cell-driven epithelial cell death in Crohn's Disease. *Front Immunol* **13**, 6786 (2022).

277. Nozaki, K. *et al.* Co-culture with intestinal epithelial organoids allows efficient expansion and motility analysis of intraepithelial lymphocytes. *J Gastroenterol* **51**, 206 (2016).

278. Cattaneo, C. M. *et al.* Tumor organoid–T-cell coculture systems. *Nat Protoc* **15**, 15–39 (2020).

279. Qin, X. *et al.* Cell-type-specific signaling networks in heterocellular organoids. *Nat Methods* **17**, 335–342 (2020).

280. Min, S., Kim, S. & Cho, S. W. Gastrointestinal tract modeling using organoids engineered with cellular and microbiota niches. *Experimental and Molecular Medicine* vol. 52 227–237 Preprint at https://doi.org/10.1038/s12276-020-0386-0 (2020).

281. Wang, T. *et al.* Probiotics Lactobacillus reuteri Abrogates Immune Checkpoint Blockade-Associated Colitis by Inhibiting Group 3 Innate Lymphoid Cells. *Front Immunol* **10**, (2019).

282. Bayless, N. L. *et al.* Development of preclinical and clinical models for immune-related adverse events following checkpoint immunotherapy: a perspective from SITC and AACR. *J Immunother Cancer* **9**, e002627 (2021).

283. Heul, A. ver & Stappenbeck, T. Establishing A Mouse Model of Checkpoint Inhibitor-Induced Colitis: Pilot Experiments and Future Directions. *Journal of Allergy and Clinical Immunology* **141**, AB119 (2018).

284. Adam, K., Iuga, A., Tocheva, A. S. & Mor, A. A novel mouse model for checkpoint inhibitor-induced adverse events. *PLoS One* **16**, (2021).

285. Cheung, V. *et al.* Immunotherapy-related hepatitis: Real-world experience from a tertiary centre. *Frontline Gastroenterol* **10**, 364–371 (2019).

286. Cheung, V. *et al.* Immunotherapy-related hepatitis: real-world experience from a tertiary centre. *Gastroenterology* 1–8 (2019) doi:10.1136/flgastro-2018-101146.

287. Cheung, V. *et al.* PTU-061 Immunotherapy-related gastritis in a tertiary oncology centre. *Gut* **68**, A147–A147 (2019).

288. CosMx SMI Overview - Single-Cell Imaging - NanoString. https://nanostring.com/products/cosmx-spatial-molecular-imager/single-cell-imaging-overview/?gclid=Cj0KCQiAtvSdBhD0ARIsAPf8oNmF1TVD8lE4034AMqr83OPWO5qCjiRjabr6SoB_HnXlcPDfMoVBbK8aAv7yEALw_wcB.

289. PRISE study: Investigating Cancer Immunotherapy Colitis - Oxford Talks. https://talks.ox.ac.uk/talks/id/ff3bb537-f931-48a8-9099-419d730355c2/.

290. Gopalakrishnan, V. *et al.* Gut microbiome modulates response to anti–PD-1 immunotherapy in melanoma patients. *Science* **359**, 97 (2018).

291. Balachandran, V. P. *et al.* Identification of unique neoantigen qualities in long-term survivors of pancreatic cancer. *Nature* **551**, S12–S16 (2017).

292. Li, X., Zhang, S., Guo, G., Han, J. & Yu, J. Gut microbiome in modulating immune checkpoint inhibitors. *EBioMedicine* **82**, 104163 (2022).

293. Choi, Y. *et al.* Immune checkpoint blockade induces gut microbiota translocation that augments extraintestinal anti-tumor immunity. *bioRxiv* 2022.01.26.477865 (2022) doi:10.1101/2022.01.26.477865.

294. Iida, N. *et al.* Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science* **342**, 967–970 (2013).

295. Peng, Y. *et al.* An immunodominant NP105–113-B*07:02 cytotoxic T cell response controls viral replication and is associated with less severe COVID-19 disease. *Nature Immunology 2021 23:1* **23**, 50–61 (2021).

296. Perkel, J. M. Single-cell analysis enters the multiomics age. *Nature* **595**, 614–616 (2021).

297. Single Cell Multiome ATAC + Gene Expression - 10x Genomics. https://www.10xgenomics.com/products/single-cell-multiome-atac-plus-gene-expression.

298. Kaech, S. M. & Cui, W. Transcriptional control of effector and memory CD8+ T cell differentiation. *Nat Rev Immunol* **12**, 749 (2012).

299. Chen, Y., Zander, R., Khatun, A., Schauder, D. M. & Cui, W. Transcriptional and Epigenetic Regulation of Effector and Memory CD8 T Cell Differentiation. *Front Immunol* **9**, 2826 (2018).

300. Belarif, L. *et al.* IL-7 receptor blockade blunts antigen-specific memory T cell responses and chronic inflammation in primates. *Nature Communications 2018 9:1* **9**, 1–13 (2018).

301. Christie, D. & Zhu, J. Transcriptional regulatory networks for CD4 T cell differentiation. *Curr Top Microbiol Immunol* **381**, 125–172 (2014).

302. Koizumi, S. & Ishikawa, H. Transcriptional Regulation of Differentiation and Functions of Effector T Regulatory Cells. *Cells* **8**, 939 (2019).

303. Shevyrev, D. & Tereshchenko, V. Treg Heterogeneity, Function, and Homeostasis. *Front Immunol* **10**, 3100 (2019).

304. Behr, F. M., Chuwonpad, A., Stark, R. & Gisbergen, K. P. J. M. van. Armed and Ready: Transcriptional Regulation of Tissue-Resident Memory CD8 T Cells. *Front Immunol* **9**, 1770 (2018).

305. Mielke, L. A. *et al.* TCF-1 limits the formation of Tc17 cells via repression of the MAF–RORγt axis. *J Exp Med* **216**, 1682 (2019).

306. Yamauchi, T. *et al.* T-cell CX3CR1 expression as a dynamic blood-based biomarker of response to immune checkpoint inhibitors. *Nature Communications 2021 12:1* **12**, 1–14 (2021).

307. Zheng, C. *et al.* Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell* **169**, 1342-1356.e16 (2017).

308. Leeansyah, E. *et al.* Arming of MAIT Cell Cytolytic Antimicrobial Activity Is Induced by IL-7 and Defective in HIV-1 Infection. *PLoS Pathog* **11**, (2015).

309. Gérart, S. *et al.* Human iNKT and MAIT cells exhibit a PLZF-dependent proapoptotic propensity that is counterbalanced by XIAP. *Blood* **121**, 614–623 (2013).

310. Arbogast, A. *et al.* Failure of T Lymphocytes from Elderly Humans to Enter the Cell Cycle Is Associated with Low Cdk6 Activity and Impaired Phosphorylation of Rb Protein. *Cell Immunol* **197**, 46–54 (1999).

311. Shi, M., Lin, T. H., Appell, K. C. & Berg, L. J. Cell cycle progression following naïve T cell activation is independent of Jak3/γc cytokine signals. *J Immunol* **183**, 4493 (2009).

312. Parikh, K. *et al.* Colonic epithelial cell diversity in health and inflammatory bowel disease. *Nature* **567**, 49–55 (2019).

313. Shoshkes-Carmel, M. *et al.* Subepithelial telocytes are an important source of Wnts that supports intestinal crypts. *Nature* **557**, 242 (2018).

314. Pilling, D., Vakil, V., Cox, N. & Gomer, R. H. TNF-α–stimulated fibroblasts secrete lumican to promote fibrocyte differentiation. *Proc Natl Acad Sci U S A* **112**, 11929 (2015).

315. Kosinski, C. *et al.* Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. *Proc Natl Acad Sci U S A* **104**, 15418 (2007).

316. Brügger, M. D., Valenta, T., Fazilaty, H., Hausmann, G. & Basler, K. Distinct populations of crypt-associated fibroblasts act as signaling hubs to control colon homeostasis. *PLoS Biol* **18**, (2020).

317. Huan, C. *et al.* Gremlin2 Activates Fibroblasts to Promote Pulmonary Fibrosis Through the Bone Morphogenic Protein Pathway. *Front Mol Biosci* **0**, 619 (2021).

318. Uhlitz, F. *et al.* Mitogen-activated protein kinase activity drives cell trajectories in colorectal cancer. *bioRxiv* 2020.01.10.901579 (2021) doi:10.1101/2020.01.10.901579.

319. Travaglini, K. J. *et al.* A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature 2020 587:7835* **587**, 619–625 (2020).

320. Halpern, K. B. *et al.* Lgr5+ telocytes are a signaling source at the intestinal villus tip. *Nat Commun* **11**, (2020).

321. Melissari, M.-T. *et al.* Col6a1+/CD201+ mesenchymal cells regulate intestinal morphogenesis and homeostasis. *bioRxiv* 2021.02.16.431453 (2021) doi:10.1101/2021.02.16.431453.

322. Popescu, L. M. & Faussone-Pellegrini, M.-S. TELOCYTES – a case of serendipity: the winding way from Interstitial Cells of Cajal (ICC), via Interstitial Cajal-Like Cells (ICLC) to TELOCYTES. *J Cell Mol Med* **14**, 729 (2010).

323. Qi, Z. *et al.* BMP restricts stemness of intestinal Lgr5+ stem cells by directly suppressing their signature genes. *Nat Commun* **8**, (2017).

324. Kosinski, C. *et al.* Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. *Proc Natl Acad Sci U S A* **104**, 15418 (2007).

325. El-Salhy, M. & Hausken, T. The role of the neuropeptide Y (NPY) family in the pathophysiology of inflammatory bowel disease (IBD). *Neuropeptides* **55**, 137–144 (2016).

326. Chen, W. *et al.* Neuropeptide Y Is an Immunomodulatory Factor: Direct and Indirect. *Front Immunol* **11**, (2020).

327. Cox, H. M. Neuropeptide Y receptors; antisecretory control of intestinal epithelial function. *Auton Neurosci* **133**, 76–85 (2007).

328. Harnack, C. *et al.* R-spondin 3 promotes stem cell recovery and epithelial regeneration in the colon. *Nature Communications 2019 10:1* **10**, 1–15 (2019).

329. Villar, J. & Segura, E. Decoding the Heterogeneity of Human Dendritic Cell Subsets. *Trends Immunol* **41**, 1062–1071 (2020).

330. Gautier, E. L. *et al.* Gene expression profiles and transcriptional regulatory pathways underlying mouse tissue macrophage identity and diversity. *Nat Immunol* **13**, 1118 (2012).

331. da Silva, F. A. R. *et al.* Whole transcriptional analysis identifies markers of B, T and plasma cell signaling pathways in the mesenteric adipose tissue associated with Crohn's disease. *Journal of Translational Medicine 2020 18:1* **18**, 1–14 (2020).

332. Stewart, A. *et al.* Single-Cell Transcriptomic Analyses Define Distinct Peripheral B Cell Subsets and Discrete Development Pathways. *Front Immunol* **12**, (2021).

333. Sanz, I. *et al.* Challenges and Opportunities for Consistent Classification of Human B Cell and Plasma Cell Populations. *Front Immunol* **10**, (2019).

334. Jaeger, N. *et al.* Single-cell analyses of Crohn's disease tissues reveal intestinal intraepithelial T cells heterogeneity and altered subset distributions. *Nat Commun* **12**, (2021).

335. Capone, A. & Volpe, E. Transcriptional Regulators of T Helper 17 Cell Differentiation in Health and Autoimmune Diseases. *Front Immunol* **11**, (2020).

336. Yamazaki, S. *et al.* The AP-1 transcription factor JunB is required for Th17 cell differentiation. *Sci Rep* **7**, (2017).

337. Zhou, L. How smart can it be: transcriptional regulation of T helper cells by SMAR1. *Mucosal Immunol* **8**, 1181–1183 (2015).

338. Evans, C. & Jenner, R. Transcription factor interplay in T helper cell differentiation. *Brief Funct Genomics* **12**, 499–511 (2013).

339. Oestreich, K. & Weinmann, A. Transcriptional mechanisms that regulate T helper 1 cell differentiation. *Curr Opin Immunol* **24**, 191–195 (2012).

340. Egawa, T., Tillman, R. E., Naoe, Y., Taniuchi, I. & Littman, D. R. The role of the Runx transcription factors in thymocyte differentiation and in homeostasis of naive T cells. *J Exp Med* **204**, 1945 (2007).

341. Willinger, T. *et al.* Human naive CD8 T cells down-regulate expression of the WNT pathway transcription factors lymphoid enhancer binding factor 1 and transcription factor 7 (T cell factor-1) following antigen encounter in vitro and in vivo. *J Immunol* **176**, 1439–1446 (2006).

342. Feng, X. *et al.* Transcription factor Foxp1 exerts essential cell-intrinsic regulation of the quiescence of naive T cells. *Nat Immunol* **12**, 544 (2011).

343. Parker, M. E. & Ciofani, M. Regulation of γδ T Cell Effector Diversification in the Thymus. *Front Immunol* **11**, 42 (2020).

344. Barros-Martins, J. *et al.* Effector γδ T Cell Differentiation Relies on Master but Not Auxiliary Th Cell Transcription Factors. *The Journal of Immunology* **196**, 3642–3652 (2016).

345. Vandereyken, M., James, O. J. & Swamy, M. Mechanisms of activation of innate-like intraepithelial T lymphocytes. *Mucosal Immunology 2020 13:5* **13**, 721–731 (2020).

346. Kumar, B. v. *et al.* Human tissue-resident memory T cells are defined by core transcriptional and functional signatures in lymphoid and mucosal sites. *Cell Rep* **20**, 2921 (2017).

347. Kragten, N. A. M. *et al.* Blimp-1 induces and Hobit maintains the cytotoxic mediator granzyme B in CD8 T cells. *Eur J Immunol* **48**, 1644–1662 (2018).

348. Caushi, J. X. *et al.* Transcriptional programs of neoantigen-specific TIL in anti-PD-1-treated lung cancers. *Nature 2021 596:7870* **596**, 126–132 (2021).

349. Viganò, S. *et al.* CD160-Associated CD8 T-Cell Functional Impairment Is Independent of PD-1 Expression. *PLoS Pathog* **10**, (2014).

350. Scott, C. L. & Omilusik, K. D. ZEBs: Novel Players in Immune Cell Development and Function. *Trends Immunol* **40**, 431–446 (2019).

351. Parish, I. A. & Kaech, S. M. Diversity in CD8+ T cell differentiation. *Curr Opin Immunol* **21**, 291 (2009).

352. Gegonne, A. *et al.* The General Transcription Factor TAF7 Is Essential for Embryonic Development but Not Essential for the Survival or Differentiation of Mature T Cells. *Mol Cell Biol* **32**, 1984 (2012).

353. Luo, C. T. *et al.* Ets transcription factor GABP controls T cell homeostasis and immunity. *Nat Commun* **8**, (2017).

354. Guo, Y. *et al.* GABPA is a master regulator of luminal identity and restrains aggressive diseases in bladder cancer. *Cell Death Differ* **27**, 1862 (2020).

355. Xue, H.-H. *et al.* Targeting the GA Binding Protein β1L Isoform Does Not Perturb Lymphocyte Development and Function. *Mol Cell Biol* **28**, 4300 (2008).

356. Su, G. H. *et al.* The Ets protein Spi-B is expressed exclusively in B cells and T cells during development. *J Exp Med* **184**, 203 (1996).

357. Kanaya, T. *et al.* The Ets Transcription Factor Spi-B Is Essential for the Differentiation of Intestinal Microfold (M) Cells. *Nat Immunol* **13**, 729 (2012).

358. Willis, S. N. *et al.* Environmental sensing by mature B cells is controlled by the transcription factors PU.1 and SpiB. *Nat Commun* **8**, (2017).

359. David-Fung, E.-S. *et al.* Transcription factor expression dynamics of early T-lymphocyte specification and commitment. *Dev Biol* **325**, 444 (2009).

360. Bhullar, J. & Sollars, V. E. YBX1 expression and function in early hematopoiesis and leukemic cells. *Immunogenetics 2011 63:6* **63**, 337–350 (2011).

361. Cobaleda, C., Jochum, W. & Busslinger, M. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature 2007 449:7161* **449**, 473–477 (2007).

362. Seo, W., Jerin, C. & Nishikawa, H. Transcriptional regulatory network for the establishment of CD8+ T cell exhaustion. *Exp Mol Med* **53**, 202 (2021).

363. Chen, Z. *et al.* TCF-1-centered transcriptional network drives an effector versus exhausted CD8 T cell fate decision. *Immunity* **51**, 840 (2019).

364. Li, J., He, Y., Hao, J., Ni, L. & Dong, C. High Levels of Eomes Promote Exhaustion of Anti-tumor CD8+ T Cells. *Front Immunol* **9**, 2981 (2018).

365. Castro, G. *et al.* RORγt and RORα signature genes in human Th17 cells. *PLoS One* **12**, (2017).

366. Gavins, F. N. E. & Hickey, M. J. Annexin A1 and the regulation of innate and adaptive immunity. *Front Immunol* **3**, (2012).

367. Ramesh, R. *et al.* Pro-inflammatory human Th17 cells selectively express P-glycoprotein and are refractory to glucocorticoids. *J Exp Med* **211**, 89 (2014).

368. Singh, S. P., Zhang, H. H., Foley, J. F., Hedrick, M. N. & Farber, J. M. Human T Cells That Are Able to Produce IL-17 Express the Chemokine Receptor CCR6. *The Journal of Immunology* **180**, 214–221 (2008).

369. Hudak, S. *et al.* Immune Surveillance and Effector Functions of CCR10+ Skin Homing T Cells. *The Journal of Immunology* **169**, 1189–1196 (2002).

370. Andrew, D. P. *et al.* C-C Chemokine Receptor 4 Expression Defines a Major Subset of Circulating Nonintestinal Memory T Cells of Both Th1 and Th2 Potential. *The Journal of Immunology* **166**, 103–111 (2001).

371. Nakatani, T. *et al.* CCR4+ memory CD4+ T lymphocytes are increased in peripheral blood and lesional skin from patients with atopic dermatitis. *Journal of Allergy and Clinical Immunology* **107**, 353–358 (2001).

372. Soler, D., Humphreys, T. L., Spinola, S. M. & Campbell, J. J. CCR4 versus CCR10 in human cutaneous TH lymphocyte trafficking. *Blood* **101**, 1677–1683 (2003).

373. Batista, N. v. *et al.* T Cell–Intrinsic CX3CR1 Marks the Most Differentiated Effector CD4+ T Cells, but Is Largely Dispensable for CD4+ T Cell Responses during Chronic Viral Infection. *Immunohorizons* **4**, 701–712 (2020).

374. Gerlach, C. *et al.* The chemokine receptor CX3CR1 defines three antigen-experienced CD8 T cell subsets with distinct roles in immune surveillance and homeostasis. *Immunity* **45**, 1270 (2016).

375. Nishimura, M. *et al.* Dual Functions of Fractalkine/CX3C Ligand 1 in Trafficking of Perforin+/Granzyme B+ Cytotoxic Effector Lymphocytes That Are Defined by CX3CR1 Expression. *The Journal of Immunology* **168**, 6173–6180 (2002).

376. Kobayashi, T. *et al.* Exclusive increase of CX3CR1+CD28-CD4+ T cells in inflammatory bowel disease and their recruitment as intraepithelial lymphocytes. *Inflamm Bowel Dis* **13**, 837–846 (2007).

377. Omilusik, K. D. *et al.* Transcriptional repressor ZEB2 promotes terminal differentiation of CD8+ effector and memory T cell populations during infection. *J Exp Med* **212**, 2027 (2015).

378. Li, H. *et al.* Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell* **176**, 775-789.e18 (2019).

379. Viallard, J. F. *et al.* HLA-DR expression on lymphocyte subsets as a marker of disease activity in patients with systemic lupus erythematosus. *Clin Exp Immunol* **125**, 485 (2001).

380. Chiba, A., Murayama, G. & Miyake, S. Mucosal-Associated Invariant T Cells in Autoimmune Diseases. *Front Immunol* **0**, 1333 (2018).