Identification of Crohn's disease immunopathotypes

Hannah Moana Baer

BSc (Hons) in Biochemistry

MSc in Immunology and Inflammatory Diseases

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

School of Infection and Immunity

College of Medical, Veterinary and Life Sciences

University of Glasgow

October 2022

Abstract

Crohn's disease (CD) describes an inflammatory condition characterised by chronic gastrointestinal (GI) inflammation. Currently, >1 in 150 people are estimated to be affected in the UK, with the incidence rising. The pathology of CD is linked to an aberrant immune response against the gut microbiota. CD is a heterogenous disease with different disease phenotypes (stricturing and fistulating disease) and locations; it can affect any part of the GI tract. A range of drugs is approved for CD treatment. However, response rates to these specific therapies are low, and 46% of patients still require bowel resective surgery.

The high rates of surgical interventions and unpredictable disease relapse pose challenges for the clinical management of CD. Little is known about the molecular mechanisms underlying the different treatment responses, and cost-effective and feasible approaches to predict these responses are still a long way off. Disease relapse has been linked with decreased T cell immunoreceptor with Ig and ITIM (TIGIT) on $CD38^+$ $CD4^+$ T cells in the peripheral blood of paediatric patients. Other clinical markers of inflammation have been found to lack specificity in predicting disease progression. Several biomarkers predicting non-response to frontline therapeutics blocking the pro-inflammatory cytokine tumour necrosis factor (TNF) have been proposed, including the levels of luminal membrane-bound TNF and gene expression of the pro-inflammatory cytokines oncostatin-M (OSM) and TNF itself. None of these markers are currently used clinically, and there is still a need for reliable ways to predict disease progression and treatment responses. Based on these findings I hypothesised that distinct immunopathotypes in CD drive distinct disease phenotypes and underlie distinct treatment responses to biologic therapies, such as anti-TNF.

My project aimed to elucidate distinct immunopathotypes in the blood and intestines of CD patients and link them to disease phenotype, state of inflammation and treatment response. Two distinct immunopathotypes were identified. One showed decreased circulating T cells and increased intestinal $CD8^+$ and cytotoxic immune responses. The second immunopathotype was defined by IL-1β-driven inflammation and granuloma formation in the intestines of people and was linked to anti-TNF refractory disease.

Table of Content

List of Tables

List of Figures

Acknowledgements

First of all, I would like to thank Simon for being an amazing supervisor and mentor for the last four years. The biggest thank you for supporting (most of) my crazy science ideas and ventures (even when totally unrelated to my PhD), only mildly judging me for being hungover and calming me down when I randomly get stressed about ethics. I am glad I got to work with you and am extremely thankful for all the time you have invested in me. Thank you also to Rob for making this PhD project possible and giving great advice and positive reassurance whenever needed. Dan, I already miss my Friday clinic adventures and chats with you and the nurses at the GRI. You have made a huge contribution to my project, and without all your hard work and support, my research would not have been possible, so thank you! The same goes for all the endoscopy staff at the GRI, thanks for always welcoming me with open arms and making Fridays a day to look forward to. I would also especially like to thank Ann-Marie from the gastroenterology day unit - you have taken a lot of time out of your very busy schedule to supply me with samples (and biscuits).

I also owe a big thank you to the staff from the histology and flow core facilities. Thanks for keeping me company and letting me work in your histology realm, Fiona, and for always having an ear and solution for my IHC problems. I am glad we got to spend so much time together. Diane and Alana, thanks so much for all your hard work around keeping our flow core as amazing as it is, for giving the best knitting tips and Diane, thanks so much for sharing your incredible flow knowledge with me whenever needed. John and Umer, thank you for your constant bioinformatics advice and mentoring.

To all the Mowlings, Milloys and MMPs (past and present), and my ladies from the science chat, you guys have been amazing! I am very thankful that I got to get to know so many of you outside of work, which made my PhD experience even more special. Kymbo, my human partner in crime, I could have literally not done this without you. Thanks for all your moral and physical support, your enthusiasm and most importantly, for watching me do the scary multiplexes so I won't have to do them myself. Anna Banana, you have been there through my entire PhD journey. Thanks for always giving amazing advice and for making The Chip roof part of my

life. Bella (I am still working on Annabelle), thanks for all the work and boy chats and for making my life a lot more organized. And a big thank you to my new lab peeps in Vancouver for getting me through the final writing period – Chad, Stacy, Shruti and Andrew you have been the best (eh).

I owe the biggest thank you to my Glasgow fam (also known as the cheese chat). I am so glad I met you all during my time in Glasgow, thanks so much for getting me through this with unlimited supplies of prosecco and love. George, thank you for sharing a workspace and home with me for most of my PhD and all the work and life pep talks (and gins). Han and Chris, there is literally not enough space to thank you guys enough for everything have done for me.

To my family at home, thanks for supporting me from afar. Lauri and Nikki, I could not have done any of this without the three musketeers. Thanks for all the pep talks and pandemic game nights. Suse and Romi, thanks for always believing in me. Finally, Mama and Papa, thanks for always just being a phone call away, and your unlimited support and love. You two are the best!

Impact of the COVID-19 Pandemic

The COVID-19 pandemic has significantly impacted my PhD project. After obtaining ethical approval a year into my work (October 2019) to be able to study all of my hypotheses, research sample collection was put on hold on the 21st of March 2020. Sample collection could be temporarily resumed from October 2020 onwards. However, since my presence in the hospital was required for the collection, this was only possible during certain lockdown phases. Due to staff shortage and illness, the gastroenterology clinic at the GRI, which previously supplied me with blood samples, had to be discontinued until the end of 2021. Additionally, the specimens I received in the endoscopy unit in the meantime included human intestinal tissue and blood samples which have previously been shown to carry the SARS-CoV-2 virus upon infection. Processing of these specimen types had to be reassessed based on health and safety by the research institute to ensure adequate handling of potentially COVID-19-infected samples. This review was finalised in December 2020. For most of 2020 and 2021, the university policy was to try and work from home when possible. Lab space was limited due to social distancing, which significantly reduced my productivity even when processing banked samples. Overall, I tried to mitigate these issues by using publicly deposited datasets, improving my computational biology skills and optimising histology protocols in samples provided by the biobank.

Author's declaration

"I declare that, except where reference is made to the contribution of others, this thesis is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution".

Signature: Hannah Moana Baer

Abbreviations

2-ΔΔCT – Delta-Delta Cycle Threshold

5-ASA – aminosalicylates

A

Ab – antibody

ACTB – actin beta

Ada – Adalimumab

AIC – Akaike Information Criterion

AMP - antimicrobial peptide

ANOVA – Analysis of variance

APC – antigen presenting cell

ASCA – anti–Saccharomyces cerevisiae antibody

ATG16L1 – autophagy-related 16 like 1

AUC – Area Under Curve

B

BATF – basic leucine zipper activating transcription factor (ATF)-like transcription factor

C

C1s – complement component 1S

C3 – complement component 3

C4 – complement component 4

CARD15 – caspase recruitment domain family member 15

CCL – chemokine ligand

CCR – chemokine receptor

CD – Crohn’s disease

CD1 – Crohn's disease subset 1

CD11c – integrin alpha subunit X

CD14 – cluster of differentiation 14

CD16 – FC-gamma receptor 3A

CD2 – Crohn's disease subset 2

CD25 – IL-2 receptor alpha chain

CD27 – cluster of differentiation 27

CD28 – cluster of differentiation 28

CD3D – CD3 delta chain

CD3E – CD3 epsilon chain

CD3G – CD3 gamma chain

CD38 – ADP-ribose cyclase 1

CD4 – cluster of differentiation 4

CD45 – leukocyte common antigen

CD45RA – isoform of protein tyrosine phosphatase receptor type C

CD56 – neural cell adhesion molecule

CD62L – L-selectin

CD64 – FC gamma receptor 1

CD7 – cluster of differentiation 7

CD8 – cluster of differentiation 8

CD8A – CD8 alpha chain

CD8B – CD8 beta chain

CDAI – Crohn’s disease activity index

CDEIS – Crohn’s disease endoscopy index of severity

CDG – chronic granulomatous disease

cDNA – complementary deoxyribonucleic acid

CDR – Crohn's disease remission

CHI3L1 – Chitinase-like 3 protein 1

CM – central memory

CRP – C-reactive protein

CRTH2 – prostaglandin D2 receptor

CSF1 – colony-stimulating factor 1

CTL – cytotoxic T lymphocyte

CXCL – C-X-C Motif Chemokine Ligand

CXCR – C-X-C Motif Chemokine Receptor

D

DAB – 3,3'-Diaminobenzidine

DAMP – damage-associated molecular pattern

dapB – L-2,3-dihydrodipicolinate reductase

DC – dendritic cell

DEG – differential expressed genes

DIABLO – Data integration Analysis for Biomarker discovery using Latent variable approaches for Omics studies

DP – double positive

E

E. coli – *Escherichia coli*

EBI3 - Epstein-Barr virus-inducible protein 3

ECCO – European Crohn's and Colitis Organisation

EDTA – ethylenediaminetetraacetic acid

EEN – exclusive enteral nutrition

EIM – extraintestinal manifestation

ELISA – enzyme-linked immunosorbent assay

EM – effector memory

EOMES – eomesodermin

ERK – extra-signal-regulated protein kinase

ESGAR – European Society of Gastrointestinal and Abdominal radiology

F

FACS – fluorescent-activated cell sorting

FC – faecal calprotectin

FCGR2A – FC gamma receptor IIa

FFPE – formalin-fixed paraffin-embedded

FGR – FGR proto-oncogene

FOXP3 – forkhead box P3

FSC-A – forward scatter area

FSC-H – forward scatter height

G

G-CSF – granulocyte stimulating factor

GALT – gut-associated lymphoid tissue

GATA3 – GATA binding protein 3

GBRC – Glasgow biomedical research centre

GI – gastrointestinal

GO – gene ontology

GRI – Glasgow Royal Infirmary

GSDMB – gasdermin-B

GSE – gene set enrichment

GSE – gene set enrichment

GWAS – Genome-wide association study

H

HBI – Harvey Bradshaw Index

HLA-DR – major histocompatibility complex II

HRP – horseradish peroxidase

I

IBD – inflammatory bowel disease

IBDU – IBD undetermined

IFN – Interferon
ICAM1 – intercellular adhesion molecule 1
IEC – intestinal epithelial cell
IEL – intraepithelial lymphocyte
Ig – immunoglobulin
IGHG – immunoglobulin heavy constant gamma
IGSF6 – Immunoglobulin superfamily member 6
IHC – immunohistochemistry
IL – interleukin
IL-12RB1 – IL-12 receptor beta subunit 1
IL-12RB2 – IL-12 receptor beta subunit 2
IL21R – IL-21 receptor
IL23P19 – IL23 subunit alpha
ITGAL – lymphocyte function-associated antigen 1
ITGA4 – integrin subunit alpha 4
ITGB2 – integrin subunit beta 2
ITGAL – lymphocyte function-associated antigen 1
ITGAM – macrophage-antigen 1
IRF – interferon regulatory factor
IRIS – Immune Response In Silico
ISH – in situ hybridisation

K

KEGG – Kyoto Encyclopedia of Genes and Genomes

L

L/D – viability dye
LAG3/CD223 – lymphocyte activating gene 3
LPMC – lamina propria mononuclear cell
LCK – Lymphocyte Cell-Specific Protein-Tyrosine Kinase
LN – Lymph nodes
LPS – lipopolysaccharide
LT – low peripheral blood T cells

M

M-CSF – colony stimulating factor 1
MADCAM-1 – mucosal vascular addressin cell adhesion molecule-1
MDP – myramyl dipeptide
MeSH – medical subject heading
mgsA – SMY methylglyoxal synthase
MHC – major histocompatibility complex
MLN – mesenteric LN
MNP – mononuclear phagocyte
MPO – myeloperoxidase
Msigdb – Molecular signature database
mTNF – membrane-bound TNF
MTOR – mammalian target of rapamycin
MTX – methotrexate
MyD88 – Myeloid differentiation primary response 88

N

NC – non-IBD controls
NCBI – national centre for biotechnology information
NF-kB – nuclear factor kappa B
NICE - UK National Institute for Health and Care Excellence
NK – Natural Killer

NLCR3 – NOD-like receptor family CARD domain containing 3
NLR – NOD-like receptor
NLRP3 – NOD-, LRR- and pyrin domain-containing protein 3
NOD2 – Nucleotide Binding Oligomerization Domain Containing 2
NOT – no treatment
NT – normal peripheral blood T cells

O

OOB – out-of-bag error
OSM – oncostatin-M
OSMR – oncostatin-M receptor
OX40L – OX40 ligand

P

p.adjust – adjusted p-value
PAMP – pathogen-associated molecular pattern
pANCA – perinuclear anti-neutrophil cytoplasmic antibody
PB – peripheral blood
PBMC – peripheral blood mononuclear cell
PBS – phosphate-buffered saline
PBST – phosphate-buffered saline with Tween®
PCA – principal component analysis
Peripheral blood – PB
PPIB – peptidylprolyl isomerase B (cyclophilin B)
PYCARD – Apoptosis-Associated Speck-Like Protein Containing A CARD

Q

qPCR – quantitative polymerase chain reaction

R

RA – rheumatoid arthritis
RNA – ribonucleic acid
RNAseq – RNA sequencing
ROC – Receiver Operating Characteristic
RORγt – retinoid-related orphan receptor gamma t
ROS – reactive oxygen species
RPLPO – large ribosomal protein
RT – room temperature
RTB – research tissue bank

S

scRNAseq – single cell RNA sequencing
SELPLG – Cutaneous Lymphocyte-Associated Associated Antigen
SES-CD – simple-endoscopy score for CD
SHIP – Inositol polyphosphate 5'-phosphatase–deficient
siRNA – small interfering RNA
SLE – systemic lupus erythematosus
sPLS-DA – sparse Least Square-Discriminant Analysis
SPP1 - osteopontin
SSC-A – sideward scatter area
STAT – Signal Transducer And Activator Of Transcription

T

T-bet – T box expressed in T cells
TBS – Tris-buffered saline
TBST – Tris-buffered saline with Tween®
TCR – T cell receptor
TGF-β – transforming growth factor β

Th – T helper

TIGIT – T-cell immunoglobulin and ITIM domain

TLR – toll-like receptor

TNFAIP6 – TNF alpha induced protein 6

TNFR – TNF receptor

TNFRSF1A – TNF receptor superfamily 1A

Treg – T regulatory cell

TREM-1 – triggering receptor of myeloid cells-1

Trm – tissue resident memory T cell

TSA – Tyramide signal amplification

U-Z

UC – ulcerative colitis

Vedo – Vedolizumab

WBC – white blood cell

YKL-40 – Chitinase-like 3 protein 1

ZAP70 – zeta chain of TCR-associated protein kinase 70

Chapter 1: General Introduction

1.1 Inflammatory bowel disease

Crohn's disease (CD) is one of the two main types of inflammatory bowel disease (IBD). IBD is an umbrella term describing a range of conditions that exhibit chronic gastrointestinal (GI) inflammation. The two most common forms of IBD are CD and ulcerative colitis (UC), which differ in disease location, behaviour, and inflammatory response. Inflammation in IBD, like most chronic inflammatory conditions, including rheumatoid arthritis (RA), follows a pattern of relapse and remission. During a relapse or flare, the patients experience a lot of discomfort, including abdominal pain, bloating, rectal bleeding, fatigue, nausea, and extreme weight loss (Hornecker and Veauthier, 2018). When left untreated, patients can experience life-threatening complications, including perforation of the intestinal wall, malnutrition, colonic cancer, bowel obstructions and toxic megacolon, a condition predominant in UC which leads to rapid swelling of the gut (MFMER, 2022).

The human intestine harbours around 100 trillion microbes (Ley, Peterson and Gordon, 2006), collectively called the gut microbiota, which usually generate beneficial metabolites from digested food and occupy niches to prevent the invasion of pathogens. A hallmark of IBD is an increase in intestinal permeability, also referred to as "leaky gut", allowing the translocation of microbiota which in turn triggers an immune response (Linares *et al.*, 2021). Bacteria or their DNA fragments have been quantified in lymph nodes (LNs) (O'Brien *et al.*, 2014) as well as peripheral blood (PB) of CD patients (Guti-acerrez *et al.*, 2009). Some species may dislocate into the mucosal layer during CD or adhere to the luminal epithelium. During active disease, the large number of microbiota serves as a continuous target for a complex immune response. IBD occurs in genetically susceptible individuals. More than 200 susceptibility genes (Jostins *et al.*, 2012; Liu *et al.*, 2015) have been identified so far, suggesting that complex interactions between genetic and environmental factors play important roles in IBD onset (*Figure 1.1*)

The first case of UC was reported in 1859 by Wilks (Wilks, 1859). Crohn *et al.* (1932) discovered CD 73 years later. IBD was only associated with chronic inflammation in 1959 after the identification of autoantibodies in the serum of diseased children (Broberger and Perlmann, 1959). Since the 20th century, a global surge in IBD cases has been observed. Initially, this increase was observed in western countries, including Europe, North America (Canada and the USA) and Oceania (New Zealand and Australia). However, since China reported its first IBD case in 1956, newly industrialised country are also exhibiting an increase in prevalence and incidence of the disease (Kaplan and Ng, 2017). The incidence of IBD in the UK has been found to have increased by 33.8% between 2006 to 2016 (Freeman *et al.*, 2021). Currently, at least 1 in 123 people are affected with IBD in the UK (Crohn's & Colitis UK, 2022).

1.2 Crohn's disease

1.2.1 CD phenotypes

CD affects any part of the GI tract and can appear in multiple locations. UC manifests as a continuous stretch of inflamed tissue being restricted to the colon and rectum. Most CD patients, approximately 50%, display inflamed tissue in both the ileum and colon, which is referred to as ileocolonic CD (Cheifetz, 2013; Feuerstein and Cheifetz, 2017). However, CD can also display solely as ileal or colonic disease. CD can cause inflammation in any part of the GI; inflammation in the perianal and upper regions can occur in a small proportion of patients (<10%).

Unlike UC, which only affects the mucosal surface, CD inflammation penetrates the mucosal layer causing a transmural inflammatory response (Panaccione, 2013). The behaviour of CD varies between patients and can change during disease progression. Patients can develop fistulas, abnormal connections between two organs, and stricturing due to fibrotic tissue formation and immune cell infiltration, preventing the passing of digestive matter (Satsangi *et al.*, 2006). Both complications cannot be reversed and require resective surgery. The most common phenotype in CD is strictly inflammatory behaviour without strictures or fistulating disease (Satsangi *et al.*, 2006). However, Louis *et al.* (2001) monitored

the disease progression of 125 CD patients over 10 years and found that 27.1% developed strictures and 29.4% fistulae, although they initially exhibited a solely inflammatory phenotype.

Commonly most patients are diagnosed between the ages of 17 and 40 (Quezada, Steinberger and Cross, 2013), but paediatric disease is more predominant year by year (Sýkora *et al.*, 2018). The age of diagnosis has been shown to be an indicator of disease severity. For example, patients that have been diagnosed above the age of 60 are more likely to have inflammation restricted to the colon (Crohn's colitis) (Saad *et al.*, 2016), which is usually less severe than ileal or ileocolonic disease. Additionally, paediatric disease onset (<17 years) has a higher proportion of complications than patients diagnosed later in life (Saad *et al.*, 2016). This could be due to a longer period of chronic inflammation or stronger genetic and/or environmental factors driving the disease.

Approximately 6-40% of IBD patients exhibit complications outside the intestine, referred to as extraintestinal manifestations (EIMs) (Rankin *et al.*, 1979; Bernstein *et al.*, 2001; Rogler *et al.*, 2021). The percentage of patients with EIMs varies massively between cohorts, again demonstrating the heterogeneity of IBD. EIMs can occur in many different locations and can, in fact, affect any organ. The most common ones (as defined by inflammatory scores defined in *section 1.3.1*) are inflammation in the eyes (uveitis), joints (arthralgia and arthritis), fat tissue (erythema nodosum) and skin (pyoderma gangraenosum) (Best *et al.*, 1976; Harvey and Bradshaw, 1980).

1.2.2 Genetic predisposition to CD

Genome-wide association studies (GWAS) have identified over 200 gene variants enriched in patients with IBD (Jostins *et al.*, 2012; Liu *et al.*, 2015). Some of these genes have been repeatedly connected with CD and provide information about its immunopathogenesis, which involves loss of epithelial barrier integrity, aberrant autophagy, and T cell responses. The caspase recruitment domain family member 15 (*CARD15*), also known as NOD2, was discovered as one of the first genetic risk factors and its polymorphisms have been identified as a marker to distinguish between CD and UC (Hampe *et al.*, 2001; Ogura *et al.*, 2001; Tsianos, Katsanos

and Tsianos, 2012). Its importance in CD pathology is demonstrated by its association with early onset of the disease and a more severe phenotype that often presents during relapse of steroid treatment (Posovszky *et al.*, 2013; Horowitz *et al.*, 2021). CARD15/NOD2 is a NOD-like receptor (NLR) and functions as a pattern recognition receptor (PRR) (further discussed in *section 1.4.7*) recognising bacterial peptides (Sidiq *et al.*, 2016). A *CARD15* mutation frequently found in CD patients reduced the expression of anti-microbial peptides (AMPs) *in vitro* (Voss *et al.*, 2005), which are crucial for maintaining the epithelial barrier. Another *in vitro* study in intestinal epithelial cell lines has demonstrated that the introduction of CD-associated *CARD15* polymorphisms prevents effective clearance of the bacterium *Salmonellum typhimurium* compared to the wildtype gene (Hisamatsu *et al.*, 2003). CARD15 recruits autophagy-related 16 like 1 (ATG16L1) protein to the membrane (Travassos *et al.*, 2010), which has been identified as another CD susceptibility gene and plays an important role in autophagy (Rioux *et al.*, 2007). During autophagy, cells degrade faulty cellular components in their cytoplasm. Together with the aberrant function of NOD2, dysfunctional autophagy has developed as a hallmark of CD immunopathology. Homer *et al.* (2010) demonstrated that kidney-derived epithelial cell lines with the CD-associated *ATG16L1* and *NOD2* variants failed to induce autophagy upon *Salmonella enterica* infection, which did not allow for bacterial clearance. These findings further supported the link between ATG16L1, NOD2, CD, and autophagy.

Toll-like receptors (TLRs), the second group of PRRs, are widely expressed on innate immune and epithelial cells and required for fast anti-bacterial responses (Takeda and Akira, 2005). Various gene variants of TLRs have been linked to CD, leading to aberrant activation, the most commonly affected one being the *TLR4* gene (Franchimont *et al.*, 2004; Oostenbrug *et al.*, 2005). Interestingly, CD-associated *CARD15* mutations impair negative TLR regulation, contributing to TLR hyperactivation. In turn, CARD15 signalling has been found to be induced by TLR activation, creating a feedback loop exacerbating inflammatory responses (Watanabe, Kitani and Strober, 2005).

Duerr *et al.* (2006) were the first to identify polymorphism in the IL-23 receptor (IL23R)-encoding gene as another major risk factor for developing CD. This receptor is crucial during the T helper (Th) 17 response (further discussed in

section 1.4.2.1), an important mechanism for defence against bacteria and, therefore, also crucial in keeping the intestinal microbiota from translocation into the intestinal tissue (Ahern *et al.*, 2008).

1.2.3 Environmental risk factors in CD

Even though many gene variants have been enriched in individuals with CD, most of them have low penetration within the patient cohorts (Rioux *et al.*, 2007). The most prominent age category for developing CD is between 17 and 40 years. However, studies have reported a surge in not only paediatric but also geriatric patients (>60 years of age) (Quezada, Steinberger and Cross, 2013). Early disease onset may relate to genetic predisposition. However, children in the western world exhibit a surge in many autoimmune and inflammatory diseases, such as diabetes type 2 (Rosenbloom *et al.*, 1999), which has been linked to environmental factors, including diet (Pulgaron and Delamater, 2014). Increasing disease onset late in life is another indicator that the environment and not necessarily genetics can cause CD.

The western lifestyle has been shown to decrease diversity within the intestinal microbiome when compared to individuals from rural and pre-industrial regions. Epidemiologically IBD cases started rising in newly industrialised countries when they started adopting the westernised diet and habits (Ng *et al.*, 2013; Ng, Tang and Leong, 2015). When affected with CD, individuals have less bacterial diversity in their faeces than those with a non-IBD phenotype (Olaisen *et al.*, 2021), specifically, they exhibit fewer *Firmicutes* and *Bacteroides* (Peterson *et al.*, 2008), which make up the majority of bacteria in a healthy gut (Qin *et al.*, 2010). Reduction in microbiome diversity is a phenomenon also observed in humans living in the USA compared to individuals from developing countries such as Malawi and Venezuela (Moeller *et al.*, 2014), supporting the hypothesis that the western diet contributes to CD development potentially *via* changes in the microbiome composition. CD can be managed using exclusive enteral nutrition (EEN), a liquid diet, which has proven to be very successful in treating children (de Bie, Kindermann and Escher, 2013). Svolos *et al.* (2018) have replicated the nutritional composition of EENs in a specific food-based diet plan called CD TREAT and found it effective in inducing remission in 80% of paediatric patients. The same study

demonstrated that intestinal microbial diversity is reduced during administration of EEN and CD TREAT, supporting the hypothesis that the microbiome contributes to inflammation in CD. Another dietary factor that has been linked to IBD is vitamin D. For example, high levels of vitamin D reduced the predisposition to CD in a cohort of 122 women (Ananthakrishnan *et al.*, 2012). Additional environmental risk factors, also linked to the western lifestyle, include smoking, air pollution, and high-stress levels (Berg, Dam and Farraye, 2013). A study comparing CD patients with different smoking statuses, continuing smokers (n =59) or quitters (n =59) and non-smokers (n =59) found that cessation of smoking reduced the risk of intestinal inflammation flare-ups (Cosnes *et al.*, 2001).

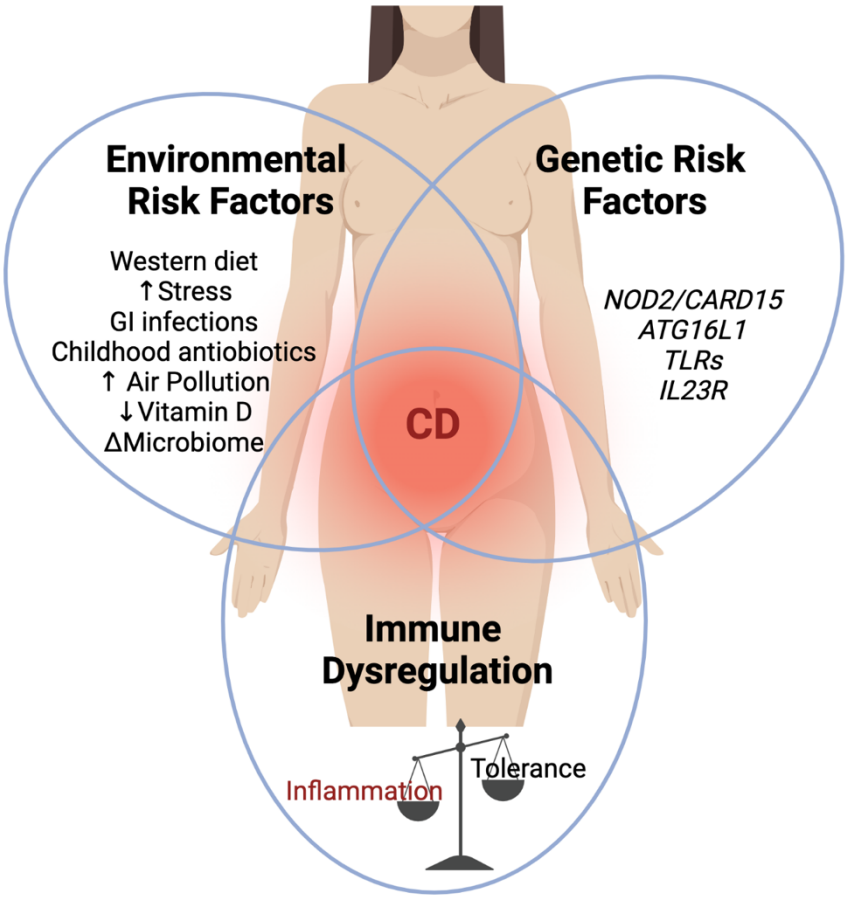


Figure 1.1: ***Overview of CD pathology.***

CD is a complex condition with many contributing factors, which can be classified into environmental and genetic risk factors and dysregulation of the immune system. Environmental factors include the nutritional value of the westernised diet, increased stress, air pollution, reduced exposure to vitamin D, infections affecting the GI tract, administration of childhood antibiotics and overall changes in the microbiome diversity and composition. >200 susceptibility genes have been identified as genetic risk factors, including polymorphisms in *NOD2/CARD15*, *ATG16L1*, TLR-encoding genes and *IL23R*. Another hallmark of CD immunopathology is the shift from tolerance to the commensal microbiome towards an aberrant anti-microbial pro-inflammatory response. This figure was created using BioRender.com.

Bacterial infections are another potential driver of CD disease onset and relapse. An indicator for this was the finding of DNA of pathogenic bacteria in lymph nodes (Lamps *et al.*, 2003) and the presence of anti-bacterial antibodies in patient serum (Tabaqchali, O'Donoghue and Bettelheim, 1978). Additionally, antibiotic treatments in child- and adulthood (Kronman *et al.*, 2012; Aniwan *et al.*, 2018) have been associated with CD, supporting that bacterial infections might promote disease and/or that changes in the intestinal microbiota contribute to CD onset. The bacterium *Clostridium difficile* has been linked with mortality of in-patient IBD patients, and therefore its infection might underlie severe disease progression (Ananthakrishnan, McGinley and Binion, 2008). In contrast to these findings, helminth infections during childhood have been allocated with potential protective capabilities against CD (Chu *et al.*, 2013).

1.2.4 Inflammation scores

The Crohn's Disease Activity Index (CDAI) is a questionnaire-based method to assess inflammation status in CD. It was developed by the National Cooperative Crohn's disease study group in 1976 in response to the need for a single index determining whether a patient was inflamed or not (Best *et al.*, 1976). They studied a cohort of 187 CD patients with different disease locations and selected the following eight variables as most predictive for inflammation in CD: the number of liquid stools, abdominal pain, general well-being, use of anti-diarrhoeal drugs in the last 7 days, abdominal mass, the proportion of red blood cells, body weight and complications (e.g., comorbidities, fissures, fistulae). The score range is 0-600, and the cut-off for active inflammation is >150. The Harvey Bradshaw Index (HBI) is a modification of the CDAI and was generated by Harvey and Bradshaw in 1980. It is a simplified version of the CDAI, which only scores general well-being, level of abdominal pain, the number of liquid stools, abdominal mass and complications. Even though it only covers five of the variables in the CDAI, Best (2006) demonstrated in a cohort of 224 that CDAI and HBI have a positive association with a correlation coefficient of 0.93. Overall, these well-being-based methods are very useful because they do not require invasive interventions, such as endoscopies, to estimate the level of inflammation in CD. However, a major limitation is a patient bias. Different patients may have different levels of pain

tolerance, may not remember the number of liquid stools over the past 24 hours correctly or may generally over/under exaggerate their symptoms.

The Crohn's Disease Endoscopic Index of Severity (CDEIS) was introduced in 1989 by the Groupe d'Etudes Thérapeutiques des Affections Inflammatoires du Tube Digestif (GETAID) (Mary and Modigliani, 1989). They recruited 75 CD patients and assigned pairs of clinicians to simultaneously assess the level of inflammation based on type, size and number of lesions and ulcerations to develop the index. Subsequently, an additional 103 patients with active CD (CDAI>150) were scored using the CDEIS, and the outcome compared to the general severity of lesions showed a significant positive correlation ($p<0.001$). Like the HBI, the Simple-Endoscopy Score for Crohn's Disease (SES-CD) is a simplified version of the CDEIS focussing on the extent, number, and size of ulcers, not taking the type of lesion into consideration (Daperno *et al.*, 2004). The score significantly correlated with CDEIS and CDAI in a 121 CD patient cohort.

1.2.5 Laboratory markers

Laboratory markers of inflammation in CD involve peripheral blood leukocyte counts and the measure of faecal calprotectin, a protein shed by specific immune cells during inflammation.

Increased C-reactive protein (CRP) levels indicate intestinal inflammation in CD (Solem *et al.*, 2005). CRP is a plasma protein produced by hepatocytes in response to pro-inflammatory stimuli such as interleukin (IL)-1ß and IL-6 cytokine secretion in response to trauma, inflammation or infection (Hurlimann, Thorbecke and Hochwald, 1966; Morley and Kushner, 1982). CRP is one of the main acute phase response proteins, which describes a state of a broad initial response of the immune system after its activation, and harbours anti-microbial functions (Richardson, Gray and Shankland, 1991). Even though CRP is routinely used in the clinical management of CD, it has been found that 63% of patients with inflamed ileocolonic disease did not display an increase in the protein in their blood (Solem *et al.*, 2005). A major limitation of CRP, and potentially also a contributing factor to its low sensitivity, is its upregulation during any kind of inflammatory response. For example, it has been recently shown that CRP also acts as a biomarker for severe influenza infection (Vasileva and Badawi, 2019). Infected patients with

severe disease had 1.6-fold higher CRP blood levels when compared to those with a less debilitating disease. Similarly, the average CRP levels in individuals with influenza-dependent hospitalisation and/or death were 140mg/L CRP, 50mg/L higher than patients classified as experiencing severe influenza.

Lymphocyte count (type of immune cells further discussed in *section 1.4.2.1*) is routinely used to monitor disease activity in CD. A decrease in lymphocyte count in PB has been frequently linked to active CD (Auer *et al.*, 1979; Heimann, Bolnick and Aufses, 1986; Neubauer, Woźniak-Stolarska and Krzystek-Korpacka, 2018; Kosoy *et al.*, 2021). However, another study did not detect any difference between remitting and inflamed patients (Xu *et al.*, 2019). Reports about changes in the total white blood cell (WBC) count of patients in response to inflammation have also been controversial. One study reported no differences between active and inactive disease (Liu *et al.*, 2012), another study demonstrated an increase in total WBC in inflamed CD patients (Xu *et al.*, 2019). An additional marker used to predict disease activity in CD is the neutrophil count (type of immune cell further discussed in *sections 1.5.3* and *1.5.4*), which is often used to calculate the PB neutrophil-to-lymphocyte ratio. When looking at CD, an increase in both neutrophil as well as the neutrophil-to-lymphocyte ratio has been linked with ongoing inflammation (Xu *et al.*, 2019; Langley *et al.*, 2021). When comparing PB markers of inflammation, CRP has been suggested to be the most effective predictive blood marker for inflammation in CD (Yazar *et al.*, 2020).

Faecal markers are an attractive target in CD since the main focus of damaging inflammation is the intestines. Calprotectin is a complex made up of S100A8 and S100A9 calcium zinc-binding proteins, which are mainly expressed by neutrophils and monocytes (see *section 1.5*)(Bjarnason, 2017). This heterodimer is released during inflammation to promote immune cell recruitment and induce cytokine stimulation. Faecal calprotectin has been found to be a reliable marker in paediatric as well as adult IBD (Bunn *et al.*, 2001; Gaya *et al.*, 2005; Penna *et al.*, 2020). Depending on the FC cut-off in stools selected to distinguish between non-inflamed and inflamed CD patients, this biomarker predicted activity with a sensitivity between 80-96% and a specificity of 67-78% (Gaya *et al.*, 2005; Penna *et al.*, 2020). In this instance, a higher cut-off contributed to better accuracy of FC in predicting intestinal inflammation. However, FC is not necessarily specific

to IBD-related intestinal inflammation and has also been shown to function as a predictive marker for acute bacterial gastroenteritis (Duman *et al.*, 2015).

A study by Penna *et al.* (2020) compared CRP and FC with the inflammation scores CDAI and SES-CD in a total of 80 endoscopies performed on 65 CD patients. They identified CRP and FC to be more sensitive in distinguishing endoscopic remission from active disease than the CDAI. Furthermore, FC was found to be the marker discriminating between different levels of inflammation with the highest sensitivity. Research by Schoepfer *et al.* (2010) found FC to be more accurate than CRP as well as lymphocyte count.

1.3 Organisation of the intestinal mucosa

The intestinal mucosa forms the barrier between the gut lumen and the rest of our body, which is essential to prevent pathogens as well as commensal microbiota from entering. The first layer of the mucosa is mainly composed of intestinal epithelial cells (IECs) with a ruffled luminal appearance, referred to as the crypts (Okumura and Takeda, 2017). In the small bowel, separated into duodenum, jejunum and ileum (from stomach to colon), the crypts are separated by villi structures. The villi of the colonic epithelium are less protruding, giving it a smoother appearance. To maintain the barrier, IECs have to be frequently replenished and/or replaced when old/damaged, which is facilitated by self-renewing stem cells located at the bottom of the crypts (Booth, O'Shea and Potten, 1999; He *et al.*, 2004). In order to protect the IECs, they are coated with a layer of mucus, which consists of mucin proteins produced predominantly by goblet cells (specific type of IEC)(Birchenough *et al.*, 2015). Another population of IECs are Paneth cells which are located in the stem cell area of the crypts and are important because of their anti-microbial activities (Ouellette, 2005), such as the production of AMPs. In addition to IECs, intraepithelial lymphocytes, an intestine-specific type of immune cell, are spread throughout the epithelial layer, which can induce quick inflammatory responses at pathogens that are trying to penetrate the luminal barrier (Hu and Edelblum, 2017).

Just below the thin layer of epithelium, the lamina propria begins, which is home to a complex immune system (further discussed in *sections 1.5* and *1.6*). This

involves innate immune cells, type of leukocytes that can induce fast immune responses, adaptive immune cells, which require priming and activation but carry out specific responses to pathogens, and secondary as well as tertiary lymphoid organs, referred to as gut-associated lymphoid tissues (GALTs). GALTs are dominated by B cells (an adaptive immune cell known for its antibody-producing characteristics) and play the main role in inducing immunoglobulin (Ig) A production and patrol specialised IECs called M cells (Murch, 2021). The IgA antibody is crucial for an effective anti-microbial response in the intestines (Gutzeit, Magri and Cerutti, 2014). M cells are a type of IEC that allows bacterial peptides and food protein to enter the lamina propria so that they can be sampled by immune cells for surveillance (Gebert, Rothkötter and Pabst, 1996). The intestine is drained by mesenteric lymph nodes (MLNs) that enable the immune interactions necessary to initiate and maintain the adaptive immune response (Mayer *et al.*, 2020).

1.4 Innate immunity in CD

The innate immune system forms the first line of defence against pathogenic infections and is characterised by fast immune responses triggered by invariant receptors. The cells highlighted in this section are relevant for this thesis, although many other innate immune cells have been identified and linked to CD.

1.4.1 What are macrophages?

Macrophages play a crucial role in maintaining intestinal tissue homeostasis (Bain and Schridde, 2018). They function as mononuclear phagocytes (MNPs), a type of immune cell that can engulf and lyse pathogens and can clear bacterial components without triggering major pro-inflammatory responses (Smythies *et al.*, 2005; Bujko *et al.*, 2018). In addition to their phagocytic capacities, intestinal macrophages have antigen-presenting cell (APC) characteristics, meaning they can sample, process, and carry pathogenic peptides to then present them to other immune cells using major histocompatibility complex (MHC) I or II (Platt and Mowat, 2008; Bain *et al.*, 2013, Domanska *et al.*, 2022). They are activated upon ligation of their PRRs by damage (DAMPs) or pathogen-associated molecular patterns (PAMPs) (Zhou *et al.*, 2016).

Based on their function during immune responses, macrophages have been subset into pro-inflammatory/classically activated and immunoregulatory/alternatively activated phenotypes, a paradigm coined by Charles Mills and colleagues (Mills *et al.*, 2000). Classical macrophages, also known as M1 macrophages, are polarised by cytokine signals, such as granulocyte-macrophage colony-stimulating factor (GM-CSF), tumour necrosis factor-alpha (TNF-α) and interferon-gamma (IFN-γ) (Nathan *et al.*, 1983; Albina *et al.*, 1990). Alternative macrophages, also known as M2 macrophages, undergo polarisation in response to IL-4 and IL-13 (Mills *et al.*, 1992; Stein *et al.*, 1992; Doyle *et al.*, 1994). In terms of cytokines secretion, classical macrophages produce pro-inflammatory cytokines such as interleukin (IL)-1ß, IL-6, IL-12, IL-23 and TNF-α. Alternative macrophages function as their counterpart and express the immunoregulatory cytokines IL-10 and transforming growth factor (TGF)-ß. Based on their cytokine profiles, they are linked to T helper (Th1) 1 and Th2 immune responses, which are further discussed in *section 1.6*.

This paradigm has since evolved, and further macrophage subsets have been identified. Mantovani *et al*. (2004) proposed distinguishing M2 macrophages into three distinct phenotypes based on their activation stimuli and effector functions. M2a macrophages relate to the classic M2/alternative macrophage phenotype and are polarised by IL-4 and IL-13. They have type 2 effector functions, which involve immune responses against helminth infections and inflammation in allergy. M2b macrophages differentiate upon stimulation with a TLR ligand and an immune complex (antigen:antibody complex). They release pro-inflammatory cytokines (such as IL-1, TNF, and IL-6) but also high levels of IL-10. M2c macrophages are induced by IL-10 and glucocorticoids (used as anti-inflammatory treatments) and are involved in tissue remodelling and immunoregulation. Overall, Mantovani *et al*. (2004) described diverse phenotypes within the M2/alternative macrophage subpopulation and highlighted the importance of potentially stepping away from the M1/M2 paradigm.

Additionally, even though classically and alternatively macrophages have distinct functions, their phenotypes are plastic. For example, sequential treatment of murine bone marrow-derived and peritoneal macrophages with different cytokines *in vitro* has shown their ability to adapt to their cytokine environment (Stout *et al.*, 2005). IL-10 treatment post-IL-4 and pre-LPS stimulation increased IL-10

production and dampened TNF-α and IFN-γ production. This effect was reversed by IL-4 or IL-12 stimulation prior to LPS treatment. Gratchev *et al.* (2006) isolated human blood monocytes and differentiated them in Th1 or Th2 conditions with IFN-γ or IL-4 or IL-4 and dexamethasone (a glucocorticoid hormone). Interestingly, IFN-γ polarised macrophages responded to IL-4 and IL-10 similarly to those primed with IL-4/IL-4+ dexamethasone, suggesting that M1/classical macrophages can adapt an M2/alternative phenotype. This macrophage function is especially important in the intestine due to the constant exposure to commensals and food antigen.

1.4.2 What are monocytes?

Monocyte differentiation is driven by a selection of stimuli. After egress into the intestinal tissue, human monocytes have high expression of CD14 and CD11c and the calprotectin complex proteins (S100A8/9) and have strong pro-inflammatory characteristics. Stimulation with the colony stimulation factor 1 (CSF1/M-CSF) induces their differentiation and potentially also promotes cell survival (MacDonald *et al.*, 2010). Independent of CSF-1 exposure, monocytes start upregulating MHC II protein HLA-DR (Bernardo *et al.*, 2018), essential for antigen processing and presentation. In humans, the regulatory cytokines IL-10 and transforming growth factor (TGF)-β have been found to be the main mediators for the phenotypic change from immature/pro-inflammatory monocyte to a mature intestinal macrophage with the CD64 positive, HLA-DR^{high} and CD14^{low} expression pattern (Smythies *et al.*, 2010; Maheshwari *et al.*, 2011).

1.4.3 Role of macrophages and monocytes in intestinal homeostasis and infection

Intestinal macrophages are seeded during embryonic development or differentiate from infiltrating blood monocytes later in life. Using techniques such as fate mapping in mice, where cell types are fluorescently labelled and followed during embryonic development, researchers have provided evidence that intestinal macrophages can be generated and deposited in early developmental stages (Yona *et al.*, 2013; Sheng, Ruedl and Karjalainen, 2015). Blood monocytes are mononuclear cells also from the myeloid lineage and patrol the circulation as a

crucial first line of defence. Even though intestinal macrophages derived from embryonic precursors are thought to be able to self-renew (Hashimoto *et al.*, 2013), this cycle is limited in terms of speed and progenitor resources (Bain *et al.*, 2014). Large numbers of monocytes enter the intestines from the bloodstream and undergo differentiation to become a mature functional intestinal macrophage. Knockout of the chemokine receptor type 2 (CCR2) in mice has been shown to impact intestinal macrophage levels (Bain *et al.*, 2013), showing that monocytes are indeed homing to the gut for differentiation into macrophages and linking monocyte recruitment to CCR2:CCL2 (chemokine ligand 2) axis.

It has been difficult to distinguish intestinal macrophages from phenotypically similar dendritic cells (DCs). The expression of Fc gamma receptor 1 (CD64) has been identified as a marker for macrophages in both humans and mice (De Calisto, Villablanca and Mora, 2012). Intestinal macrophages consistently strengthen the barrier by clearing apoptotic epithelial cells and remodelling the luminal wall (Cummings *et al.*, 2016). Using their APC and phagocytic capabilities hand in hand, intestinal macrophages silently remove bacteria that adhered to the epithelium or translocated into the mucosa and induce local antigen-specific T cell responses (Smythies *et al.*, 2005). This allows for the immune system to react quickly to penetrating bacteria. In contrast to this pro-inflammatory response, alternative macrophages release IL-10, which is required for immunoregulation and repair mechanisms. In mice, IL-10 derived from intestinal macrophages has been shown to protect from mortality upon *Citrobacter rodentium*, a type of bacteria infecting the intestines, by reducing IL-23A secretion, a cytokine driving anti-bacterial immune responses (Krause *et al.*, 2015). Another mouse study showed that macrophage-derived IL-10 is required to heal the small intestine after chemically-induced damage (Morhardt *et al.*, 2019).

IL-10 production of intestinal macrophages has been related to transforming growth factor-beta (TGF-ß) signaling. The use of floxed TGF-ß receptor expression in integrin alpha subunit X $(CD11c)^{+}$ cells, a common marker for monocytes/macrophages, showed that intestinal macrophages in Cre^{+} mice had reduced levels of IL-10 expression compared to Cre^{-} mice (Schridde *et al.*, 2017). In addition to the cytokine environment, the microbiota has also been found to be required for shaping intestinal macrophage phenotype in homeostasis (Sorini *et*

al., 2018). Flow cytometry analysis showed a decrease in colonic macrophages in germ free mice (mice lacking microbiota) compared to those with commensals. However, co-housing of the two mice strains restored colonic macrophage levels in the germ free phenotype.

1.4.4 Role of macrophages and monocytes in CD

The importance of macrophage-derived IL-10 in the maintenance of intestinal homeostasis is supported by its role in IBD pathology. Patients with IBD can exhibit mutations in IL-10 receptor encoding genes (Glocker *et al.*, 2009; Shim and Seo, 2014; Khoshnevisan *et al.*, 2019). Macrophages generated *in vitro* from PB monocytes from patients with IL-10 polymorphisms generate higher levels of pro-inflammatory cytokines in response to TLR stimulation than control individuals (Shouval *et al.*, 2014). This suggests that a dysfunction in the IL-10-pathway in monocytes-derived macrophages in CD patients may contribute to aberrant inflammation. Additionally, CD patients have exhibited high levels of $CD14^{high}$ cells, a marker for immature monocytes, with the capacity to produce the pro-inflammatory cytokines such as IL-23, IL-6 and TNF-α in their lamina propria (Kamada *et al.*, 2008). Bernardo *et al.* (2018) identified two subsets of $CD64^{+}CD14^{+}HLA\text{-}DR^{+}$ macrophages in the human intestine defined by the expression of CD11c and CCR2, the latter being the receptor to the monocyte chemoattractant CCL2. $CD11c^{high}CCR2^{high}$ macrophages secreted higher levels of IL-1β and lower levels of IL-10 than $CD11c^{dim}CCR2^{dim}$ macrophages in response to LPS stimulation. Interestingly, patients with active CD showed a significant increase in $CD11c^{high}CCR2^{high}$ macrophages in inflamed tissue compared to quiescent tissue, patients with remitting disease and non-IBD controls. Overall, these studies suggest that CD patients exhibit infiltration and expansion of pro-inflammatory macrophages, facilitated by the CCR2:CCL2 axis.

An additional hallmark of CD macrophages is their hyperresponsiveness to TLR stimulation, which is a characteristic displayed by infiltrating blood monocytes prior to maturation in the intestine (Maheshwari *et al.*, 2011). This indicates that immature, highly pro-inflammatory macrophages fail to mature and contribute to CD pathology. In addition to their aberrant activation, macrophages and monocytes have been suggested to also have impaired bacteria-clearing capacities

in CD. Vazeille *et al.* (2015) isolated monocytes from PB of 24 CD patients and 12 healthy controls and cultured them with pathogenic and non-pathogenic *Escherichia coli (E. coli)* strains to understand their potential to phagocytose and subsequently lyse the bacteria. Prolonged exposure for 1hr resulted in a significantly higher pathogenic bacterial load in CD patient cells when compared to the non-IBD samples. No difference in bacterial load was detected between the sample groups after infection with the non-pathogenic *E. coli* strain, suggesting that blood-derived monocytes from CD patients have impaired lysogenic capabilities.

1.4.5 What are neutrophils?

Neutrophils are polymorphonuclear leukocytes characterised by their distinct segmented nuclear shape and granule expression in the cytoplasm. They derive from the granulocyte-monocyte progenitor and are generated and matured in the bone marrow (BM) (Vietinghoff and Ley, 2008). Their differentiation is driven by the granulocyte colony-stimulation factor (G-CSF). The BM stroma expresses high levels of the chemokine receptor C-X-C Motif Chemokine Ligand 12 (CXCL12), which binds C-X-C Motif Chemokine Receptor 4 (CXCR4) on neutrophils, preventing them from exiting. High levels of G-CSF have been shown to interfere with the CXCL12:CXCR4 axis, allowing for neutrophils to be released into the circulation (Hyun *et al.*, 2006). Granules of neutrophils store high levels of AMPs and reactive oxygen species (ROS), which are released upon pathogen encounter (West *et al.*, 1974). These AMPs harm microbes in different ways, disrupting their membrane structure, biosynthesis or differentiation (Huan *et al.*, 2020). Additionally, the release of ROS, a reactive group of molecules, induces DNA and/or RNA damage by oxidisation (Nathan and Shiloh, 2000).

Neutrophils are highly abundant in PB (50-70% of PB leukocytes in humans) and function as the first line of defence during infections (Kolaczkowska and Kubes, 2013). The segmented morphology of their nucleus allows for them to migrate quickly because they can easily squeeze through tissues to reach the site of infection or damage. In intestinal homeostasis, neutrophils like macrophages are crucial to fight translocated bacteria. They respond to chemokine signals released by macrophages that were previously triggered by a danger signal (pathogen or

damage). (Knall, Worthen and Johnson, 1997; Tester *et al.*, 2007; Yamamoto *et al.*, 2008). Once at the site of inflammation, neutrophils degranulate, releasing proteases with diverse functions, including inducing the production of monocyte-recruiting factors by macrophages and endothelial cells (Chertov *et al.*, 1997).

1.4.6 Role of neutrophils in CD

Human studies about the role of neutrophils in CD are rare. However, neutrophil infiltration is a common feature in UC and is used as a measure in scoring for disease activity (Bressenot *et al.*, 2015). Marks *et al.* (2006) tested the potential of neutrophil recruitment in CD by sampling tissue from the same intestinal site 6h apart in 6 patients with quiescent CD as well as 9 non-IBD controls and 3 UC patients. After inducing trauma by extracting a pinch biopsy on the sampling site, it was evaluated how many neutrophils accumulated during the 6h window. Immunohistochemistry (IHC) staining against myeloperoxidase (MPO), a marker for activated neutrophils (Sarr *et al.*, 2021), and IL-8 were used to identify neutrophils and neutrophil-recruiting cells, respectively. Rectal and ileal tissue derived from CD patients showed significantly fewer MPO^+ cells post-trauma than both tissue types of non-IBD controls and rectal tissue samples from UC patients, even though the baseline numbers were similar. Additionally, they also displayed fewer IL-8 positive cells when compared to controls in both rectum and ileum post-trauma. However, *in vitro* migration assays of human neutrophils derived from CD patients have shown that neutrophils display normal migratory behaviour (Morain *et al.*, 1981; Wandall, 1985). It is currently unknown which CD-induced factors prevent neutrophil infiltration during disease. G-CSF as a potential CD treatment has been tested in a selection of patients, initially showing promising results (Dejaco *et al.*, 2003; Korzenik and Dieckgraefe, 2005). These findings suggest a defective neutrophil phenotype in CD and a normal function in UC.

In addition to their antimicrobial function in intestinal homeostasis and inflammation, neutrophils have been suggested to recruit Th17 cells (further discussed in *section 1.6.3*) and promote their differentiation. Immunofluorescent staining showed colocalization of the transcription factor (TF) retinoid-related orphan receptor gamma t (RORγt), highly expressed on the Th17 subset, and neutrophils in inflamed intestinal tissue of CD patients (Pelletier *et al.*, 2010).

Neutrophils that infiltrated the mucosa of IBD patients have also been found to express high levels of IL-23, a cytokine driving Th17 cell expansion (Kvedaraite *et al.*, 2016). In healthy individuals and during homeostasis, this interaction may contribute to vital anti-microbial Th17 responses, but in CD, even though neutrophil infiltration has been found to be impaired, it may be contributing to detrimental chronic inflammation.

1.4.7 Pattern Recognition Receptors in CD

PRRs play an important role in pathogen sensing. GWAS have predominantly linked two types of PRRs, TLRs and NLRs, to CD pathogenesis. They recognise pathogen PAMPs and DAMPs and induce fast pro-inflammatory responses.

1.4.7.1 What are toll-like receptors?

TLRs can be located on the cell surface or within the cytoplasm (Takeda and Akira, 2005). They respond to bacterial and viral antigens and trigger the nuclear factor-kB (NF-kB) cascade and/or interferon regulatory factors (IRFs). NF-kB is a transcription factor that is held in the cytoplasm by inhibitory proteins during a quiescent state (T. Liu *et al.*, 2017). Activation of the NF-kB signalling cascade involves the cleavage of these inhibitory proteins and subsequent translocation of the functional transcription factor to the nucleus. IRFs are phosphorylated downstream of TLR activation in the cytoplasm, causing conformational changes in form of dimerization required for nuclear translocation of the NF-kB transcription factor (Mogensen, 2019). TNF-α and IL-6, pro-inflammatory cytokines which are highly expressed during active CD, are products of the NF-kB cascade (Libermann and Baltimore, 1990; Schütze *et al.*, 1995).

1.4.7.2 Role of toll-like receptors in CD

The relationship between TLR signalling and CD is complex. Activation of these receptors has been found to significantly contribute to chronic intestinal disease as well as have preventative functions. As discussed in *section 1.2.2*, polymorphisms in TLRs are common genetic risk factors for CD, demonstrating that their aberrant function contributes to disease. A study by Hausmann *et al.* (2002) has shown by IHC staining that TLR4 is upregulated in intestinal tissue of CD

patients with active disease when compared to non-IBD controls and remission samples. The staining was concentrated around intestinal crypts, an area dominated by epithelial cells as well as macrophages. Similar observations have been made about TLR2 (Hausmann *et al.*, 2002); however, these are not consistent throughout the literature, and another study has reported no change in TLR2 expression between healthy and diseased tissue (Cario and Podolsky, 2000). The levels of the anti-viral TLR3 are reduced in CD patients, a potential response of the innate immune system to the less diverse microbiome (Cario and Podolsky, 2000).

In contrast to their CD-inducing characteristics, TLRs have also been found to have a protective function in mouse models of intestinal inflammation. Mouse models with chemically induced colitis in their intestines show severe outcomes upon TLR2, TLR4 as well as MyD88 gene knockout when compared to mice with functional TLR signaling (Araki *et al.*, 2005; Fukata *et al.*, 2005; Cario, Gerken and Podolsky, 2007). The stimulation of TLR2 in the same colitis mouse model also exhibited a protective effect (Cario, Gerken and Podolsky, 2007). Moreover, mice lacking TLR4 and MyD88 exhibited higher amounts of pathogenic bacterial DNA in their MLNs when compared to wild-type mice and translocation of the bacteria into the luminal epithelium which was not detected in the control group (Fukata *et al.*, 2005). These findings indicate the complex role of TLRs in intestinal disease since they have both a pro-inflammatory but also protective role depending, suggesting TLR hypersensitivity as a hallmark of IBD.

1.4.7.3 What are NOD-like receptors?

NLRs are located intracellularly and, similar to TLRs, can trigger the NF-kB signaling cascade but also induce inflammasome activation. The inflammasome complex forms in the cytosol and enables the activation of the enzyme caspase-1 by cleavage of its initial non-functional form. Caspase-1 has proteolytic characteristics and cleaves pro-IL-1ß and pro-IL-18, two cytokines both crucial for pro-inflammatory immune responses, into their active forms.

1.4.7.4 Role of NOD-like receptors in CD

As mentioned in *section 1.2.2,* variants of the gene encoding the NLR CARD15 are prominent in CD patients. CARD15 recognises muramyl dipeptide (MDP) found in the cell wall of most bacteria (Girardin *et al.*, 2003). Excessive MDP stimulation of human blood-derived monocytes and intestinal macrophages showed that the pre-exposure to the bacterial peptide dampen TLR2 and 4 signalling *in vitro* (Hedl *et al.*, 2007). The same effect was seen in blood-derived monocytes of CD patients carrying a disease-related *CARD15* polymorphism without MDP pre-treatment, suggesting aberrant CARD15/TLR interaction in these individuals. Therefore, loss of function mutations in *CARD15*/*NOD2* potentially contribute to the hyper-activation of TLR during CD. Another link between NLR and TLR signalling can be found during inflammasome responses. The protein complex required for inflammasome activation includes three main components (a sensor, an adaptor, and an effector), which can vary in different inflammasome phenotypes (Swanson, Deng and Ting, 2019). The NOD-, LRR- and pyrin domain-containing protein 3 (NLRP3)-dependent inflammasome has been found to be an important driver of inflammation in IBD. Bauer *et al.* (2010) demonstrated that knockout of the NLRP3 gene in a murine model of colitis induced protection against intestinal inflammation. Moreover, *NLRP3* expression was increased in CD patients as well as autoimmune-prone mice (L. Liu *et al.*, 2017). The microbial sensor of this inflammasome, NLRP3, is produced downstream of NF-kB activation, which is driven by NLRs but also TLR activation. NLRP3 then binds its adaptor protein PYCARD (Apoptosis-Associated Speck-Like Protein Containing A CARD) and effector protein caspase-1, and the inflammasome response generates active IL-1ß and IL-18 (Swanson, Deng and Ting, 2019). IL-1 ß secretion in particular has been linked to CD inflammation. For example, a study using colonic explant cultures (cultures of pinch biopsies in media for a 24h period) derived from 20 CD patients and 13 non-IBD controls demonstrated a significant increase of IL-1ß in diseased tissue (Ludwiczek *et al.*, 2004). Additionally, Reinecker *et al.* (1993) demonstrated that lamina propria mononuclear cells (LPMCs) isolated form colonic pinch biopsies of IBD patients with inflamed tissue (n_{UC}= 12, n_{CD}= 12), secreted higher levels of IL-1ß than tissue from patients with disease in remission (n_{UC}= 17, n_{CD}= 26). Neutralisation of LMPC macrophages using an anti-CD68 antibody significantly

reduced IL-1β levels, suggesting them as one of the main producers of the cytokine in the intestine.

1.5 Adaptive immunity in CD

The innate immune system is vital in maintaining intestinal tissue homeostasis. Adaptive immunity is activated by the innate immune system to help clear infectious agents. Macrophages and DCs are both APCs and can present antigens to adaptive immune cells. In intestinal homeostasis, the production of IL-10 by macrophages and DCs have been found to induce immune regulatory adaptive immune cells. However, during infection or disease, they facilitate the activation of antigen-specific effector cells and immune responses by releasing pro-inflammatory cytokines. The two main effector cell phenotypes facilitating adaptive immune responses are B and T cells.

1.5.1 What are T cells?

T cells are lymphocytes that derive from the early thymic progenitor (Donskoy and Goldschneider, 1992) and mature in the thymus once they leave the bone marrow (Anderson *et al.*, 1993). During their thymic development, T cells undergo a distinct succession of maturation steps to ensure they are released into the periphery as functional non-self-reactive immune cells (Spits, 2002; Kumar, Connors and Farber, 2018). After maturation, the main T cell subsets are characterised by the type of their antigen-specific T cell receptor (TCR) and its co-receptors. The TCR can be made up of $\alpha\beta$ or $\gamma\delta$ chains (Geisler, Larsen and Plesner, 1988), the latter being highly abundant amongst IELs (Deusch *et al.*, 1991) but in general less frequently than their $\alpha\beta$ counterparts (less than 5% of human T cells) (Vroom *et al.*, 1991). Both types of TCR require two CD3 heterodimers (CD3ε/CD3γ, CD3ε/CD3δ) responsible for the intracellular signaling (Dadi, Simon and Roifman, 2003). $\alpha\beta$ T cells also express one of two types of co-receptors essential for their activation, which are the glycoproteins CD4 and CD8 (Janeway, 1992).

1.5.2 CD4$^+$ T cell activation

The TCR binds to antigenic peptides bound to the major histocompatibility complex (MHC) protein expressed abundantly on the surface of APCs. T cells circulate in the periphery, constantly entering and exiting the secondary lymphoid organs waiting to encounter an APC that carries an antigen matching the specificity of their TCR. DCs migrate from the site of inflammation or infection carrying antigen to the closest draining LN (Tai *et al.*, 2018). Once they meet the cognate T cell, the activation process begins. Step one of T cell activation is the binding of the TCR to antigen peptide carried by the MHC on the APC, and the co-receptor CD4 or CD8 stabilises this MHC:TCR complex. Step two requires the interaction of co-stimulatory surface proteins and their receptors on APCs and T cells, which are upregulated in response to pathogen encounter and TCR ligation, respectively. These molecules include CD28 and OX40 on T cells as well as their ligands CD80 and CD86 (binding CD28) and OX40 ligand (OX40L) (Rudd, Taylor and Schneider, 2009). The third and final signal requires the secretion of cytokines, which is essential to maintain T cell survival, and drive their differentiation and proliferation (Coffman, 1998). After this primary activation, activated CD4$^+$ T cell clones eventually generate expanded antigen-specific memory cohorts, which allow for faster specific responses upon re-exposure to their specific antigen (Swain, 1994; Yoshimoto *et al.*, 1996).

1.5.3 CD4$^+$ T helper subset differentiation

CD4$^+$ T cells, also referred to as Th cells, are a very plastic immune cell phenotype (Ivanova and Orekhov, 2015). Their TCRs recognise antigens carried by the MHC class II complex on APCs, and subsequently, different cytokine stimuli can induce differentiation into distinct Th subsets. Prior to activation, they are referred to as Th0 cells. Over the years, many different Th subpopulations have been identified. There is controversy amongst T cell biologists about whether all subsets are indeed final stages of Th differentiation or whether some are transient phases that exist due to high plasticity amongst the subpopulations (Ivanova and Orekhov, 2015). The main Th subsets involved in IBD are Th1, Th2 and Th17 (*Figure 1.2*).

Initially, it was found that stimulation of Th0 cells with IL-12 and IL-27 during activation phosphorylates Signal Transducer And Activator Of Transcription (STAT) 4 signaling downstream of the IL-12R promotes Th1 differentiation (Nishikomori *et al.*, 2002; Pflanz *et al.*, 2002). However, Szabo *et al.* (2000) discovered the transcription factor T box expressed in T cells (T-bet), which was solely expressed in T cells under Th1 polarising condition and is now used as one of their main markers. In addition to IL-12 and IL-27, phosphorylation of STAT1 by IFN-γ also drives Th1 differentiation (Ma, Huang and Huang, 2010). Th1 cells can be identified by their cytokine expression profile which is characterised by IFN-γ, TNF-α and IL-2. Production of IFN-γ by Th1 cells forms an autocrine stimulus, promoting and maintaining the lineage. Th1 cells are part of the type 1 immune response, important for clearing intracellular pathogens. This is reflected by the fact that individuals with a deficiency in T-bet develop mycobacterial infections, the type of bacterium causing tuberculosis (Yang *et al.*, 2020).

The cytokine IL-4 induces Th2 cell differentiation by STAT6-dependent GATA binding protein 3 (GATA3) upregulation (Kotanides and Reich, 1996; Zheng and Flavell, 1997). A downstream effect of GATA3 is enhanced accessibility of the IL-4 encoding gene, inducing a similar autocrine loop as seen in Th1 cells. Apart from IL-4 Th2 cells produce IL-5 and IL-13 (McKenzie *et al.*, 1998) and are known to drive anti-helminth responses. For example, mice deficient in IL-13 failed to effectively clear the infection of the gastrointestinal parasite *Nippostrongylus brasiliensis* (McKenzie *et al.*, 1998). Aberrant Th2 responses have been linked to allergies and asthma. Anti-IL-5 treatment is currently trialled for patients affected by the latter disease (Bagnasco *et al.*, 2018).

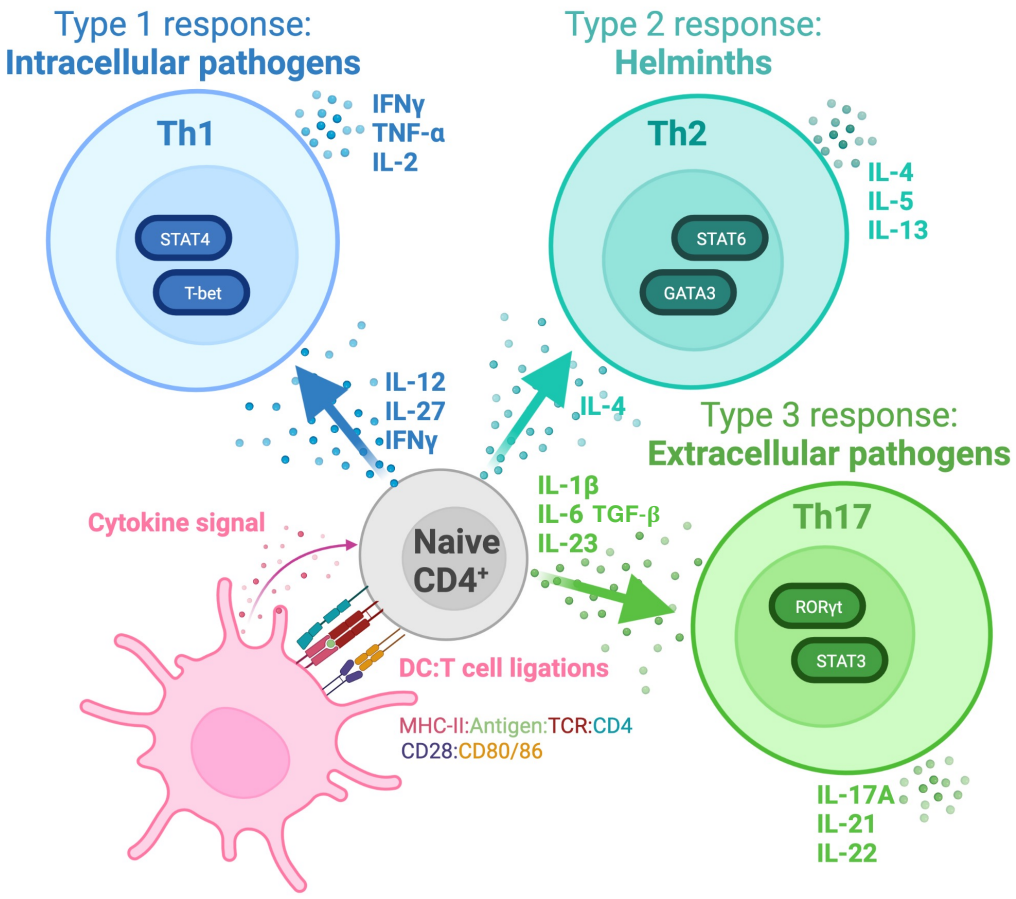


Figure 1.2: ***CD4⁺ T cell activation and polarisation of Th1, Th2 and Th17 subsets.***

The cytokine environment during the activation of naïve CD4⁺ T cells differentiates them into distinct phenotypes. Th1, Th2 and Th17 have been identified to drive IBD pathogenesis. Initial activation of T cells involves ligation of their TCR and CD4-co receptor with the MHC II:antigen complex on the DC (or other APC). Subsequently, CD28 on the APC binds with CD80/CD86 on the T cell surface to stabilise the complex. Th1 cells are polarised by IL-12, IL-27 and IFN-γ and have been characterised by an increase in STAT4, T-bet and production of IFN-γ, TNF-α and IL-2. Th2 cell differentiation is driven by IL-4 stimulation, followed by an upregulated of STAT6 and GATA3 and secretion of IL-4, IL-5 and IL-13. Th17 cells are induced by TGF-β, IL-1β, IL-6 and IL-23-dependent upregulation of RORγt and STAT3 and release of IL-17, IL-21 and IL-22. This figure was created using BioRender.com.

IL1-β, IL-6 and IL-23 differentiate Th0 cells into the Th17 phenotype via STAT3 signaling (Aggarwal *et al.*, 2003; Acosta-Rodriguez *et al.*, 2007; Wang *et al.*, 2013). The characteristic transcription factor upregulated in Th17 cells is the TF RORγt (Yang *et al.*, 2007) and their cytokine secretion profile involves IL-17A/F and IL-21, IL-22 (Liang *et al.*, 2006; Wei *et al.*, 2007). The upregulation of RORγt can also be promoted by TGF-β, which induces the degradation of the SKI protein bound to the transcription factor SMAD4, a state that suppresses RORγt transcription (Zhang *et al.*, 2017). Th17 cells are part of type 3 immune responses, which protect against extracellular pathogens and fungi. The deficiency of STAT3, and therefore

Th17 differentiation, makes humans susceptible to chronic mucocutaneous candidiasis, the chronic mucosal infection by the opportunistic fungus *Candida Albicans* (Sowerwine, Holland and Freeman, 2012). Similar findings supporting the importance of Th17 in anti-fungal defence were found in mice lacking the IL-17A receptor, exhibiting enhanced systemic *Candida Albicans* infection (Huang *et al.*, 2004). Y. Wang *et al.* (2019) extracted LMPCs from mice and co-cultured them with anti-CD3 antibodies and splenic APCs. Stimulating the LPMCs with bacteria-derived flagellins generated IL-17A, indicating that Th17 cells proliferated in response to the extracellular bacterial stimulus.

1.5.4 Regulatory T cell Activation and Differentiation

In addition to their pro-inflammatory effector functions, $CD4^+$ T cells can also exhibit immunoregulatory phenotypes in the form of T regulatory T cells (Tregs). Tregs are characterised by the expression of the transcription factor forkhead box P3 (FOXP3) and IL-2 receptor alpha chain (CD25), as well as the production of the anti-inflammatory cytokine IL-10 (Fontenot, Gavin and Rudensky, 2003). Thymic Tregs were identified as $CD25^+CD4^+CD8^-$ thymocytes by Itoh *et al.* (1999). *In vitro* stimulation showed that these cells were unresponsive to antigen (anergic). Furthermore, depletion of these cells in thymocyte suspension and adoptive subsequent transfer into mice lacking a functional thymus caused autoimmunity affecting many organs, including the intestines. In addition to those derived from the thymus (referred to as natural Tregs), Tregs can also differentiate in the periphery from naïve $CD4^+$ T cells (induced Tregs). The induced phenotype expresses the characteristic Treg markers but lacks the expression of the TF Helios, which is upregulated in $CD25^+FOXP3^+CD4^+$ thymocytes but not splenocytes during murine development (Thornton *et al.*, 2010). Induced Tregs differentiate in response to IL-2 and TGF-β. The main producers of TGF-β are DCs, which secrete the cytokine in response to stimulation of commensal bacteria (Kashiwagi *et al.*, 2015). IL-2 and TGF-β induce demethylation of enhancer sequences on the *Foxp3* gene locus of T cells, which allows for transcription of the TF (Ogawa *et al.*, 2014). Lack of functional FOXP3, and in consequence functional Tregs, in mice causes uncontrolled proliferation of lymphocytes and severe inflammation (Brunkow *et al.*, 2001). Similarly, mutations in the *FOXP3* gene have been linked to the immune dysregulation, poliendocrinopathy, enteropathy, X-linked syndrome, which

exhibits fatal systemic autoimmunity in humans (Bennett *et al.*, 2001). Additionally, the same study showed that family members of affected individuals in 2 out of 3 tested families showed mutations in *FOXP3*.

1.5.5 CD8$^+$ T cell Activation and Differentiation

CD8$^+$ T cells mostly act as part of type 1 immune responses and are crucial for anti-viral immunity. Different to CD4$^+$ T cells, they recognise intracellular peptides presented on MHC I, and not MHC II, complex (Townsend, Gotch and Davey, 1985). MHC I is expressed on all mammalian cells carrying a nucleus and is therefore useful for immune surveillance e.g., for the detection of intracellular viruses (Townsend, Gotch and Davey, 1985; Hewitt, 2003). Upon activation, CD8$^+$ T cells differentiate into cytotoxic T lymphocytes (CTLs), which, as their name suggests, exhibit cytotoxic characteristics that release lysogenic proteins during their degranulation (Peters *et al.*, 1991). Cytokines promoting CTL activation are the Th1 cytokines IL-2, TNF-α and IFN-γ (Leist *et al.*, 1989; Gately *et al.*, 1994; Pardoux *et al.*, 1997; Kasahara *et al.*, 2003). Besides granule production, they also secrete IFN-γ and TNF-α (Ratner and Clark, 1993; Gately *et al.*, 1994). The importance of anti-viral immunity of CTLs has very been shown in relation to the immune response against the COVID-19-causing virus SARS-Cov-2. Impaired CTL effector functions either due to hypo or hyperfunction caused vulnerability to severe disease or the potentially severe disease itself (Westmeier *et al.*, 2020; Kusnadi *et al.*, 2021). Hypofunctional CTLs may cause a lack of protection against pathogens, and hyperfunction may cause high levels of damage.

1.5.6 Role of T cells in CD

CD4$^+$ and CD8$^+$ T cells are increased in the mucosa of active CD patients (Fuss *et al.*, 1996; Müller *et al.*, 1998). T cells in the intestinal microenvironment are exposed to many foreign antigens, including food and commensal bacteria-derived molecules. IL-10 production by Tregs has been found to be essential to maintain tolerance towards the commensal microbiota by e.g., acting on T cells as well as macrophages (Rubtsov *et al.*, 2008). During intestinal inflammation in CD, an aberrant shift toward Th1 and Th17 activity and away from Treg maintained homeostasis has been observed. Breese *et al.* (1993) identified an accumulation

of IL-2 secreting intestinal T cells ex vivo in colonic biopsies derived from a small cohort of paediatric CD patients (n =18) compared to non-IBD controls (n =14) ($p < 0.002$). Two additional studies found an increase in IFN-γ producing $CD4^+$ T cells as well as IL-12 production in the lamina propria of CD patient (Fuss *et al.*, 1996; Parronchi *et al.*, 1997). The IL12B gene was identified as one of the genetic risk loci for CD, further supporting the hypothesis that an aberrant Th17 response is involved in its pathology (Barrett *et al.*, 2009). Additionally, Monteleone *et al.* (2005) found that IL-12 production in isolated LPMCs induces IL-21 production in active CD lesions, a hallmark of Th17 response (Wei *et al.*, 2007). They also found that when stimulated with IL-21, LPMCs increase their IL-12 production, suggesting a positive feedback loop between IL-12 secreting Th1 and IL-21 producing Th17 cells. Th17 cells are involved in the pathology of both CD and UC. GWAS linked a selection of Th17 genes to IBD, including the gene encoding IL-23R (Zhu *et al.*, 2020). Th17 cells have been identified in the tissue of many chronic inflammatory diseases, such as CD, psoriasis, RA and severe asthma, based on their cytokine and TF profile (i.e., IL-17^+RORγt^+) using flow cytometry and qPCR analyses, linking them to autoimmunity (Pène *et al.*, 2008). Fujino *et al.* (2003) quantified IL-17 expression in intestinal tissue and serum of UC (n =20) and CD (n =20) patients as well as control groups, including infectious colitis (n =5), ischaemic colitis (n =8, temporarily disruption in blood flow supplying parts of the colon) and unaffected volunteers (n =15). Interestingly, only individuals with active IBD showed elevated levels of IL-17 and no change was found in patients with non-chronic colitis. In terms of serum IL-17 expression, active CD patients exhibited a significantly higher protein load compared to individuals with active UC.

Based on these findings, anti-IL17 treatment was trialled in a randomised CD cohort. The study used the anti-IL-17 therapeutic secukinumab, a monoclonal antibody binding IL-17A cytokine with high affinity, which had to be discontinued due to an accumulation of severe adverse side effects (Hueber *et al.*, 2012). This suggested that even though IL-17 promotes CD it also has protective characteristics in CD. Supporting this hypothesis is the finding that during a normal intestinal immune response, Th17 cells interact with Tregs and intestinal tissue-resident cells to prevent commensal translocation without mounting inflammatory immune responses (Omenetti and Pizarro, 2015). Thus, while Th17 cells may

contribute significantly to pathology during CD, IL-17 also appears to have additional non-redundant roles in maintaining intestinal homeostasis.

1.5.7 Peripheral T cell-homing to the Intestines

Since T cells (apart from IELs and tissue resident memory T cells) are not tissue resident, they require gut-homing markers to enhance their migration to the intestines. Different receptor-ligand interactions have been found to promote migration to the intestines, with some being selective for small intestine or colon. The α4β7 integrin expressed on T cells, has been found to facilitate their binding to mucosal vascular addressin cell adhesion molecule-1 (MADCAM-1). Mouse studies have shown that α4β7:MADCAM-1 selectively induces homing to the small intestine (Artis *et al.*, 2000; Kuhbandner *et al.*, 2019). Wang *et al.* (2009) studied the effect of C-C chemokine receptor type 6 (CCR6) and CCR9 on the migration behaviour of Th17 cells to the intestines in mice. They found no difference in tissue distribution of $CCR6^+$ Th17 cells between small intestine and colon, however, they found that the $CCR9^+$ cells were more accumulated in the small bowel. CCR9 and α4β7 have been investigated as potential therapeutic targets in CD. The monoclonal antibody against α4β7 called vedolizumab has been a licenced treatment since 2014 (Ha and Kornbluth, 2014; Ranasinghe and Eri, 2018). The GEMINI II trial, a randomised placebo-controlled trial, studied the response to vedolizumab in CD patients. They found a significant increase in clinical remission (defined by CDAI$\leq$150) in vedolizumab treated patients compared to the placebo group ($p < 0.01$) 52 weeks after treatment initiation.

1.6 Summary of intestinal immunopathology in CD

As highlighted in *sections 1.4* and *1.5*, CD harbours as very complex intestinal immunopathology (partially summarised in *Figure 1.3*), which involves mainly different cellular players (not all mentioned in this thesis). Microbiota contribute to tissue homeostasis by shaping the phenotype of intestinal macrophages (Sorini *et al.*, 2018). IL-10 production by lamina propria macrophages has been found to expand $FOXP3^+$ Tregs in mice together with additional administration of TGF-ß (Denning *et al.*, 2007). The source of TGF-ß the intestines are most likely DCs, which secrete the cytokine upon stimulation by commensal-derived antigen

(Kashiwagi *et al.*, 2015). Treg-derived IL-10 in turn provides a tolerogenic cytokine environment for DCs and macrophages and phosphorylates STAT3, inhibiting differentiation of Th17 cells (Chaudhry *et al.*, 2011).

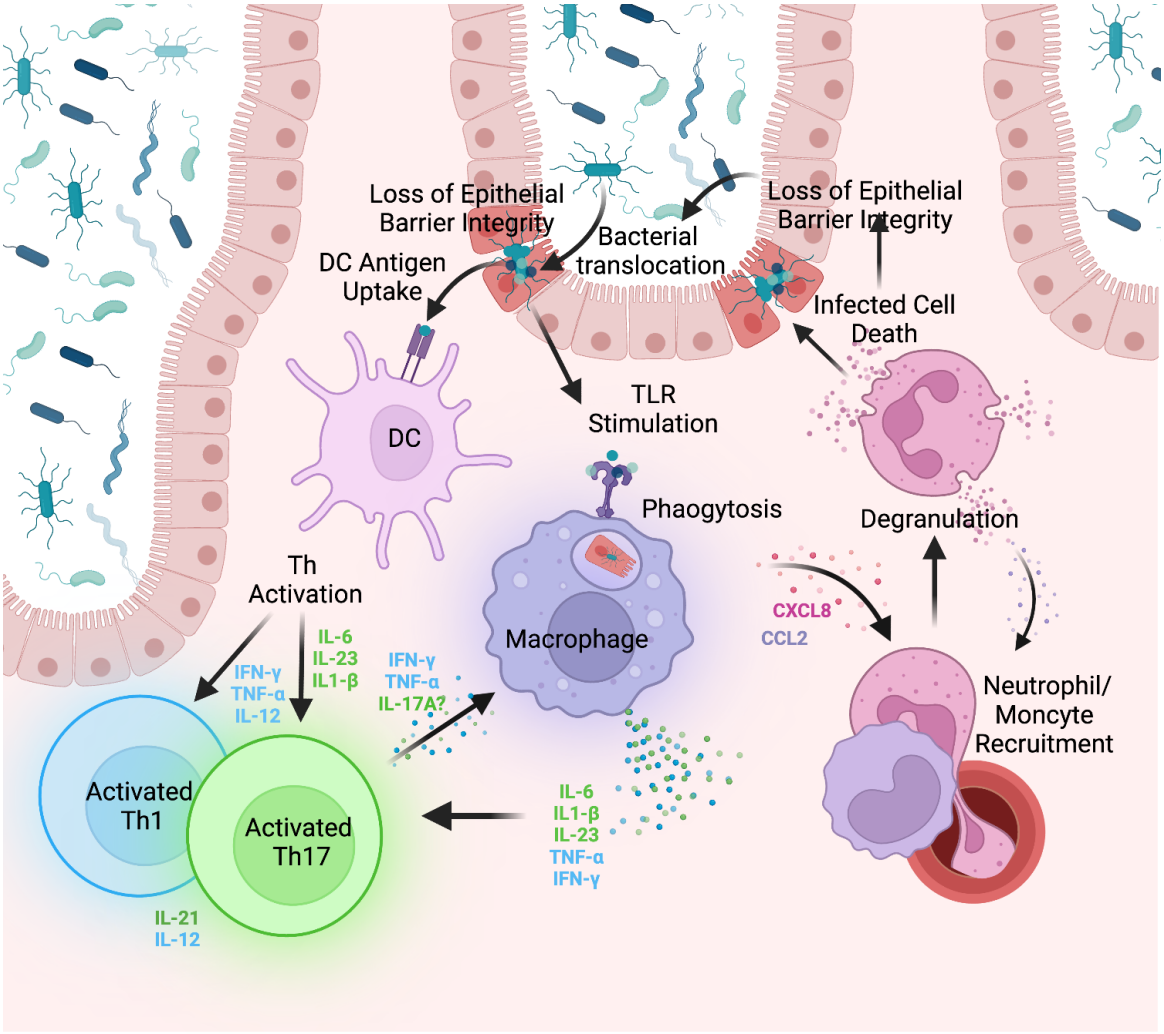


Figure 1.3: The role of mononuclear Phagocytes, CD4⁺ T cells and neutrophils in CD immunopathology.

Microbiota contribute to tissue homeostasis with low level TLR activation of macrophages and DCs and subsequent TGF-beta and IL-10 secretion. These cytokines promote FOXP3+ Tregs which in turn secrete IL-10 contributing to the immunoregulatory environment. In IBD, bacterial translocation and loss of epithelial barrier integrity causes activation of innate immune cells. Their activation causes a shift in the cytokine microenvironment (from IL-10 to Th1 and Th17 cytokines) driving differentiation of Th1 and Th17 effector populations, pro-inflammatory macrophage responses and monocyte/neutrophil infiltration. This figure was created using BioRender.com.

Patients with CD exhibit dysbiosis between the microbiota and host immune system, reflected by less diversity in their intestinal microbiome (Peterson *et al.*, 2008). In contrast to those in homeostasis, intestinal macrophages derived from CD express high levels of pro-inflammatory cytokines, such as *TNF-α, IL-6, IFN-γ, IL-23 and* IL-1β (Kamada *et al.*, 2008; Bernardo *et al.*, 2018). Although they have the capacity to mount an immune response against the translocation bacteria, monocyte-derived macrophages in CD might have impaired ability to effectively

clear bacterial loads (Vazeille *et al.*, 2015). Even though less neutrophil migration has been identified in CD patients (Marks *et al.*, 2006), IL-8 production from intestinal macrophages might still recruit them to the intestine where they kill infected cells and help facilitate Th17 immune responses (Pelletier *et al.*, 2010). The shift from a tolerogenic TGF-ß and IL-10 dominated to a Th1/Th17 pro-inflammatory microenvironment which is further driven by the interaction of the Th subsets. Once activated by DCs and polarised cytokines, LPMCs derived from CD patients secrete IL-12 when stimulated with IL-21 (Monteleone *et al.*, 2005), potentially describing Th1 cells driving Th17 differentiation. Additionally, Th1 and Th17 cells and macrophages might be forming a feedback loop with macrophages with the latter driving Th1 proliferation with *TNF-α* secretion and Th17 with IL-6 and IL-1ß. Th1 cells releasing *IFN-γ* inducing pro-inflammatory macrophage and Th1 responses.

1.7 Granuloma formation in CD

In *chapter 5,* I describe an intestinal CD immunopathotype characterised by granuloma formation. Granulomas are histological structures made up of an accumulation of macrophages in response to a persistent antigen (Pagán and Ramakrishnan, 2018) and are sometimes used as histological marker to distinguish CD from other intestinal inflammatory conditions (Heresbach *et al.*, 2005). They are displayed in 30-40% of CD patients that underwent resective surgery (Wolfson *et al.* 1982) and have been found to be linked to severe disease, defined by higher rates of hospitalisation and surgeries (Molnár *et al.*, 2005).

Granulomas can progress into different structures depending on environmental stimuli as well as the type of antigen (Pagán and Ramakrishnan, 2018). Macrophages can undergo epithelioid changes where the immune cells link their membranes and fuse them, forming a giant cell with multiple nuclei. Additionally, they can form a necrotic core after triggering necrosis, a type of cell death that occurs spontaneously or in response to injury. In the case of macrophages, it can be triggered by over-eating i.e., the cells being overloaded with pathogenic material which they cannot clear using phagocytosis. The precise antigens involved in granuloma formation in CD are unknown. However, bacterial DNA was found in the mucosa of 80% of patients with granuloma exhibiting CD, which

suggests bacterial translocation into the mucosal layer as a potential trigger (Ryan *et al.*, 2004). Granulomas in CD usually remain epithelioid and non-necrotic and are useful to identify the disease and distinguish it from other types of intestinal inflammation (Feakins, 2014). Heresbach *et al.* (2005) studied the prevalence of granulomas as well as the sub-phenotypes in 56 CD patients. 37% of patients exhibited granuloma. 33.3% of granuloma-positive patients developed the epithelioid phenotype.

Granuloma formation and progression have been shown to be influenced by many different cytokines, including IL-4, IL-13, TGF-ß and TNF-α (Pagán and Ramakrishnan, 2018). TNF-α plays an important role in CD pathogenesis, with macrophages being one of its major producers. During mycobacterial infection of mice, TNF-α has been found to drive granuloma responses by inducing the release of macrophage chemoattractants, and its knockout resulted in a poor granuloma response in mice (Roach *et al.*, 2002).

1.8 Role of Chitinase 3-like protein 1 in CD

In addition to granuloma, the CD immunopathotype in *chapter 5* expressed high levels of chitinase-like 3 protein 1 (CHI3L1, also known as YKL-40). CHI3L1 is a non-enzymatic hydrolase that is part of the chitinase-like protein family. Chitinases catalyse chitin degradation; however, a mutation in the CHI3L1 protein sequence hampers its enzymatic functionality (Zhao *et al.*, 2020). Even though it is non-functional as a chitinase, it has been found to be elevated in many chronic inflammatory conditions, including CD, sarcoidosis and RA (Zhao *et al.*, 2020). CHI3L1 has been found to be expressed by mature macrophages (Rehli *et al.*, 2003), neutrophils (Volck *et al.*, 1998; Deutschmann *et al.*, 2019) and tissue-specific epithelial cells, such as colonocytes (Mizoguchi, 2006). Its expression can be induced by IL-6, TNF-α and IL-1ß (Mizoguchi, 2006). Human colonocyte cell lines infected with pathogenic bacteria have been shown to have elevated expression of CHI3L1 (Mizoguchi, 2006). Another study observed an increase in bacterial load after enhancing the expression of CHI3L1 in colonocytes *in vitro* using transfection (Kawada *et al.*, 2008). The same study showed that the secretion of chitin binding protein, secreted by bacteria to invade chitin-expressing hosts (insects and crustaceans) and to bind soluble chitin (which they

use as an energy source), was required for effective bacterial adhesion to colonocytes. Fluorescent microscopy showed binding between CHI3L1 and chitin binding protein overexpressing *E. coli*, suggesting that the interaction between these two proteins promotes bacteria to adhere.

A study by Deutschmann *et al.* (2019) identified CHI3L1 as a potential autoantigenic target in CD. PB sera were collected from 110 CD patients and 86 non-IBD controls. Neutrophil proteins were extracted from healthy donors and incubated with patient and control sera. Immunoblotting and enzyme-linked immunosorbent assays (ELISAs) showed binding of CHI3L1 by antibody in patient but not control sera. When looking at the level of serum reactivity with CHI3L1 in CD, UC, and coeliac disease (an inflammatory condition causing gluten intolerance) it was significantly upregulated in CD compared to the other two conditions. Interestingly, CHI3L1 has also been shown to induce secretion of the neutrophil chemoattractant IL-8 in colonocytes derived from UC patients (Chen *et al.*, 2011) and has also been demonstrated to induce macrophage recruitment in colorectal cancer (Kawada *et al.*, 2011).

In addition to binding chitin, CHI3L1 can act on a selection of receptors. It can bind to the IL-13 Receptor Alpha 2 (IL-13Rα2), usually a target of IL-13 (He *et al.*, 2013; Lee *et al.*, 2016). He *et al.* (2013) found that CHI3L1 and IL-13Rα2 are co-localised in the cytoplasm of murine alveolar macrophages. They also demonstrated that when treated with recombinant CHI3L1, the human THP-1 cell line (immortalised macrophages) upregulated a selection of kinase signaling cascades, including extra-signal-regulated protein kinase (ERK) 1/2, Akt and Wnt/beta-catenin pathways, all of which are crucial for cell differentiation and proliferation. These effects were impaired when simultaneously treated with IL-13Rα2 small interfering RNA (siRNAs) or soluble IL-13Rα2 protein hijacking recombinant CHI3L1. Another study by Geng *et al.* (2018) reported that selective deletion of CD44, an extracellular glycoprotein, or IL-13alpha2 receptor in human gastric cancer cell lines prevented the upregulation of these signaling pathways and showed that CHI3L1:CD44:IL-13α2 receptor form a multi-protein complex. These findings suggest that CHI3L1 plays an important role in tissue remodelling. Supporting evidence was recently published by Zhou *et al.* (2022), linking increased CHI3L1 in patient serum with an increase incidence of pulmonary

fibrosis, a condition characterised by macrophage-induced mucosal tissue remodelling (Zhou *et al.*, 2015). In the same study it was demonstrated that CHI3L1 induced collagen deposition in fibroblasts of mice by interacting with the prostaglandin D2 receptor (CRTH2). CRTH2 has been previously linked to fibrosis in a selection of organs as well as fibroblast collagen synthesis (Zuo *et al.*, 2021). CHI3L1 relation to fibrosis might elucidate its role in chronic inflammatory conditions, such as CD, which are often characterised by the presence of fibrotic tissue.

CHI3L1 has not only been measured in human intestinal cells and serum but also in faeces, making it a potentially interesting biomarker. In fact, faecal CHI3L1 has been identified as a potential marker for intestinal inflammation in IBD. The stool was collected from 48 CD and 44 UC paediatric patients with different states of inflammation and 56 non-IBD controls (Bohr *et al.*, 2015). CHI3L1 levels were significantly increased in the stool of patients with endoscopically active IBD compared to controls and remission disease sample groups. Using predictive analysis, it was found that CHI3L1 could predict inflammation with a specificity of 88.9% and sensitivity of 84.7%. Additionally, a study by Buisson *et al.* (2016) identified faecal CHI3L1 to be also predictive for inflammation in adult CD patients and to correlate with CDEIS and the commonly used faecal biomarker calprotectin.

1.9 Current CD therapeutics

There is now a broad range of therapeutics available to treat CD. They vary from general immunosuppressive drugs to biological treatments, which target specific molecules (such as cytokines) that are heavily involved in the patient immune response. The goal of CD management is to re-establish mucosal healing.

1.9.1 General immunosuppressants

Once a patient is diagnosed with CD or experiences a strong flare, the first-line treatment, as advised by the UK National Institute for Health and Care Excellence (NICE), are corticosteroids (NICE, 2022). They are fast-working ant-inflammatory agents that initially induce remission before introducing maintenance therapy. Patients should be weaned off corticosteroid treatments once remission is

achieved because of their severe side effects (NICE, 2017; Lamb *et al.*, 2019). Corticosteroids diffuse through the cell membrane and bind the glucocorticoid receptor in the cytosol (Barnes, 2006). This protein receptor complex can re-locate to the nucleus and interfere with the induction of pro-inflammatory transcriptional programs, such as the NF-kB cascade. If steroid treatment is contraindicated, e.g., due to underlying infections or young age, aminosalicylate (5-ASA) or EEN treatment can be prescribed (NICE, 2022). Other monotherapies used for long-term CD maintenance (after steroid-induced remission is reached) are thiopurine, aminosalicylates (5-ASA), and methotrexate (NICE, 2022).) Thiopurines, 5-ASA and methotrexate are general immunosuppressants that hamper immune cell differentiation (Rousseaux *et al.*, 2005; Djurić *et al.*, 2018). In the case of methotrexate, this is achieved by interfering with purine synthesis, which is especially important in T cell proliferation (Fairbanks *et al.*, 1999).

1.9.2 Biologic drugs

If immunosuppressants fail to manage the CD inflammation, another group of medications is routinely used, which is called biologics (Hommes, 2006). These are a group of anti-inflammatory drugs that target specific molecules that drive inflammation in chronic inflammatory diseases such as CD. The most commonly prescribed biologics in IBD, and also in RA, are TNF blocking agents. These drugs hijack soluble TNF, preventing it from receptor ligation and subsequently inhibiting its pro-inflammatory downstream effects. The ACCENT I randomised trial studied the efficacy of the anti-TNF antibody infliximab in 573 CD patients with a CDAIS >220, comparing different concentrations of the medication to a placebo group (Hanauer *et al.*, 2002). Disease progression was evaluated after 30 weeks of treatment, and clinical remission (defined by CDAIS <150) was achieved in significantly more patients in all groups consistently treated with anti-TNF (average 42%) compared to placebo (21%). In addition to initial nonresponse to the treatment (referred to as primary non-responders) are subsets of patients who stop responding over time and/or develop allergies/adverse reactions. Many of these adverse reactions occur because they start producing antibodies against the drug, forcing discontinuation of the therapy (secondary loss of response) (Gisbert and Panés, 2009). The commonly used anti-TNF drugs are adalimumab and infliximab.

Ustekinumab and vedolizumab are the two other types of biologic drugs licensed for CD, usually administered either as add-on therapies to general immunosuppressants or anti-TNF blocking agents or in anti-TNF refractory disease. Ustekinumab targets the p40 subunit of the cytokines IL-12 and IL-23, targeting both Th1 and Th17 effector functions (Sandborn *et al.*, 2012). Vedolizumab binds to the α4ß7 integrin, and so inhibits the migration of leukocytes to the intestine (Ha and Kornbluth, 2014).

1.10 Crohn's disease stratification

The current treatment model for CD is based on trial and error, starting patients off with steroids, followed by general immunosuppressants. Biologics are more commonly used now than they were in the past because their market value has reduced, and their benefits have become clear. However, they are usually prescribed in the UK as add-on therapy or last resort after other treatment options have failed. This step-up treatment approach can be very detrimental for patients not only in terms of their health but also in their quality of life. Overall biologic treatments have improved the clinical management of CD, but the surgery rate remains high with up to 46% (Frolkis *et al.*, 2013). Gisbert and Panés (2009) summarised the relapse rate during infliximab treatment in 16 studies (a total of 2236 patients) and estimated that 40% of patients lose response to infliximab after initial treatment success. Together with the cases of primary non-response less than 20% of CD patients can be successfully treated long-term with anti-TNF blockers. Additionally, studies have suggested that previous exposure to anti-TNF treatment increases the risk of refractory disease during therapy with follow-up biologics such as vedolizumab (Sands *et al.*, 2017; Verstockt *et al.*, 2020).

To improve the quality of CD treatment, researchers and clinicians have been trying to identify biomarkers which will first predict the progression of the disease and second predict whether patients will respond to treatments such as anti-TNF blockers. My colleagues and I have recently reviewed potential predictive markers in IBD (Wang and Baer *et al.*, 2019).

1.10.1 Prediction of disease progression

CD is a very heterogeneous disease which can spread anywhere within the gastrointestinal tract and presents unpredictable flares and complications. Predicting the progression of CD would allow clinicians to introduce prophylactic measures instead of solely reacting to the potential damage that has already been caused by the disease. Examples of disease characteristics that would benefit from biomarkers are disease location, phenotype (i.e., stricturing/fistulating), the risk for EIMs, relapse and surgery.

In addition to the serological laboratory markers of inflammation discussed in *section 1.3.2* anti–*Saccharomyces cerevisiae* antibody (ASCA) and perinuclear anti-neutrophil cytoplasmic antibody (pANCA) have been studied in the context of disease progression. A cohort of 738 CD patients was studied, and their blood was tested for multiple serological markers, including ASCA and pANCA (Ferrante *et al.*, 2007). Subsequently, the association between ASCA and pANCA and ileal or colonic inflammation was tested, which identified ASCA as a promising marker for ileal involvement ($p = 0.001$), complications (i.e., fistulating or stricturing disease, $p < 0.001$) and risk for surgery ($p < 0.001$). In contrast, pANCA expression in PB of patients was linked to colonic involvement ($p < 0.001$). The IBSEN study followed 159 patients for a 10-year and 135 patients for a 20-year time period and sampled their blood at one or both time points (113 matched samples) and collected their surgical information (Kristensen *et al.*, 2020). Univariate analysis results displayed that negative pANCA and positive ASCA antibodies are predictive of the need for surgery, which supports the findings by Ferrante *et al.* (2007). Even though these findings seem promising, recent guidelines by the European Crohn's and Colitis Organisation (ECCO) and European Society of Gastrointestinal and Abdominal radiology (ESGAR) do not recommend pANCA or ASCA as routine clinical markers due to their moderate sensitivity and specificity (Maaser *et al.*, 2019).

To improve the reliability of serological markers in predicting disease complications in CD, Lichtenstein *et al.* (2011) trialled a study using a combination of clinical information, genetic and blood biomarkers. Matching *NOD2* genetic phenotyping and PB ASCA levels were available for 385 CD patients, and they were classified based on the Montreal classification. Occurrences of stricturing and

fistulating disease at the time of sample collection were referred to as 'disease complications'. ASCA levels were increased in patients with complicated disease, which supported the observations by (Ferrante *et al.*, 2007). Subsequently, they selected the three *NOD2* polymorphisms with the highest incidence within the patient cohort. Regression between the complicated CD patient group and the variables disease duration, ASCA levels together with the presence of the selected *NOD2* polymorphism showed a significant association ($p < 0.001$). Additional, predictive modelling indicated an accuracy of around 80% for the prediction of complicated disease. Even though these findings sound promising, this study harbours a few limitations. When looking for a predictive biomarker, researchers and clinicians aim to identify complications prior to developments; however, in this study, they only linked current fistulising or stricturing with changes in ASCA and *NOD2* polymorphisms. Additionally, a biomarker has to be cost-efficient, and the use of genetic, serological analysis as well as subsequent advanced statistical analysis by a statistician might not be feasible for routine clinical care. Supporting these findings are several studies which linked significantly increased frequency of *NOD2* with severe disease. For example, *NOD2* as well as *IL23R*, and *AT16GL1* polymorphisms had a higher occurrence in CD patients with disease onset before 40 years of age, surgical intervention and complications (fistulating/stricturing disease) compared to individuals with disease onset >40 years, no surgeries and disease without complication (Weersma *et al.*, 2009).

Peripheral blood mononuclear cells (PBMCs), leukocytes extracted from the PB, have been studied in the context of relapse risk in CD. $CD4^+$ effector T cells expressing the mucosal marker ADP-ribose cyclase 1 (CD38), were found to be elevated in PB of paediatric IBD patients with active disease (n =22, defined by endoscopic examination) when compared to patients in remission (n =26) (Joosse *et al.*, 2018). This $CD38^+CD4^+$ effector T cell population was also found to upregulate T-cell immunoglobulin and ITIM domain (TIGIT) expression, a T cell inhibitory marker, after 72h co-culture with APCs in cells derived from adult healthy blood. To follow up, the predictive potential of $TIGIT^+CD38^+CD4^+$ effector T cells was studied in PB of 18 paediatric patients with ongoing inflammation and 9 non-IBD controls. TIGIT expression on $CD38^+CD4^+$ appeared to split the IBD cohort into two subsets with a cut-off at 25% $TIGIT^+CD38^+CD4^+$ T cells. The percentage of paediatric patients with low $TIGIT^+CD38^+CD4^+$ circulating T cells exhibiting relapse

within approximately 800 days of follow-up post-sampling was significantly higher than that of the high $TIGIT^+CD38^+CD4^+$ circulating T cell group. Therefore, this PB effector T cell subset might have predictive potential intestinal inflammation flares for around 2 years after sampling. For these results to be of clinical relevance, this analysis must be repeated in a larger cohort as well as in adult patients.

The faecal marker FC has not only been identified as a useful measure of inflammation in CD but also has been shown to function as a potential predictor for the length of remission periods in IBD (García-Sánchez *et al.*, 2010). 66 patients with CD and 69 patients with UC who had at least been in remission for 3 months (defined by CDAI) were recruited. They were monitored for 1-year post-sampling of FC, which showed relapse in 27% of CD and 31% of UC patients. The FC in IBD relapse was 444ug/g compared to 112ug/g in patients without relapse. However, for CD patients with isolated ileal disease, the sensitivity was only 50%, suggesting that FC might be a reliable marker in colonic CD and UC but less so for predicting relapse of inflammation in the ileum. Interestingly, this is a similar pattern to the pANCA serum antibody, which like FC is linked to neutrophil inflammation, suggesting that neutrophilic disease is linked to colonic rather than ileal disease. This could be due to a lower secretion level in the ileum or the degradation of the protein before the final stool sample is collected.

1.10.2 Prediction of treatment response to biologics

Anti-TNF is the front-line biological treatment for CD, often used as rescue therapy, which is why many researchers focus on identifying biomarkers for treatment response prior to initiation of the treatment.

Like disease progression, specific gene polymorphisms have been found to be associated with non-response to anti-TNF treatment. A systematic review by Bek *et al.* (2016) summarised the findings of 15 studies which investigated the potential of genetic risk factors in predictive anti-TNF treatment response. Polymorphism with odd ratios (OR) for responders vs non-responders higher than 2 were *TLR2* ($OR_{rs1816702}$ =2.02), *TLR4* ($OR_{rs5030728}$ =3.34) and TNF receptor superfamily 1 A (*TNFRSF1A*, $OR_{rs4149570}$ =2.39), one of the two TNF receptors. The

effects of the polymorphism for *TLR4* are currently unknown, but the other variant of *TLR2* has been linked to its overexpression on human Caucasian monocytes (Bielinski *et al.*, 2011), potentially driving the increase in TNF-α production. Similarly, the *TNFRSF1A* polymorphism rs4149570 has exhibited to induce increased expression of the receptor in PBMCs derived from children with Kawasaki disease, a paediatric inflammation of the blood vessels (Wang *et al.*, 2011), which unsurprisingly might drive enhanced inflammation via the TNF:TNFRSF1A axis. Overall, the understanding of genetic risk factors for anti-TNF refractory disease is important to elucidate the mechanisms driving non-response; however, as mentioned previously, genetic phenotyping is currently not suitable for routine clinical care.

Verstockt *et al.* (2019) studied the gene expression of a selection of potential biomarker targets for anti-TNF non-response in RNA derived from whole blood and intestinal pinch biopsies at baseline of anti-TNF treatment. They recruited 24 CD and 30 UC patients. 31 of these patients reached endoscopic remission (as defined by endoscopic examination) 6 months post-treatment initiation. Significantly increased expression of triggering receptor of myeloid cells-1 (*TREM1*), known to be expressed on intestinal pro-inflammatory macrophages during IBD inflammation (Schenk *et al.*, 2007), was found in non-responders compared to responders in both whole blood and inflamed intestinal tissue ($p < 0.001$). Prins *et al.* (2021) recently published that $CD14^+$ monocytes isolated from PBMCs of responders had lower *TREM1* gene expression than those of non-responders to anti-TNF, supporting the findings of Verstockt *et al.* (2019). Therefore, high levels of TREM-1 protein and/or gene expression in whole blood and *TREM1* expression in the intestine is a potential indicator of non-response to anti-TNF treatment.

Another gene frequently associated with non-response to anti-TNF in IBD is oncostatin-M (*OSM*). Prins *et al.* (2021) found an increase in *OSM* in inflamed mucosa of non-responders to anti-TNF at baseline compared to responders, but no differences in the serum protein levels between the patient subsets. OSM is part of the IL-6 cytokine family and has been frequently linked to chronic inflammation (Richards, 2013). Verstockt *et al.* (2021) studied the relationship between serum OSM expression, treatment response to anti-TNF ($n_{responder}$ =91, $n_{non\text{-}responder}$ =95) and Vedolizumab treatment ($n_{responder}$ =99, $n_{non\text{-}responder}$ =75) in IBD patients. No

difference was found between serum levels of OSM in the patient subsets starting either therapy supporting initial findings from Prins *et al.* (2021). In addition to OSM serum levels, they also looked at mucosal gene expression by transcriptomics (microarray and RNAseq, anti-TNF cohort: ($n_{responder}$ =13, $n_{non\text{-}responder}$ =12), vedolizumab cohort: $n_{responder}$ =23, $n_{non\text{-}responder}$ =24). They found an increase in *OSM* as well as OSM receptor (*OSMR*) in the mucosa of patients non-responsive to either biologic. These findings were confirmed *in vivo* using an anti-TNF resistant mouse model of colitis (West *et al.*, 2017). *Osm* neutralisation reduced colitis pathology and resulted in a similar pro-inflammatory cytokine profile as steady state controls, suggesting its potential as an alternative therapeutic to anti-TNF.

In addition to *TREM1* and *OSM* many genes have been found to be upregulated in the mucosa of non-responders at baseline of anti-TNF treatment. These genes include *TNF* itself and its receptor TNF receptor 2 (*TNFR2*)(Verstockt *et al.*, 2019). Additionally, membrane-bound TNF (mTNF) on the luminal side of the mucosa was visualised *in vivo* during colonoscopies of CD patients by using a fluorochrome-conjugated anti-TNF antibody (Atreya *et al.*, 2014). When comparing the level of mTNF between responders (n =13) and non-responders (n =12) low protein expression predicted non-response with a sensitivity of 84.6% and specificity of 91.7%. These results support the previous findings by Verstock *et al.* (2019) and propose that higher TNF titres indicate TNF-dependent disease, which is more susceptible to anti-TNF treatment. Genes encoding for the calprotectin protein complex (*S100A8/9*) were also upregulated (Arijs *et al.*, 2010; Leal *et al.*, 2015). Upregulation of *S100A8* and *S100A9* suggests FC as a potential faecal marker for not only inflammation but also anti-TNF non-response. However, no change in FC baseline levels between responders and non-responders was found (Guidi *et al.*, 2014), potentially due to the limitations discussed earlier in regards to FC levels and disease location.

The cytokines encoded by *IL1B, IL6, IL17A* and IL23 subunit alpha (*IL23P19*) genes are all part of the Th17 immune response (Aggarwal *et al.*, 2003; Acosta-Rodriguez *et al.*, 2007; Wang *et al.*, 2013) and were found to be upregulated in anti-TNF non-responders prior to treatment induction. Furthermore, the same behaviour was seen in the expression of the gene encoding for the IL-13Rα2, which facilitates the signaling of the Th2 cytokine IL-13. As discussed in *section 1.6.5*. Th17 T cells

facilitate pathology in CD. Th2 cells, in turn, are more associated with UC, so the potential association with anti-TNF refractory disease might be due to studies studying UC and CD as one cohort. Most of the genes in this paragraph appear to be linked to a general inflammatory response in CD/UC.

1.10.3 Proposed mechanism of primary anti-TNF non-response in CD

Martin *et al.* (2019) identified 33 lamina propria cell phenotypes in inflamed ileal tissue derived from 11 CD patients using single cell RNA sequencing (scRNAseq). Correlation of the cell phenotype frequencies across all samples revealed a signature related to IgG-producing effector B cells, inflammatory MNPs, activated T and stromal cells (connective tissue cells), which was enriched in a subset of patients. Mapping of the signature on a larger paediatric CD gene expression dataset (n =340) derived from inflamed ileum at baseline of anti-TNF non-response showed a significant increase in patients with anti-TNF refraction 18 to 24 months after treatment initiation. Amongst the key drivers of this immunopathotype in ileal CD were TNF, OSM, IL-1β secretion by blood-monocyte derived pro-inflammatory macrophages, and TNF, IL-6, IL-1β and IL-23 secretion by T cells. Thus type 1 and 3 immune responses appear to perpetuate disease during anti-TNF treatment.

An additional predictive signature for anti-TNF non-response in CD was identified by Friedrich *et al.* (2021) and included two gene modules which were linked to stromal cells, MNPs and neutrophils. Transcriptomic analysis of fluorescence-activated cell sorted (FACS) stromal, neutrophil and MNP populations showed that most genes of the modules were linked to neutrophils and stromal cells and less so MNPs. Within the neutrophil-enriched genes were *OSM* and *TREM1* genes, which were previously found to be upregulated in non-responders. The stromal cell population expressed high levels of neutrophil-chemoattractants, including *CXCL8* and *CXCL5*. Stimulation of stromal cell lines with TNF and IL-1β showed upregulation of these chemoattractants, but treatment with OSM or IL-6 did not. Co-culturing of the cell lines and media from patient-derived cell cultures had similar effects. The IL-1-targeting therapy anakinra reduced chemoattractant expression, which was not reflected in cultures treated with the anti-TNF therapeutic adalimumab. Together these two studies propose a model of immune

and stromal cell interaction that potential drives non-response *via* alternative pro-inflammatory cytokine programs to TNF, such as IL-1ß.

Overall, non-response to anti-TNF has been linked to many potential biomarker targets (summarised in *Figure 1.4*). Given the heterogeneity of CD, the mechanisms underlying non-response discussed in *section 1.8.3* can either relate to a single complex immune response or to two distinct intestinal immunopathotypes.

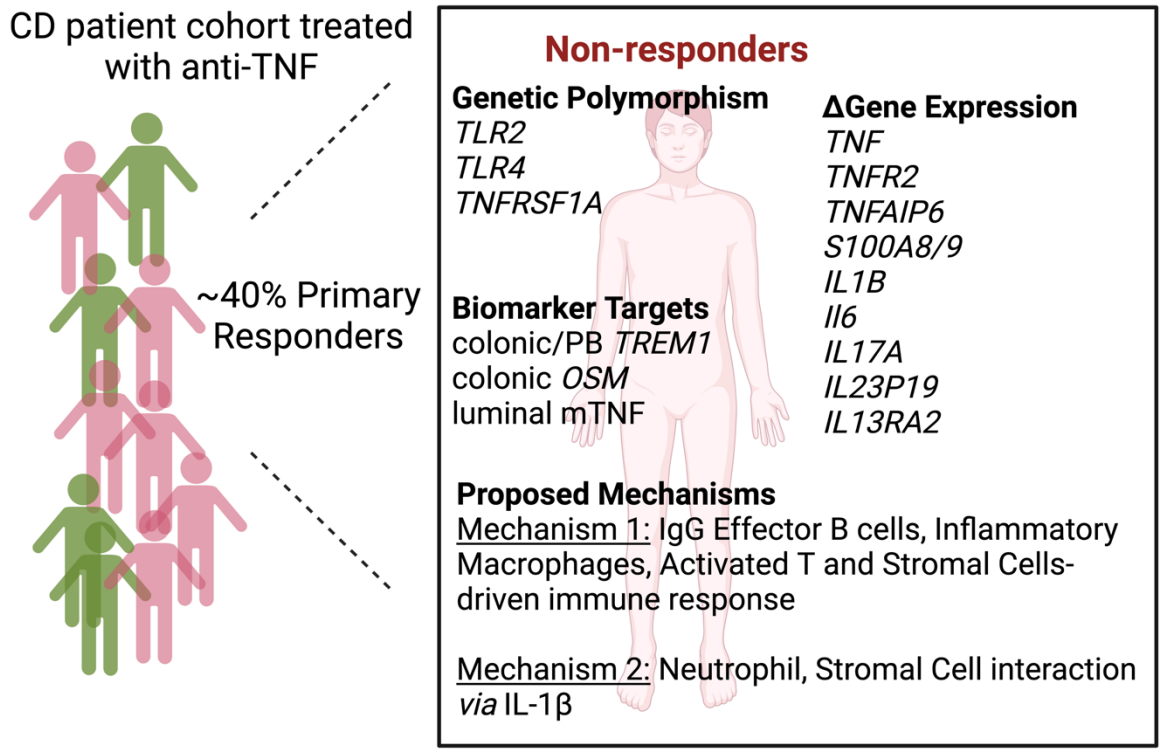


Figure 1.4: ***Summary of anti-TNF non-response biomarkers and proposed underlying mechanisms in CD.***

Currently up to 60% of CD patients cannot be treated long-term with anti-TNF agents, making a clinical biomarker for non-response for this specific treatment very desirable. Many different characteristics have been found in non-responders to anti-TNF treatment. A selection of genetic polymorphisms, which include but are not restricted to *TLR2*, *TLR4* and *TNFRSF1A*, and altered gene expression exhibiting an increase in pro-inflammatory cytokines (IL1-ß, IL-6, IL-17A, OSM, TNF and IL-23) and TNF-signalling (TNFR2, TNFAIP6). Current biomarker targets, selected because they have repeatedly been associated with anti-TNF non-response, are *TREM1* (highly expressed on pro-inflammatory intestinal macrophages), OSM and TNF. Two mechanisms driving ant-TNF non-response have recently been proposed by Martin *et al.* (2019) and Friedrich *et al.* (2021). Mechanism 1 described a complex interaction driven by IgG, pro-inflammatory macrophages together with T and stromal cell responses. Mechanism 2 described an IL1-ß-dependent feedback loop between neutrophils and stromal cells. This figure was created using BioRender.com.

1.11 Hypothesis and Aims

Even though mechanisms underlying anti-TNF non-responses in CD have been proposed, there is currently no predictive marker routinely used in clinical care. Moreover, clinicians are still unable to identify patients before relapsing, developing complications, or spreading of the disease to other parts of the GI. The studies by Martin *et al*. (2019) and Friedrich *et al*. (2021) suggest that patients with non-response have specific immunological mechanisms that drive their disease.

Based on these findings, I hypothesised that distinct immunopathotypes in CD drive distinct disease phenotypes and underlie distinct treatment responses to biologic therapies, such as anti-TNF.

To test these hypotheses, I aimed to:

1. Study T cell phenotypes in the PB of CD patients using flow cytometry analysis for cell and multiplex ELISA for cytokine quantification.
2. Compare changes in PB T cells to haematological markers of inflammation.
3. Link distinct PB phenotypes to specific intestinal immune responses using matched blood and gene expression data.
4. Stratify patients based on their intestinal immune signatures and test their association with anti-TNF refractory disease.
5. Quantify potential gene and proteins of interest using clinically available laboratory methods such as qPCR and IHC.

Chapter 2: Materials and Methods

2.1 Patients

Samples were collected under the "NHS GG&C additional biopsies, blood, urine and/or stool samples for direct use in research projects or archived for future use in the Research Tissue Bank (RTB) at the GBRC (Glasgow Biomedical Research Centre)" ethics (REC 4 approvals *14/WS/1035* and *19/WS/0111*). Informed consent was obtained by qualified medical staff prior to sampling, and samples were processed in the GBRC. Patient recruitment and ethical approval for the publicly deposited datasets GSE94648, GSE126124 and GSE16879 were previously described by Planell *et al.* (2017), Palmer *et al.* (2019) and Arijs *et al.* (2010), respectively.

2.1.1 Inclusion and exclusion criteria

Active Disease (CD inflamed)	
Inclusion Criteria	Exclusion Criteria
High Harvey Bradshaw Index (HBI) >4 OR endoscopic activity	Pregnancy Other autoimmune/inflammatory diseases /cancer Low HBI

Table 2.1: ***Inclusion and exclusion criteria CD patients with active disease.***

Inactive Disease (CD remission)	
Inclusion Criteria	Exclusion Criteria
Low HBI ≤4 OR no endoscopic activity	Pregnancy Other autoimmune/inflammatory diseases/cancer High HBI

Table 2.2: ***Inclusion and exclusion criteria for CD patients with inactive disease.***

2.1.2 Peripheral blood samples

Peripheral blood (PB) samples from patients with CD were collected in the Gastroenterology Day Unit at the Glasgow Royal Infirmary (GRI) by Ann-Marie Scott. Healthy volunteers for age and sex-matched controls were recruited according to ethics approved by the College of Medical, Veterinary and Life Sciences in the University of Glasgow at the GBRC. Samples were collected in tubes with coagulating agents lithium heparin and ethylenediaminetetraacetic acid (EDTA) for further processing. IBD patients were sampled for our analysis and for routine clinical blood tests on the same day. This enabled collection of CRP levels, total WBC neutrophil counts and lymphocyte counts, and the HBI score.

2.1.3 Intestinal tissue samples

Intestinal pinch biopsies were collected by clinicians during routine endoscopy procedures in the Endoscopy Unit of the GRI under the supervision of Dr. Daniel Gaya. Non-IBD controls were invited for the procedure by the Scottish Bowel Cancer Screening Programme or for symptom check-ups. Tissue was collected in RNA*later*™ Stabilization Solution (Cat. No. AM7020, Thermo Fisher) for RNA extraction or immediately fixed in 10% buffered Formalin (Cat. No. FOR0020AF59001, VWR) for IHC. CDEIS was taken as a measure of inflammation.

2.2 Peripheral blood mononuclear cell (PBMC) isolation

Fresh blood was collected in lithium heparin tubes and processed within 4 hours of sampling. Blood was diluted with Phosphate Buffered Saline (PBS) (1:1), layered on a gradient (Histopaque®-1077 Hybri-Max™, Cat. No. H8889-100ML, Sigma-Aldrich) and separated by centrifugation at room temperature (2100g for 30mins, without brake). Subsequently, they were washed twice at 4°C with PBS (1X 200xg for 10mins, 1X 400xg for 10mins) to remove Histopaque®-1077 Hybri-Max™, plasma, and red blood cell residues and kept on ice for further processing.

2.3 Plasma isolation

Fresh blood collected in EDTA tubes was separated into blood cells and plasma by centrifuging at room temperature (1200xg 20mins). Subsequently, the plasma layer was removed and stored in cryovials at -20°C for short-term and -80°C for longer-term storage.

2.4 Bulk RNA sequencing (preliminary data)

These preliminary data were generated by Boehringer Ingelheim using colonic CD patient-derived colonic tissue samples in collaboration with The Milling Group. Illumina Sequencing using a poly-A tail library was used for RNA sequencing. The raw sequencing reads were aligned using TopHat alignment (BaseSpace App), and FastqC (BaseSpace Labs App) was applied for quality control and adaptor trimming (courtesy of John Cole, University of Glasgow). The Bioconductor DESeq package was used for differential gene expression analysis in R studio.

2.5 Extracellular flow cytometry

PBMCs were stained on ice for 30mins with live dead dye (eBioscience, eFlour780; 1:1000) to identify live cell populations and human FCR Binding inhibitor block (eBioscience, 1:200) to prevent non-specific antibody binding. After one PBS wash (400xg, 5mins at 4°C) the extracellular antibody panel summarised in *Table 2*.3 was added to $2x10^6$ cells/ml in 200µl staining volume at room temperature for an additional 30mins. After the extracellular antibody staining, the wash step was repeated twice with PBS (+5%FBS). Cells were then resuspended in 100ul PBS (+5%FBS) and analysed using a LSR Fortessa cytometer (BD Bioscience). Flow cytometry data was quantified in FlowJo™ (Version 10.8 and below).

Target	Fluorophore (Clone)	Catalogue No. (Company)	Dilution
CD45	BV510 (HI30)	304036 (BioLegend)	1/100

CD3	AF700 (HIT3a)	300324 (BioLegend)	1/100
CD4	BV785 (RPA-T4)	300554 (BioLegend)	1/100
CD8	BV605 (RPA-T8)	301040 (BioLegend)	1/100
CD62L	PE (DREG-56)	304806 (BioLegend)	1/100
CD45RA	PerCPCy5.5 (HI100)	304122 (BioLegend)	1/100
CD16	FITC (3G8)	302006 (BioLegend)	1/40
CD56	AF647 (MEM-188)	304612 (BioLegend)	1/40
CD25	BV421 (BC96)	302630 (BioLegend)	1/50

Table 2.3: ***Extracellular FACS antibody panel.***

2.6 Luminex multiplex assay

The Luminex Multi-Analyte Profiling technology from ThermoFisher Scientific uses a magnetic-labelled bead assay to allow for the quantification of multiple protein targets. 12 cytokine protein targets were selected for the Luminex assay: IL-1ß, IL-6, CCL2, IL-21, IFN-γ, IL-10, IL-23, TNF-α, IL-17/17A, IL-12p70 subunit, IL-31, and IL-4. The assay was performed as per the manufacturer's instructions and the output was measured using the Bio-Plex 200 instrument (Bio-Rad).

2.7 Immunohistochemistry (IHC)

Intestinal pinch biopsies were fixed in 10% buffered Formalin (Cat. No. FOR0020AF59001, VWR) for 24 to 72hrs prior to embedding into paraffin blocks. Tissue was sectioned with a thickness of 5µm using a microtone and subsequently loaded on KP-PLUS positively charged frosted slides (Cat. No. KP-3056, PFM). The sections were dried onto slides overnight at room temperature and baked for 35mins at 60°C prior to rehydration. The de-waxing and re-hydration process

included the following steps: 2X 5min Xylene, 2X 3mins 100% Ethanol, 2X 3mins 90% Ethanol, 2X 3mins 70% Ethanol, 3mins dH_20. Subsequently, the heat-induced antigen retrieval method was employed for 5mins under pressure in full power setting of the microwave. Antigen retrieval buffer varied between 0.01M citrate buffer pH 6.0 (TCS HDS05-100) and 10% EDTA buffer (pH 8.5), depending on the primary antibody. After cooling down, slides were washed in dH_20 for 10mins, and sections were incubated with BLOXALL® Endogenous Blocking Solution (Cat. No. SP-6000-100, Vector Laboratories) for 10mins at room temperature to quench tissue endogenous peroxidases and prevent their reaction during the developing stage of IHC. Slides were washed in dH_20 for 10mins, and tissue was blocked for 1hr at room temperature with 1XCasein solution (Cat. No. SP-5020-250, Vector Laboratories) in DAKO antibody diluent (Cat. No. S0809, DAKO) to prevent non-specific binding of secondary antibody. The blocking solution was removed without a washing step. Primary antibodies were applied for 1hr at room temperature or overnight at 4°C with the following targets and concentrations: mouse anti-CD68 (DAKO, GA613), 0.5ug/mL rabbit anti-CHI3L1 (abcam, ab77528), and 100ug/mL goat anti-TLR2 (abcam, ab1655). The subsequent washing step included 2X 5mins in tris-buffered saline (TBS). ImmPRESS polymer detection kits for IHC (Vector Laboratories) were used un-diluted as secondary antibody and incubated on section for 30mins at room temperature with a specificity against host-species of primary antibody. The subsequent washing step included 2X 5mins in TBS. At developing stage DAB substrate kit, with 1 drop diluent per 1ml of diluent, was incubated with sections for 5mins or until saturated (SK-4100, Vector Laboratories). Slides had a final wash in dH_20 for 10mins. Sections were counterstained in haematoxylin for 1min followed by 2 dips in acetic acid and 30s in Scott's Tap water solution. Finally, slides were dehydrated the following way: 1min 70% Ethanol, 2X 30s 90% Ethanol, 2X 3mins 100% Ethanol, 2X 3mins Xylene. Slides were mounted with a coverslip using VectaMount® Mounting Media for IHC (H-5700-60, Vector Laboratories). IHC signal was quantified using the QuPath software as previously published (Bankhead *et al.*, 2017).

2.8 *IL1B* RNAscope *in situ* hybridisation with anti-CD68 IHC co-staining

Intestinal pinch biopsies were pre-processed and sectioned as indicated in *section 2.9*. Subsequently, the sections were mounted onto SuperFrost® microscope slides (631-0113, VWR) to ensure better adhesion of tissue to slides during the RNAscope process. RNAscope® was performed using the Biotechne assay with all reagents summarised in *Table 2.4*.

Reagents List for RNAscope®	
Catalogue Number	Reagent
322350	RNAscope® 2.5 High Definition (HD)- RED Assay
310361	RNAscope® Probe - Hs-IL1B - Homo sapiens IL-1 beta (*IL1B*), mRNA
310043	RNAscope® Negative Control Probe - DapB - Bacillus subtilis strain SMY methylglyoxal synthase (mgsA) gene
313901	RNAscope® Positive Control Probe - Hs-PPIB - Homo sapiens peptidylprolyl isomerase B (cyclophilin B) (*PPIB*), mRNA
310045	HeLa Positive Control Slides
323180	RNA-Protein Co-Detection Ancillary Kit

Table 2.4: ***Reagents and kits used for RNAscope® assay as instructed by supplier.***

The sections were dried onto slides for two days at room temperature and baked for 60mis at 60°C prior to rehydration. De-waxing and re-hydration had the following steps: 2X 5mins Xylene, 2X 1min 100% alcohol. All steps were performed with slight agitation by moving the slide rack up and down. The slides were left to air dry for 5mins at room temperature. For quenching of endogenous peroxidases, the section was covered with RNAscope®Hydrogen Peroxide (1-2 drops) and incubated for 10mins at room temperature. The quenching solution was washed twice in dH_20 for 1min with agitation. Heat-induced antigen retrieval in a steamer was used to un-mask the target epitope as per manual instruction. Summarised, two slide holders were placed in the water-filled steamer, one with

dH_20 and one with the RNAscope® 1X Target Retrieval Reagent. Once the Target Retrieval Reagent reached 99°C, slides were added for 15mins as recommended by the supplier. Slides were rinsed with distilled water for 15secs and fixed in 100% Ethanol for 3mins. Primary antibody against CD69 for the IHC co-stain was diluted in Co-Detection Antibody diluent (*see section 2*.9) and incubated overnight at 4°C in HybEZ™ Slide Rack. For fixation of the primary antibody to prevent cleavage during protease treatment, slides were warmed up to room temperature for 30mins and washed in PBS-T (PBS 0.01%Tween-20) 3X for 2mins. Subsequently, they were incubated in fresh 10% Neutral Buffered Formalin for a minimum of 30mins at room temperature. The previous washing step in PBS-T was repeated. Slides were fixed in 100% Ethanol for 3mins and air dried. To gain access to the target RNA sequence, protease treatment was required removing proteins surrounding the nucleic acids. Slides were added to HybEZ™ Slide Rack and treated with RNAscope® Protease Plus for 30mins at 40°C in the pre-heated HybEZ™ Oven. Slides were washed 2X for 1min in dH20 on a plate shaker under constant agitation. The probe was hybridised with the *IL1B*-specific or negative/positive control probes (*Table 2.4)* for 2hrs at 40°C in the HybEZ™ Oven. Slides were washed 2X for 2mins in Wash Buffer on a plate shaker under constant agitation. The following steps were used for amplification of the probe:

1. Sections were covered with *AMP1* solution and incubated in HybEZ™ Oven for 30mins at 40°C. Wash Buffer washing step was repeated.

2. Sections were covered *AMP2* solution and incubated in HybEZ™ Oven for 15mins at 40°C. Wash Buffer washing step was repeated.

3. Sections were covered *AMP3* solution and incubated in HybEZ™ Oven for 30mins at 40°C. Wash Buffer washing step was repeated.

4. Sections were covered *AMP4* solution and incubated in HybEZ™ Oven for 15mins at 40°C. Wash Buffer washing step was repeated.

5. Sections were covered *AMP5* solution and incubated at room temperature for 30mins. Wash Buffer washing step was repeated.

6. Sections were covered *AMP6* solution and incubated in room temperature for 15mins. Wash Buffer washing step was repeated.

After amplification, the signal of the probe was detected with the Fast RED solution (1:60 Fast RED-B in Fast RED-A) for 10mins at room temperature. Slides were washed 2X for 1min in dH_20, followed by 2X 2mins in Wash Buffer on a plate shaker under constant agitation. Tissue was blocked with the Co-Detection Blocker for 15mins at 40°C in the HybEZ™ Oven prior to secondary antibody staining to prevent non-specific binding. Slides were washed for 1mins in PBS-T. The ImmPRESS® HRP Horse Anti-Mouse IgG Polymer Detection Kit (MP-7402-15, Vector Laboratories) was diluted 1:5 in Co-Detection Antibody Diluent and incubated with sections for 60mins at room temperature. Slides were washed 2X for 2mins in PBS-T. For the development of the IHC signal, the DAB substrate kit (SK-4100, Vector Laboratories), with 1 drop diluent per 1ml of diluent, was incubated with sections for 5mins. Excess DAB was washed off in dH_20 for 10mins and slides were counterstained for 2mins in 50% Haematoxylin I at room temperature. Slides were washed 2X for 1min with dH_20 and incubated with Scott's Tap Water for 10sec. Slides were washed for 1min with dH_20 and then dried in an oven at 60°C for 30mins. After a 5mins cool-down at room temperature, slides were dipped in Xylene and immediately mounted with VectaMount® Mounting Media for IHC (H-5700-60, Vector Laboratories) and sealed with a coverslip. Signal was quantified using the QuPath software according to the Bio-Techne "MK 51-154 Using QuPath (Bankhead *et al.*, 2017)to analyze RNAscope BaseScope and miRNAscope images" manual.

2.9 RNA isolation from intestinal tissue

Biopsies were stored in 1.5ml of RNA*later*™ Stabilization Solution (Cat. No. AM7020, Thermo Fisher) per sample at 4°C for 7 days and subsequently at -80°C. Samples were homogenised in 350ul in RLT buffer (RNeasy Mini Kit, Cat. No. 74106, Qiagen) using metal beads (2X5mm, 1X7mm) and TissueLyser LT (Cat. No. 85600, Qiagen) for 20mins. After the tissue was homogenised, instructions of the RNeasy Mini Kit instruction (Cat. No. 74106, Qiagen) were followed for RNA extraction. To minimise risk of genomic DNA contamination, an on-column DNAse digest was performed according to RNeasy Mini Kit instructions (Cat. No. 74106, Qiagen).

2.10 RNA to cDNA conversion

RNA was converted to cDNA using the Applied Biosciences High-Capacity RNA-to-cDNA™ Kit (Cat. No. 4387406) following the manual instructions with total of 10ng of RNA per reaction and a total volume of reaction with 20µl. Reaction composition is summarised in *Table 2.5* and reaction steps in *Table 2.6*. cDNA was kept at -20°C for short term and -80°C for long-term storage.

Component	Volume per reaction
10X RT Buffer	2.0 µl
25X dNTP Mix	0.8 µl
10X RT Random Primers	2.0 µl
MultiScribe Reverse Transcriptase	1.0 µl
Nuclease-free dH_20	3.2 µl
Sample RNA	11 µl
Total	*20* µl

Table 2.5: ***Reagents for RNA to cDNA transcription in µl per reaction.***

Settings	Step 1	Step 2	Step 3	Step 4
Temp.	25°C	37°C	85°C	4°C
Time	10min	120mins	5mins	∞

Table 2.6: ***Set-up of thermal cycler for RNA to cDNA transcription.***

2.11 Quantitative polymerase chain reaction (qPCR)

Reaction mixture composition is summarised in *Table 2.7*. 2ng of cDNA were used for each qPCR reaction. The PCR included the following steps: 10mins at 95°C, and 40 cycles of 15secs at 95°C and 1min at 60°C. Melt curves were generated as quality control. Samples were normalised to the large ribosomal protein gene (*RPLPO*) and analysed using the Delta-Delta Cycle Threshold ($2^{-\Delta\Delta}CT$) method. Primer sequences for gene targets were shown in *Table 2.8*.

Component	Volume per reaction
Fast SYBR Green Master Mix	5.0 µl
Target specific forward primer	1.0 µl
Target specific reverse primer	1.0 µl
Nuclease-free dH_20	1.0 µl
Sample cDNA (1ng/ul)	2.0 µl
Total	*10* µl

Table 2.7: ***Reagents for qPCR per reaction in µl.***

Target	FWD Sequence	REV Sequence
CHI3L1	TGA TGT GAC GCT CTA CG	AAT GGC ACT GAC TTG ATG
IL1B	CCT AAA CAG ATG AAG TGC TCC T	TGA AGC CCT TGC TGT AGT G
RPLPO1	TGG TCA TCC AGC AGG TGT TCG A	ACA GAC ACT GGC AAC ATT GCG G

Table 2.8: ***Forward (FWD) and reverse (REV) primer sequences for target genes (5' to 3').***

2.12 Microarray analyses

2.12.1 GSE94648

The normalised count matrix of this dataset was publicly deposited with the NCBI accession number GSE94648 Planell *et al*. (2017). Peripheral blood samples of 48 adult CD patients and 22 adult age and sex-matched non-IBD controls (NCs) were included in this analysis. Whole blood-derived cDNA was hybridised to Affymetrix Gene Chip Human Genome UI33 Plus 2.0 Arrays in two separate batches. Raw data was processed using the Bioconductor tool in R studio and batch effect was corrected as demonstrated by Johnson, Li and Rabinovic (2007). Additionally, this study by Planell *et al*. (2017) recruited 25 UC patients which were not included in my analysis.

2.12.2 GSE126124

The normalised count matrix of this dataset was publicly deposited with the national centre for biotechnology information (NCBI) accession number GSE126124 by Palmer *et al*. (2019). Peripheral blood was collected from 37 paediatric CD patients and 32 paediatric age and sex-matched NCs were included in this analysis. Matched intestinal biopsies were collected from 36 CD patient and 18 NCs. The recruitment age was set between 8-18 years. The PAXgene Blood RNA system was used to isolate RNA from whole blood. Matched intestinal pinch biopsies were immediately fixed in RNA*later*™ Stabilization Solution (Cat. No. AM7020, Thermo Fisher) until RNA extraction. Subsequently, cDNA libraries were generated and hybridised to Affymetrix Human Gene 1.0 T arrays. Intestinal and PB samples were analysed on separate arrays. The Affymetrix Power Tools software was used for pre-processing and normalisation. Detailed workflow was followed as stated by Palmer *et al*. (2019). Initially, *n*= 39 patients and NCs were recruited; however, two patients were initially misdiagnosed, and 7 controls were excluded due to underlying infections or conditions such as coeliac disease. 18 UC and two IBD unidentified (IBDU) patients were not included in my analysis to focus on CD pathogenesis.

2.12.3 GSE16879

The normalised count matrix of this dataset was publicly deposited with the NCBI accession number GSE16879 by Arijs *et al*. (2010). Colonic pinch biopsies were collected from 19 adult CD patients and 6 NCs. CD patients were sampled at baseline of infliximab administration i.e., within a week of the first dose, which was followed up 4 to 6 weeks post-treatment induction. RNeasy Mini Kit (Qiagen) was used for RNA extraction prior to hybridisation to Affymetrix Human Genome U133 Plus 2.0 Array. Raw data were processed using the Bioconductor tool (under R version 2.7.2) and the robust multichip method as published by Irizarry *et al*. (2003) to generate the log2 expression of each probe. 24 colonic UC and ileal 18 CD samples were excluded from my analysis to focus on colonic CD as it was done in the preliminary bulk RNA sequencing data.

2.13 Gene set enrichment and pathway overrepresentation analyses

All gene set enrichment (GSE) and pathway overrepresentation analyses were performed in R studio (R version 4.0.2 and below). Data were visualised using *heatplot()* or *barplot()* function of the *clusterProfiler* package (Wu *et al.*, 2012, 2021). Unless specified Benjamini-Hochberg cross-validation of p-value was performed to generate adjusted p-value as part of the analyses to reduce type I errors (*i.e.*, false positives).

2.13.1 WikiPathway analysis

WikiPathways analysis was used to determine significantly associated biological pathways with gene sets using the manually curated wpid2gene database (Version 20180810, organism = "Homo Sapiens") as a reference. The analysis was performed using the *enricher()* functions together with the *clusterProfiler* package.

2.13.2 Medical Subject Heading (MeSH) analysis

MeSH analysis was run using *clusterProfiler* and *meshes* packages with the *enrichMESH()* function. The manually curated database gene2pubmed by the NCBI was used as the gene reference and MeSH.Hsa.eg.db used to access MeSH classifications. The category Phenomenon and Processes (MeSH Code 'G') was selected.

2.13.3 Molecular signature database (Msigdb) analysis

MsigDB analysis was run using the *clusterProfiler* and *msigdbr* packages together with the *enricher()* function. The species was specified as 'Homo Sapiens' and category as Gene Ontology Biological Processes (MsigDB Code 'C5').

2.13.4 Gene ontology (GO) enrichment analysis

GO enrichment analysis was performed using the *enrichGO()* function of the *clusterProfiler* package using the Ensembl database as a reference. P-value cut-off was set to 0.05 instead of cross-validation.

2.13.5 Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis

KEGG analysis was performed using the *enrichKEGG()* function in the *clusterProfiler* package with homo sapiens (KEGG code 'hsa') as reference database.

2.14 Statistical analyses

2.14.1 Power and sample size calculations

Power and sample size calculations were performed on http://powerandsamplesize.com/ using the two-sample one-sided mean comparison.

2.14.2 T-test and Mann-Whitney analyses

Unpaired T-Test or Mann-Whitney was performed depending on the distribution of the data. Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests as proposed by GraphPad Prism. Statistical analysis was performed in GraphPad Prism (version 9.3.1 and below).

2.14.3 Multiple comparison analysis

Multiple comparison analyses and visualisation were performed in R studio (R version 4.0.2 and below) using the *tidyverse, ggplo2, lattice, car, multcomp tibble, dplyr, didyr, gridExtra* and *scales* packages. Data were tested for normality and equal variance using Shapiro-Wilk (*shapriro.test()* function) and Levene

(*leveneTest()* function) tests. Data were further analysed based on distribution and variance.

1. If data were normally distributed and exhibited equal variance: one-way ANOVA (*aov()* function) followed by Tukey post-test for multiple comparison (*glht()* function with STATUS = "Tukey")

2. If data were normally distributed but did not exhibited equal variance: Welsh ANOVA (*one.way()* function with *var.equal* = *FALSE*) followed by Benjamini-Hochberg post-test for multiple comparison (*pairwise.t.test()* function with *p.adjust.method* = *"BH"*)

3. If data were not normally distributed independent of the variance: Kruskal-Wallis Test (*kruskal.test()* function) followed by Wilcoxon Rank post-test for multiple comparison analysis (*pairwise.wilcox.test()* function)

Multiple comparison analysis was only performed if initial ANOVA/Kruskal-Wallis models were significant. Transcriptomic data were normalised as described in sections prior to analysis *2.4* and *2.5*.

2.15 CIBERSORTx digital cytometry

CIBERSORTx is a digital cytometry tool developed by Newman *et al*. (2019) which calculates an estimated immune cell component per sample. The signature matrices used as reference were the LM22 microarray dataset published by Newman *et al*. (2019) and the Immune Response In Silico (IRIS) dataset described by (Abbas *et al*., 2005). Bulk RNAseq and microarray inputs were formatted as instructed by Newman *et al*. (2019). For bulk RNAseq data the following parameters were used based on the developer's recommendations: enabled bulk mode batch correction (tailored to bulk RNAseq and microarray), disabled quantile normalisation, enabled absolute mode, permutations (cross-validation step) set to 100. For microarray data, the following parameters were used based on the developer's recommendations: enabled bulk mode batch correction (tailored to bulk RNAseq and microarray), enabled quantile normalisation, enabled absolute mode, and permutations (cross-validation step) set to 100. Quality control of the cibersort analysis was in the form of correlation coefficient and root mean square

error, which describes the average of the distance between the correlation model and predicted value. Only samples with a correlation coefficient $\geq$0.5 were considered for further analysis.

2.16 Linear regression

Continuous data were regressed with flow cytometry cell counts using linear subset regression. The analysis was performed using the *lm()* function as part of the *base R* package in R studio (R version 4.0.2 and below).

2.17 Logistic subset regression

Logistic subset regression was used to understand the associations between one dependent variable (e.g. clinical parameters) and multiple independent variables (i.e., flow cytometry outputs and gene signatures). Flow cytometry data were normalised using the Arcsinh transformation prior to regression analysis. All analysis was performed in R studio (R version 4.0.2 and below). Categorical data were modelled using logistic subset regression. Models with the best fit were selected based on the Akaike Information Criterion (AIC) statistic, which determines the model requiring the least independent variables explaining the most variation within the dependent variable. AIC was calculated using the *bestglm()* package following package instructions (rstudio-pubs-static, 2000; Teixeira-Pinto, 2021).

2.18 Correlations

Correlation analysis was performed using R studio (R version 4.0.2 and below) using the *rcorr()* function of the *Hmisc* package. Input data was normalised prior to correlation.

2.19 Data integration

2.19.1 sparse Least Square-Discriminant Analysis (sPLS-DA)

The sPLS-DA was performed in R studio using the *splsda()* function of the *mixOmics* package (Rohart *et al.*, 2017). This analysis allowed the identification of variables (in this case PBMC subpopulations) that have the highest discriminating power between two or more sample groups based on their variation. Subsequently, it clusters samples and variables based on their similarity. The script was provided by Dr Umer Zeeshan Ijaz and followed as previously published (Keating *et al.*, 2021; Mills *et al.*, 2021).

2.19.2 Data Integration Analysis for Biomarker Discovery using Latent Variable Approaches for Omics studies (DIABLO)

The DIABLO analysis was performed in R studio using the *block.plsda()* function of the *mixOmics* package (Rohart *et al.*, 2017). This analysis allowed the identification of variables from multiple datasets (in this case PBMC subpopulations and cytokines) that have the highest discrimination power between two or more sample groups based on their variation. Subsequently, it clustered samples and variables based on their similarity. The script was provided by Dr Umer Zeeshan Ijaz and followed as previously published (Gauchotte-Lindsay *et al.*, 2019; Frau *et al.*, 2021).

2.20 Random forest

The random forest analysis was performed in R studio (R version 4.0.2. and below) using the *randomForest()* function of the *randomForest* package. Random forest analysis can test the prediction power of a signature for clinical phenotypes using machine learning. The dataset is randomly assigned into a training and a validation dataset. After the algorithm is trained on the training dataset, the predictive strength of the model is tested on the validation dataset.

2.21 Receiver operating curve (ROC) analysis

ROC analysis was used to calculate the predictive power of predictor variables (gene expression) for treatment response to anti-TNF. The analysis was performed in R studio using the *roc()* function of the *pROC* package (Robin *et al.*, 2011). "print.auc" command was set to TRUE to calculate the area under curve (AUC) for each predictor variable.

Chapter 3: Characterisation of peripheral blood T cell phenotypes in CD patients

3.1 Introduction

T cells are main drivers of inflammation in IBD. An increase in pro-inflammatory effector T cells and a decrease of homeostatic regulatory T cells significantly contributes to the inflammatory condition. T cell infiltration in the intestines is observed during intestinal inflammation (Smids *et al.*, 2018), and quantification of leukocyte populations in peripheral blood by flow cytometry has shown an overall decrease of T cells in most CD patients compared to healthy controls (Senju *et al.*, 1991; Mack *et al.*, 2002). Blood is regularly sampled from CD patients as part of routine clinical management. Signs of relapse include changes in WBC, lymphocyte and neutrophil counts, and increases in CRP (Auer *et al.*, 1979; Heimann, Bolnick and Aufses, 1986; Neubauer, Woźniak-Stolarska and Krzystek-Korpacka, 2018; Kosoy *et al.*, 2021; Solem *et al.*, 2005). Even though a wide range of peripheral blood markers is available to measure inflammation, their changes are not necessarily representative of the level of intestinal inflammation a patient is experiencing (Mitselos *et al.*, 2018). For example, CRP has been identified as the most reliable PB biomarker of inflammation in the CD (Yazar *et al.*, 2020b) but has been found to only moderately correlate with endoscopic activity as defined by the SES-CD (r= 0.525) (Schoepfer *et al.*, 2010b). It is currently unknown why some patients exhibit signs of systemic inflammation in the form of changes in peripheral blood while others do not. Further research is required to investigate the differences between these patients so that peripheral blood markers, such as lymphocyte count, could become a valuable tool for informed disease monitoring in the future.

Aims:

In this chapter, I aimed to identify differences in the peripheral blood profile of CD patients and link them to clinical parameters, to understand how activity in the periphery may relate to distinct disease subsets.

For this purpose, I quantified PBMCs in CD patients and non-IBD controls and examined the relationship between PBMCs and clinical parameters, including intestinal inflammation status, treatment, and disease phenotype. To study differences in T cell subsets, i.e., Th1, Th2 and Th17, the levels of cytokines were measured in matched plasma samples and compared to T cell expression measured by flow cytometry.

3.2 Patient characteristics

3.2.1 Sample collection

Peripheral blood samples from CD patients were collected from the Glasgow Royal Infirmary Gastroenterology Day Unit run by Ann-Marie Scott and IBD clinic run by Dr Daniel Gaya. Patient and NC characteristics (*Table 3.1*) were collected on the day of sampling or retrospectively by the clinical research fellow Dr Muhammad Imran Khan in accordance with the ethical approval. Two patients were re-diagnosed with UC rather than CD and excluded from analyses. Healthy volunteers were recruited from the Institute of Infection, Immunity and Inflammation as non-IBD controls and sampled by qualified phlebotomists.

3.2.2 Treatment

Many immunosuppressant therapies are clinically available to effectively treat CD, some of which are known to modify the number of specific leukocyte populations in the peripheral blood. I decided to stratify patients based on their current treatment to ensure observed differences in PBMC subset level were due to differences in immune mechanisms and not effects of treatment on leukocyte quantity. Most patients in this cohort received general immunosuppressive treatments (43.75%), an umbrella term for thiopurines, 5-ASAs and methotrexate. Alternative treatments found in this cohort were adalimumab (n=2), a specific biologic agent blocking TNF, and vedolizumab (n=2), a biologic drug targeting leukocyte migration to the intestines. Three patients received combination therapies, and two patients did not receive any medication.

	IBD ***(n= 16)*** CD (*n* = 14) UC (*n* = 2)	**NC** ***(n= 16)***
Mean Age, years (Range)	42.813 (19-75)	38.214 (23-65)
Standard Deviation Age, years	18.325	14.635
F:M ratio	1.00	1.333
Mean HBI (Range) *CD only*	3.656 (1-12)	N/A
Median CRP, mg/L (Range)	3.75 (0-19)	N/A
Median Lymphocyte count, cells/μl (Range)	1.813 (0.6-3.6)	N/A
Median WBC count, cells/μL (Range)	6.525 (3.2-10.5)	N/A
Treatment **Immunosuppressants** *Thiopurines* *5-ASA* *Methotrexate* **Biologics** *Adalimumab* *Vedolizumab* **Combination Therapies** *Thiopurine + Ustekinumab* *5-ASA + Thiopurine* *5-ASA + Thiopurine + Anti-TNF* **No treatment**	 **7 (43.75%)** 3 3 1 **4 (30.77%)** 2 2 **3 (18.75%)** 1 1 1 **2 (12.5%)**	N/A
Location (%) **CD (n=14)** L1: *Ileal* L2: *Colonic* L3: *Ileocolonic* **UC (n=2)** E1: *Proctitis* E2: *Left-sided* Behaviour (%) **CD only (n=14)** B1: *Inflammatory* *+perianal* B2: *Structuring* B3: *Penetrating* Age of Onset (%) **CD only (n=14)** A1: *<16 years* A2: *17-40 years* A3: *>40 years*	 3 (21.43%) 5 (35.71%) 6 (42.86%) 1 (50.0%) 1 (50.0%) 11 (78.57%) 1 (7.14%) 1 (7.14%) 1 (7.14%) 0 8 (57.14%) 6 (42.86%)	N/A
Smoker Status (%) **CD only (n=14)** Smoker Non-smoker Ex-smoker	 4 (25%) 9 (56.25%) 1 (6.25%)	N/A

Table 3.1: ***Summary of cohort information for blood immunophenotyping.***

16 PB IBD patient samples were collected at the Glasgow Royal Infirmary and 16 non-IBD volunteers were recruited at the University of Glasgow. Markers of inflammation (HBI, CRP, lymphocyte count, WBC count), current treatments, Montreal classifications (age of onset, disease behaviour, location) and smoker status were included for IBD patients. N/A: not applicable. CRP: C-reactive protein. 5-ASA: aminosalicylates.

3.2.3 Classification

CD is a very complex pathology and can occur in different locations with different disease behaviours. The Montreal Score is a clinical tool developed to stratify patients based on their disease location, behaviour, and age of onset. Most CD patients exhibited ileocolonic disease (37.5%), describing intestinal inflammation in both the ileum and the colon. Additionally, most patients belonged to the inflammatory CD phenotype category (68.75%) without stricturing of the intestines or inflammation penetrating the mucus layer. 57.14% of CD patients were diagnosed between 17-40 years of age, and 42.86% when they were older than 40 years. Different to CD, UC only affects the colon and rectum. The UC patients in this cohort exhibited ulcerative proctitis (inflammation of the rectum lining, n=1) and left-sided colitis (disease affecting the left side of the colon below the splenic flexure, n=1).

3.2.4 Inflammatory markers

The HBI is a clinical scoring method using symptoms and patient well-being to understand disease activity. It is a simplified version of the more complex CDAI. A previous study examined the CDAI and HBI scores of 1328 CD and found a significant positive correlation between them (Vermeire *et al.*, 2010). Due to limited time in the clinic, it was decided to use the shorter HBI as a measure of disease activity. An HBI $\leq$4 indicates remission or mild disease, and a score of >4 indicates ongoing inflammation. The mean HBI in this cohort was 3.656; however, the values range from 1-12, suggesting a broad spectrum of disease activity in this cohort.

3.3 Identification of PBMC subpopulations

PBMCs were isolated from fresh blood within 4 hours of collection using a density gradient. Subsequently, they were stained for extracellular flow cytometry analysis with a selection of fluorochrome-conjugated surface markers to study different T cell subpopulations. Markers commonly used for NK cell identification were included to provide a comparator non-T cell population for the analyses.

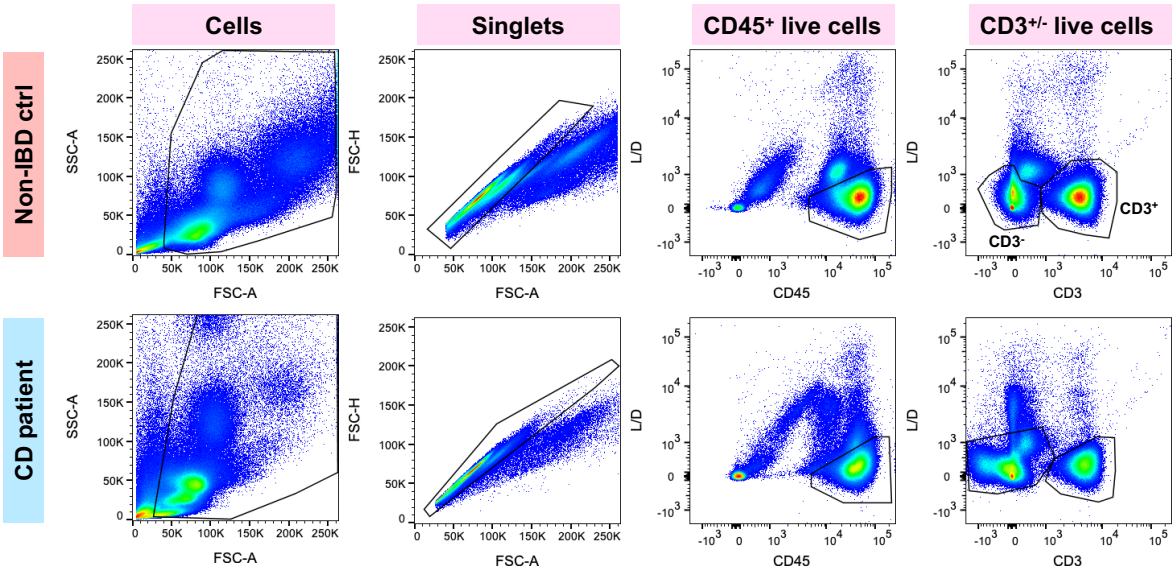


Figure 3.1: ***Gating strategy for PBMC subpopulation selection.***

PBMCs were isolated from fresh blood and stained for extracellular flow cytometry. The top panel shows the staining of PBMCs derived from a non-IBD control (NC), representative of 16 samples. The lower panel shows cells derived from a CD patient, representative of 14 CD patients. Cell debris was excluded using size (FSC-Area vs SSC-Area), followed by a selection of single cells (FSC-Area vs FSC-Height). Live $CD3^+$, $CD3^-$, and $CD45^+$ cells were selected using the surface markers CD3 and CD45 against live/dead-stain (L/D). All analysis was performed using FlowJo™ (Version 10.8).

Figure 3.1 shows the gating strategy applied for PBMC subpopulation selection. Cell debris was excluded based on its small size, using Forward Scatter-Area (FSC-A) and Side Scatter-Area, and single cells were selected by FSC-Height (FSC-H) and SSC-A to avoid contamination of cell clumps in the analysis. Subsequently, a viability dye (L/D) with DNA binding capabilities was used to exclude dead cells. Leukocyte common antigen (CD45), expressed on most hematopoietic cells, was used to quantify white blood cells. T cells were identified within the live $CD3^+$ cell gate (CD3 is a universal T cell marker).

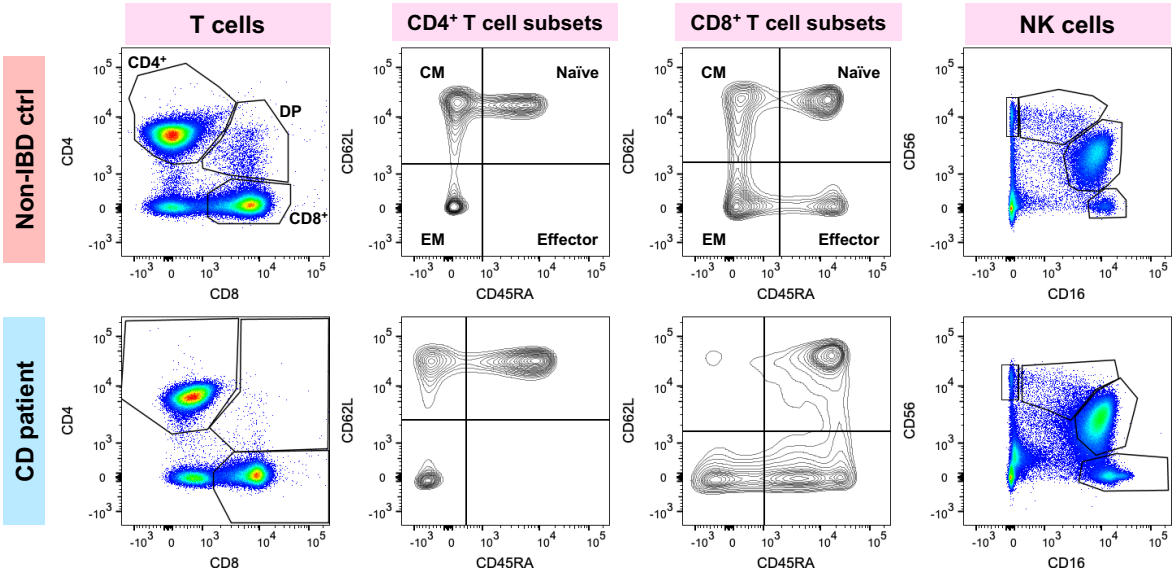


Figure 3.2: ***Gating strategy for T and NK cells subpopulation selection.***

PBMCs were isolated from fresh blood and stained for extracellular flow cytometry. The top panel shows the staining of PBMCs derived from a non-IBD control, representative of 16 samples. The lower panel shows cells derived from a CD patient, representative of 14 CD patients. $CD3^+$ live cells were divided into $CD4^+$, CD8+ and $CD4^+CD8^+$ (DP) T cells and $CD3^-$ live cells into NK cell subsets. T cells were further separated into central memory (CM, $CD45RA^-CD62L^+$), effector memory (EM, $CD45RA^-CD62L^-$), effector ($CD45RA^+CD62L^-$) and naïve T cells ($CD45RA^+CD62L^+$). NK cells were gated off the CD3- live population and subset into $CD56^+CD16^-$ NK, $CD56^{int}CD16^-$ NK, $CD56^-CD16^+$ NK and $CD56^{int}CD16^+$ NK. All analysis was performed using FlowJo™ (Version 10.8).

T cells were separated into $CD4^+$ and $CD8^+$ subsets (*Figure 3.2*). Most samples showed a $CD4^+CD8^+$ population which was included in the analysis as double positive (DP) T cells. The effector function of DP T cells is still controversial, and they have been linked to suppressive and pro-inflammatory, potentially Th2 response-driving characteristics (Eljaafari *et al.*, 2013; Bohner *et al.*, 2019). $CD3^+CD4^-CD8^-$ cells were also found in each sample. These cells are most likely referred to as NKT cells (Hu *et al.*, 2019) but were not quantified since the main focus of the study was to identify T cell-related differences (Sallusto *et al.*, 1999). L-Selectin (CD62L) is required for T cell migration to LNs (Galkina *et al.*, 2003). It is shed rapidly from the surface of activated cells, which is essential for T lymphocytes to exit LNs to carry out effector functions in the periphery. The loss of CD45RA, an isoform of protein tyrosine phosphatase receptor type C, surface expression is an indicator for memory T cells. Together with CD62L it was used to identify central memory (CM, $CD45RA^-CD62L^+$), effector memory (EM, $CD45RA^-CD62L^-$), effector ($CD45RA^+CD62L^-$) and naïve T cells ($CD45RA^+CD62L^+$). NK cells lack CD3. Therefore, their subsets were identified within the live $CD3^-$ gate based on the expression of neural cell adhesion molecule (CD56) and FC-gamma receptor 3A (CD16). These markers divide NK cells into four groups corresponding to

different states of maturity ($CD56^{bright}$) and cytotoxicity ($CD16^{bright}$). $CD56^{bright}$ NK cells have high cytokine-producing capabilities but low cytotoxic activity and are defined as immature (Kucuksezer *et al.*, 2021). They are mainly involved in immune regulation, secreting cytokines such as IFN-γ, TNF-α, or IL-10. In contrast, CD16 is a marker of mature cytotoxic NK cells, ready to clear infected or aberrant cells. Intermediate populations with an intermediate expression of CD16 and CD56, respectively, are transitional stages during the maturation (Perussia, Chen and Loza, 2005; Pesce *et al.*, 2018).

Figure 3.3 displays the IL-2 receptor alpha chain (CD25) expression on $CD4^+$ and $CD8^+$ T cell subsets. CD25 is known to be expressed and maintained on the surface 24hrs post-activation onwards and was used as an additional activation marker for T lymphocytes (Jackson *et al.*, 1990; Ferenczi *et al.*, 2000). In addition to activated T cells, $CD4^+$ T regulatory cells have also known to express CD25 (Hoffmann *et al.*, 2006). The overall percentage of Tregs within the $CD4^+$ T cell pool was approximately 5% (Jonuleit *et al.*, 2001). These data suggest that $CD25^+$ effector and memory T cell populations are most likely not Treg cells and represent activated T cells. However, the naïve population expressing IL-2 receptor alpha chains is a mixture of activated and regulatory T cells.

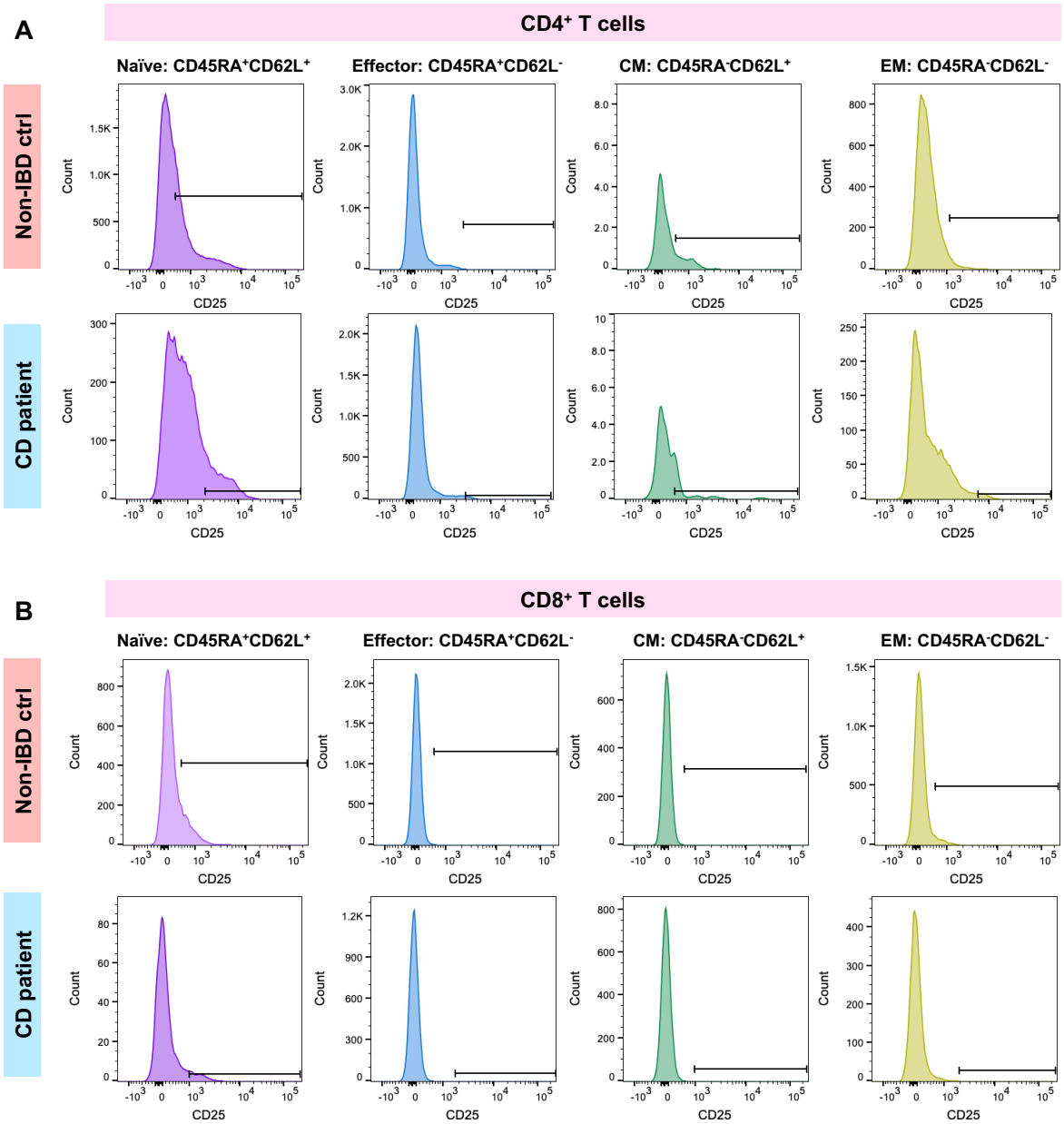


Figure 3.3: ***Gating strategy for T and NK cells subpopulation selection.***

PBMCs were isolated from fresh blood and stained for extracellular flow cytometry. The top panel shows the staining of PBMCs derived from a non-IBD control, representative of 16 samples. The lower panel shows cells derived from a CD patient, representative of 14 CD patients. The activation status of all T cell subpopulations was measured using CD25. A) CD25 expression on the surface of $CD4^+$ T cell subsets. B) CD25 expression on the surface of $CD8^+$ T cell subsets. All analysis was performed using FlowJo™ (Version 10.8).

3.4 Differences in PBMC levels between CD patients and non-IBD controls

3.4.1 Comparison of live cells, singlets, $CD45^+$ and $CD3^-$ cell levels between NC and CD patients

CD-related differences in peripheral blood leukocytes were identified by comparing the levels of PBMC subpopulations of CD patients with non-CD controls. *Figure 3.4* shows the absolute number and frequency of cells, singlets, $CD45^+$ and $CD3^-$ cells. The frequencies of cell count/ml after PBMC isolation were used to calculate the absolute numbers of cells/ml for each population. Frequencies indicate the percentage of the parent population. No significant changes between normal controls and CD patients were detected in the overall cell and singlet populations in either absolute numbers or frequencies. $CD45^+$ cells were selected to study potential differences between total leukocyte levels but showed no significant differences between the two groups. Since the aim of this chapter is to study T cells, $CD3^-$ cells are of interest to identify T cell-independent changes in the PB of CD patients. No significance was also shown between the patients and controls in this subpopulation.

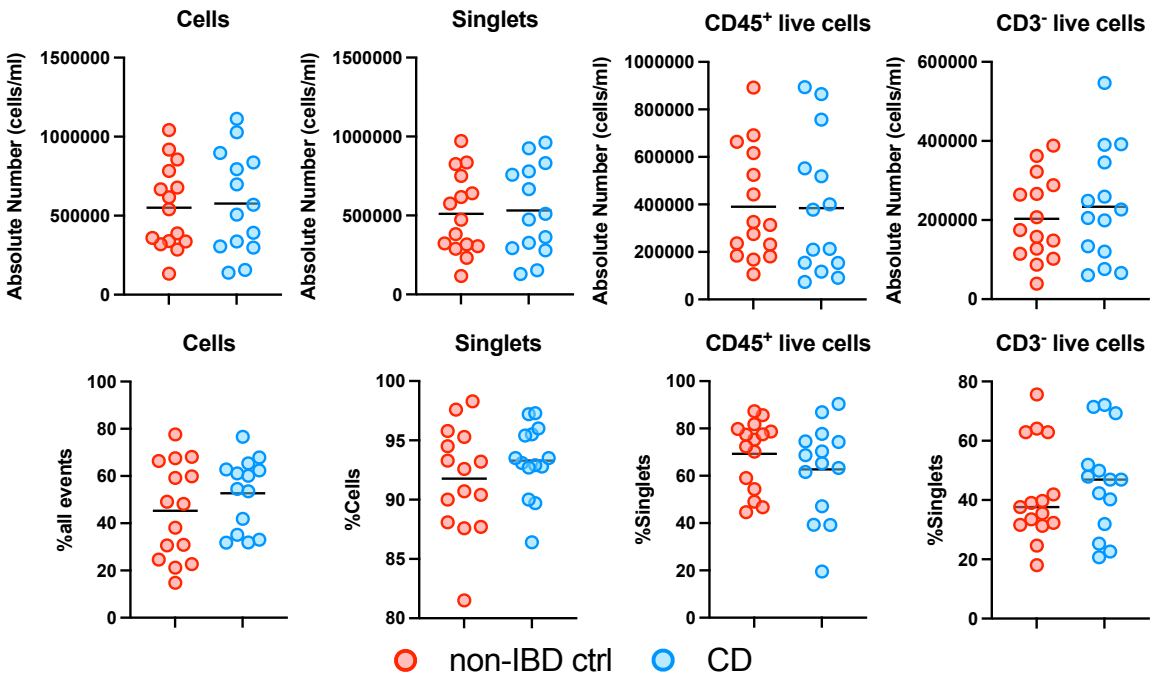


Figure 3.4: ***Comparison of live cells, singlets, CD45⁺ and CD3⁻ cell levels between non-IBD controls and CD patients.***

PBMCs of 14 CD patients and 16 non-IBD controls were analysed using flow cytometry. Absolute numbers (top) and frequencies (bottom) show the levels of cells, singlets, $CD45^+$ and $CD3^-$ live cells. Frequencies are displayed in the percentage of the parent population. Absolute numbers were calculated using frequencies and cell count/ml after PBMC isolation. Lines represent the mean. An unpaired T-Test or Mann-Whitney was performed depending on the distribution of the data. Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. Flow cytometry data were quantified in FlowJo™ (Version 10.8 and below). All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below).

3.4.2 Comparison of T cell levels between non-IBD controls and CD patients

As previously described, CD3 is expressed on T cells as a TCR co-receptor but can also be expressed on NKT cells in PB (Hu *et al.*, 2019). After cell selection based on CD3 expression, this population was further subset into $CD4^+$, $CD8^+$ and DP T cells to distinguish them from other CD3-expressing cell types, such as NKT cells (*Figure 3.5*). Frequencies and absolute numbers of T cells showed similar spread in cell numbers and frequencies. Frequencies of $CD3^+$ cells, $CD8^+$ and DP T cells, and absolute numbers showed no significant changes. Notably, one patient and one control sample had very high levels of DP T cells.

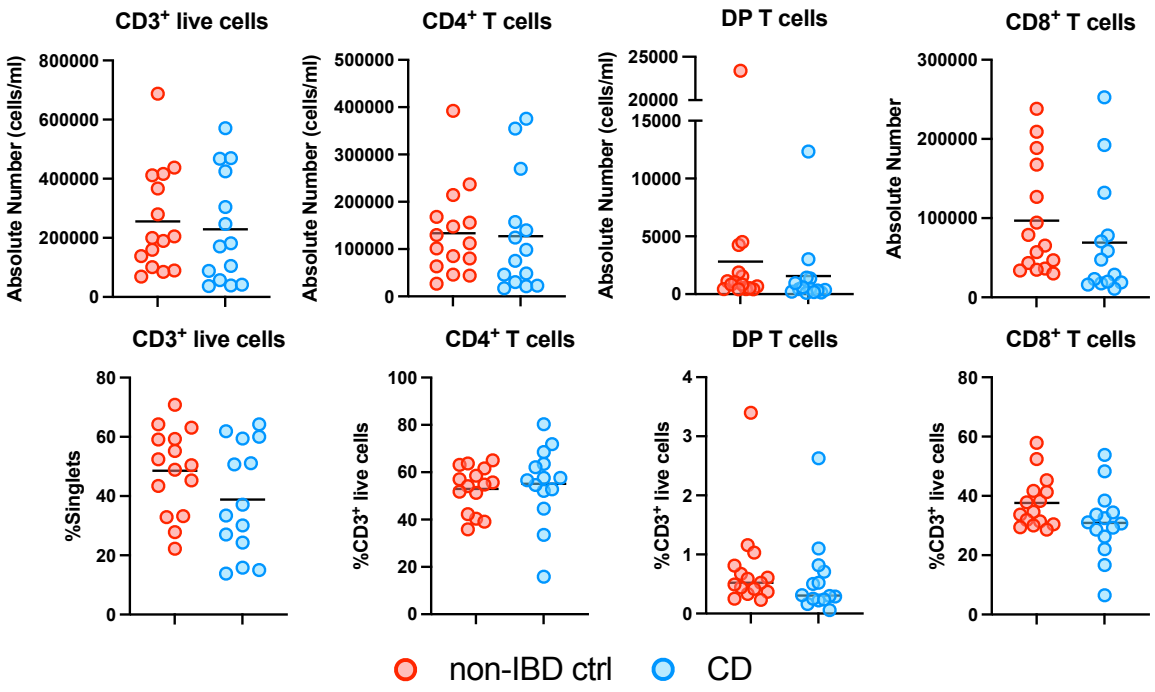


Figure 3.5: ***Comparison of PB T cell levels between non-IBD controls and CD patients.***

PBMCs of 14 CD patients and 16 non-IBD controls were analysed using flow cytometry. Absolute numbers (top) and frequencies (bottom) show the levels of all T cells ($CD3^+$ live cells), double positive (DP) T cells ($CD4^+$ and $CD8^+$), and single positive $CD4^+$ and $CD8^+$T cells. Frequencies are displayed in the percentage parent population. Absolute numbers were calculated using frequencies and cell count/ml after PBMC isolation. Lines represent the mean. An unpaired T-Test or Mann-Whitney was performed depending on the data distribution. Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. Flow cytometry data were quantified in FlowJo™ (Version 10.8 and below). All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below).

$CD4^+$ and $CD8^+$ T cells were divided into naïve, effector, CM and EM cells, as stated in *section 3.3* (*Figure 3.2*). There were no significant differences between IBD patients and non-IBD controls in any $CD4^+$ or $CD8^+$ subpopulations (*Figure 3.6*). Most populations showed a similar variation between the sample groups. Like DP T cells, one patient and one control displayed very high levels of effector $CD4^+$ T cells. The non-IBD control with an increase in effector $CD4^+$ T cells exhibited the same in the effector CD8+ T cell counterpart.

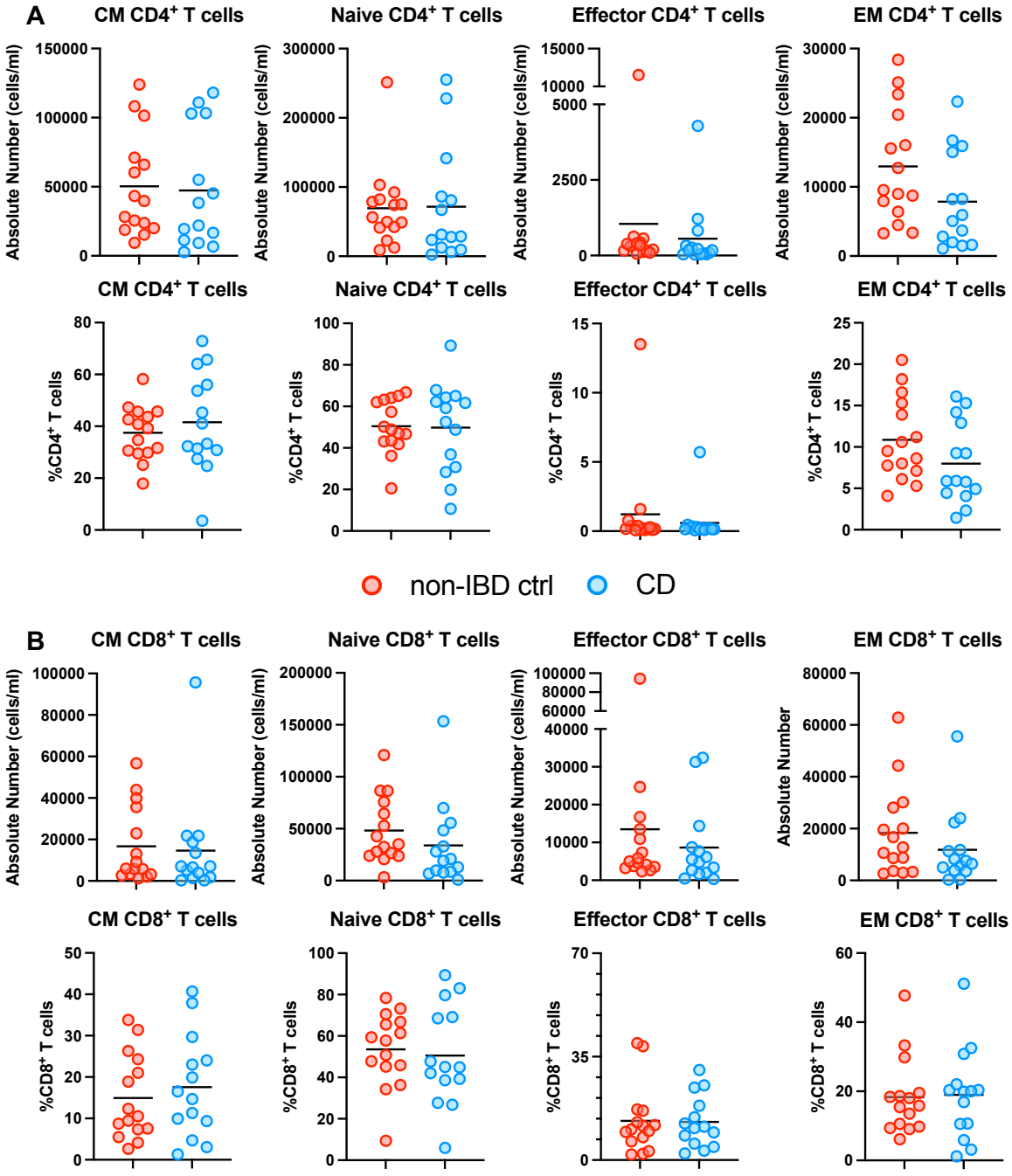


Figure 3.6: ***Comparison of PB CD4⁺ and CD8⁺ T cell levels between non-IBD controls and CD patients.***

PBMCs of 14 CD patients and 16 non-IBD controls were analysed using flow cytometry. A) Absolute numbers (top) and frequencies (bottom) showing the levels of $CD4^+$ subpopulations: naïve T cells ($CD4^+CD45RA^+CD62L^+$), central memory (CM, $CD4^+CD45RA^-CD62L^+$), effector ($CD4^+CD45RA^+CD62L^-$) and effector memory (EM, $CD4^+CD45RA^-CD62L^-$). B) Absolute numbers (top) and frequencies (bottom) showing the levels of $CD8^+$ T cell subpopulations: naïve T cells ($CD8^+CD45RA^+CD62L^+$), central memory (CM, $CD8^+CD45RA^-CD62L^+$), effector ($CD8^+CD45RA^+CD62L^-$) and effector memory (EM, $CD8^+CD45RA^-CD62L^-$). Frequencies are displayed in the percentage parent population. Absolute numbers were calculated using frequencies and cell count/ml after PBMC isolation. Lines represent the mean. An unpaired T-Test or Mann-Whitney was performed depending on the data distribution. Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. Flow cytometry data were quantified in FlowJo™ (Version 10.8 and below). All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below).

I further investigated if patients exhibited a difference in the activation status of T cells compared to non-IBD controls by quantifying the expression of the early activation marker CD25 (*Figure 3.7*). It should be noted that CD25 expression on naïve T cells as defined by CD62L and CD45RA has been shown to co-express FOXP3 and, therefore, most likely identified Tregs (Hoffmann *et al.*, 2006). The differences between the sample groups were non-significant, apart from the frequency of $CD25^+$ CM $CD8^+$ T cells which was significantly increased in CD samples compared to non-IBD controls. This finding was not replicated when the absolute numbers of these cells were compared.

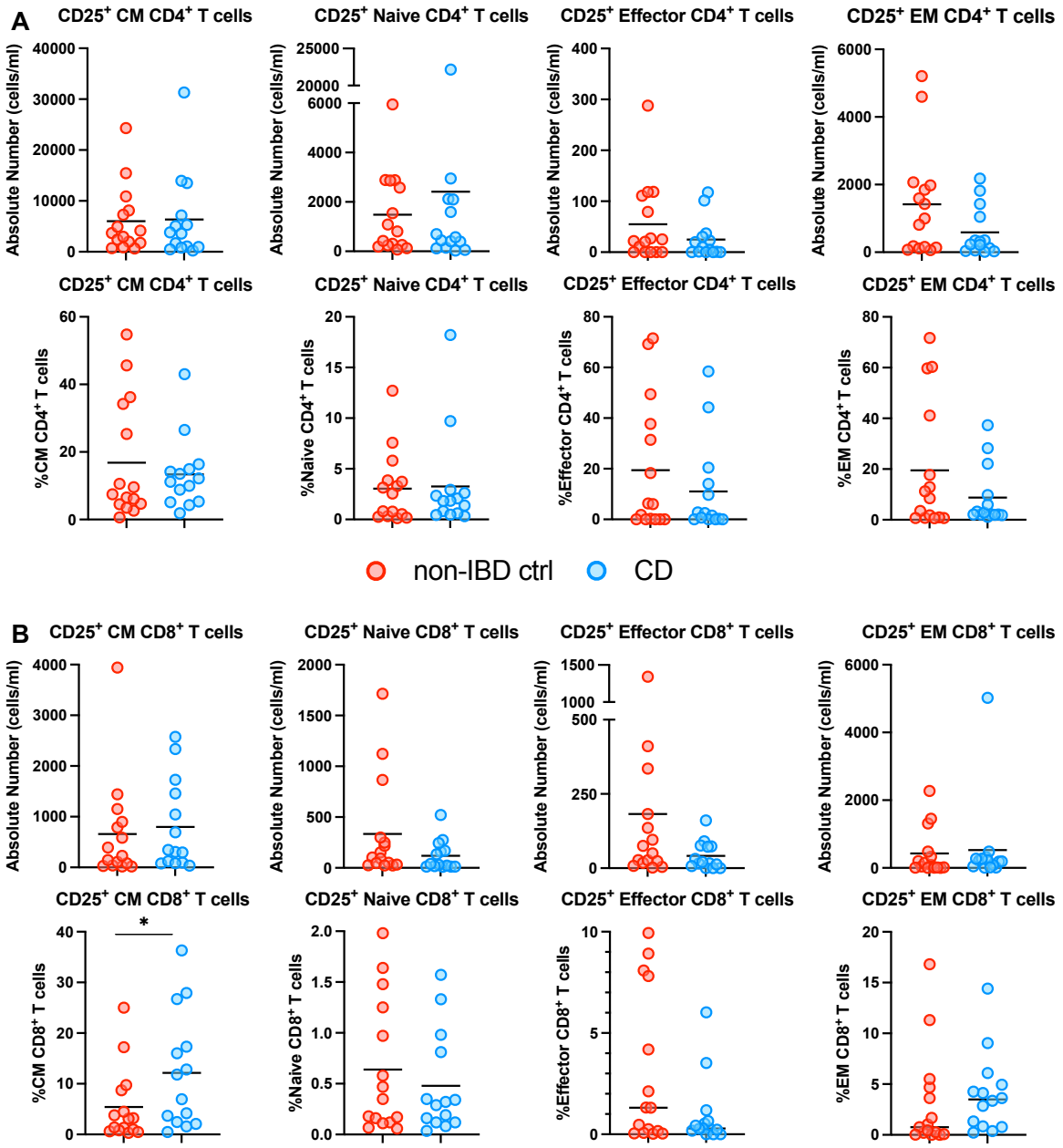


Figure 3.7: ***Comparison of PB CD25⁺ T cells between non-IBD controls and CD patients.***

PBMCs of 14 CD patients and 16 non-IBD controls were analysed using flow cytometry. Absolute numbers (top) and frequencies (bottom) showing the expression levels of IL-2 receptor (CD25) on the surface of A) $CD4^+$ and B) $CD8^+$ T cell subpopulations: naïve T cells ($CD45RA^+CD62L^+$), central memory (CM, $CD45RA^-CD62L^+$), effector ($CD45RA^+CD62L^-$) and effector memory (EM, $CD45RA^-CD62L^-$). Frequencies are displayed in the percentage parent population. Absolute numbers were calculated using frequencies and cell count/ml after PBMC isolation. Lines represent the mean. An unpaired T-Test or Mann-Whitney was performed depending on the data distribution. Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. Flow cytometry data were quantified in FlowJo™ (Version 10.8 and below). All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below). *$p<0.05$.

Due to the small sample size in this study, I performed a power calculation based on the PBMC subpopulation, with the only significant difference between the two sample groups activated CM $CD8^+$ T cells. The power was calculated using type I error rate (a)=0.95 and sample size ratio=1.071 (15/14). Means were 12.60 for CD and 5.41 for non-IBD controls, and the respective standard deviations were 7.138

and 11.40, giving a power of 0.6362. To reach a power of 0.8, a required sample size of at least 25 CD patients would be needed with a sample ratio of 1:1 (CD: NC).

3.4.3 Comparison of NK cell levels between non-IBD controls and IBD patients

To distinguish T cell-specific differences between CD patients and non-IBD controls from overall changes in leukocyte levels, I also quantified NK cells (*Figure 3.8)*. NK cells can be classified based on their cytotoxicity and maturity, as described in *section 3.3*. Statistical analysis did not show significant differences in any NK cell subsets between non-IBD controls and CD patients.

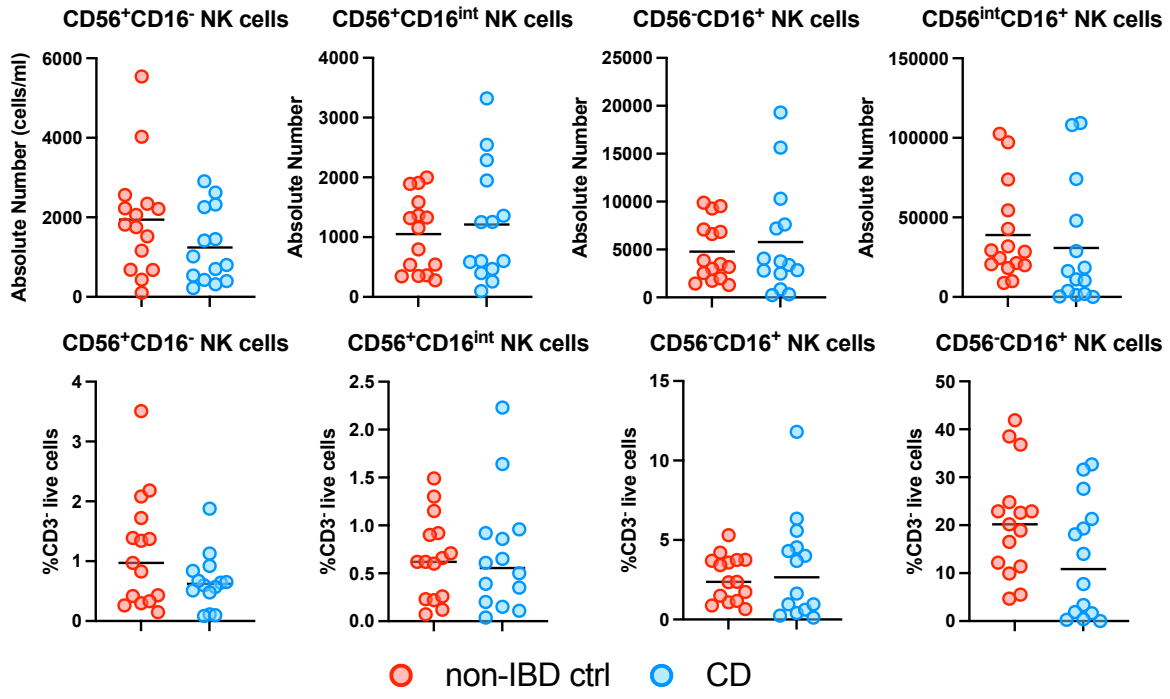


Figure 3.8: ***Comparison of PB NK cell levels between non-IBD controls and CD patients.***

PBMCs of 14 CD patients and 16 non-IBD controls were analysed using flow cytometry. Absolute numbers (top) and frequencies (bottom) are the levels of four NK cell subpopulations based on their CD56 and CD16 expression. Frequencies are displayed in the percentage parent population. Absolute numbers were calculated using frequencies and cell count/ml after PBMC isolation. Lines represent the mean. An unpaired T-Test or Mann-Whitney was performed depending on the distribution of the data. Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. Flow cytometry data were quantified in FlowJo™ (Version 10.8 and below). All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below).

3.4.4 Summary of differences in PB T and NK cells between CD and non-IBD controls

In summary, there are no significant changes in the absolute numbers of PBMC subpopulations between non-IBD controls and CD patients, and one significant increase in the frequency but not an absolute number of activated CM $CD8^+$ T cells in CD.

3.5 Associations between PBMCs levels and clinical data

I aimed to stratify patients into subsets based on their clinical characteristics, such as inflammation status, treatment history, disease phenotype and smoker status. Even though there was only one difference between non-IBD controls and CD patients in PBMC subpopulations, this stratified analysis may elucidate whether specific subsets of patients within the CD cohort show differences in their PBMC profile associated with their clinical phenotypes.

3.5.1 Multiple comparison analyses

Multiple comparison analysis was performed in *R studio*. Data were tested for normal distribution using the Shapiro-Wilk Test and equal variance using the Levene-test and were further analysed using the following conditions:

1. Data with normal distribution and equal variance were analysed using pairwise ANOVA and Tukey's HSD post-test.

2. Data with normal distribution and unequal variances were analysed using a Welsh ANOVA and Benjamini-Hochberg post-test.

3. Data without normal distribution were analysed using a Kruskal-Wallis and Wilcoxon rank sum post-test.

All analysis was performed using absolute cell numbers to ensure the results were not artefacts of measuring proportions. It should be noted that further

stratification of the patient cohort resulted in some sample groups with *n* <3. Therefore, only the models with sample groups of *n* >3 were included in the analysis. Data for the analyses of $CD25^+$ T cell populations were not shown due to the lack of a marker distinguishing activated from regulatory cells. However, it should be stated that there were no statistical differences in any of the CD subsets in terms of $CD25^+$ T cell populations.

3.5.1.1 Multiple comparison analysis of CD subsets based on their inflammation status defined by HBI

PBMC subpopulation levels were compared between non-IBD controls and CD patients with active (HBI >4) and inactive disease (HBI ≤4, *Figure 3.9*) to understand whether inflammation status influences the peripheral blood profile. No significant differences were found between the control group and the patient subsets.

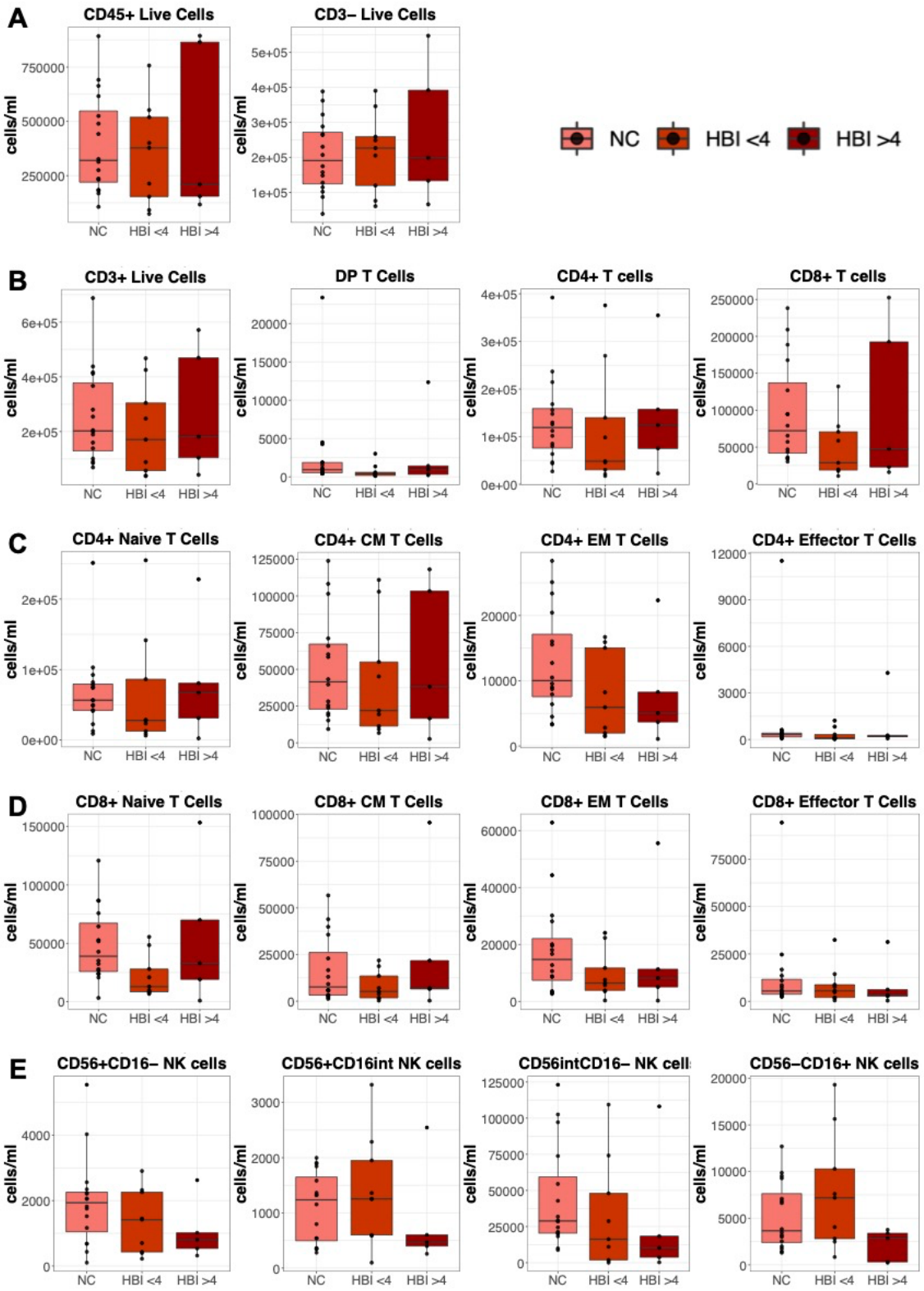


Figure 3.9: ***Comparison of absolute numbers of PBMC subpopulations non-IBD controls (NC) with CD patients categorised based on their HBI score.***

PBMCs of 14 CD patients and 16 non-IBD controls were analysed using flow cytometry. CD patients were categorised based on their HBI into ≤4 (inactive disease) and >4 (active disease). Flow cytometry data were quantified in FlowJo™ (Version 10.8 and below). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).

3.5.1.2 Multiple comparison analysis of CD subsets based on previous steroid exposure

Many treatment groups within the CD patient cohort had an *n* of less than three, which did not allow for statistical analysis. To get some indication of the effects of treatment on PB T cell levels, patients were stratified based on previous steroid exposure. Steroids are known to alter PB T cell levels, a potentially long-term effect. To rule out that any changes observed are due to previous steroid exposure, patients were stratified based on whether they have received steroid treatment in the past or not (*Figure 3.10*). Patients with steroid treatment history did not have significantly altered T or NK cell levels.

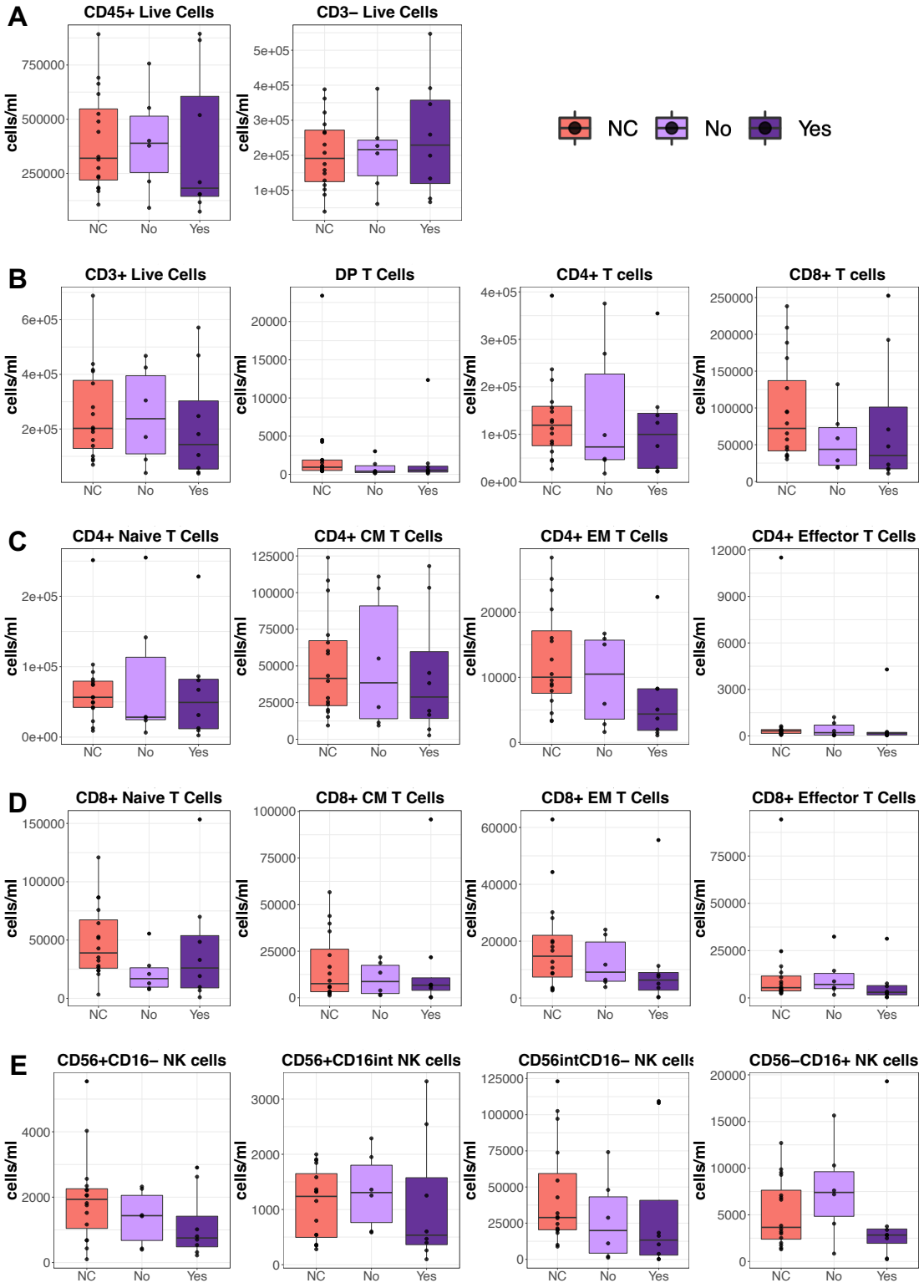


Figure 3.10: ***Comparison of absolute numbers of PBMC subpopulations non-IBD controls (NC) with CD patients categorised based on previous steroid exposure.***

PBMCs of 14 CD patients and 16 non-IBD controls were analysed using flow cytometry. CD patients were categorised based on the previous present (Yes) or absent (No) steroid exposure. Flow cytometry data were quantified in FlowJo™ (Version 10.8 and below). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).

3.5.1.3 Multiple comparison of CD subsets based on their Montreal classifications

I next investigated the link between absolute numbers of PBMC subpopulations and CD disease heterogeneity, as defined by the Montreal classification. CD is a very heterogeneous disease and vastly differs in the age of disease onset, disease behaviour and location. The age of onset is classified as A1 (<17 years), A2 (17-40 years) and A3 (>40 years). Patients diagnosed in category A2 differed from the non-IBD control and patients in category A3 (*Figure 3.11*). They displayed fewer $CD56^{-}CD16^{+}$ NK cells than A3 patients ($p = 0.024$) and fewer $CD4^{+}$ EM T cells ($p = 0.032$) and DP ($p = 0.026$) T cells than non-IBD controls. Patients in category A3 did not differ from the control group. The disease phenotype of CD can be solely inflammatory (category B1) or cause stricturing of the bowel (category B2) and/or inflammation that penetrates the intestinal mucosal surface (category B3). The involvement of the perianal area is indicated by +p and can occur in all three disease behaviours. However, due to an $n < 3$ for categories B1 and B3, multiple comparison analyses were not performed on disease behaviour. Finally, CD patients were categorised based on their disease location into ileal (category L1), colonic (category L2) and ileocolonic (category L3) groups (*Figure 3.12*). No differences in T and NK cell populations were found based on disease location.

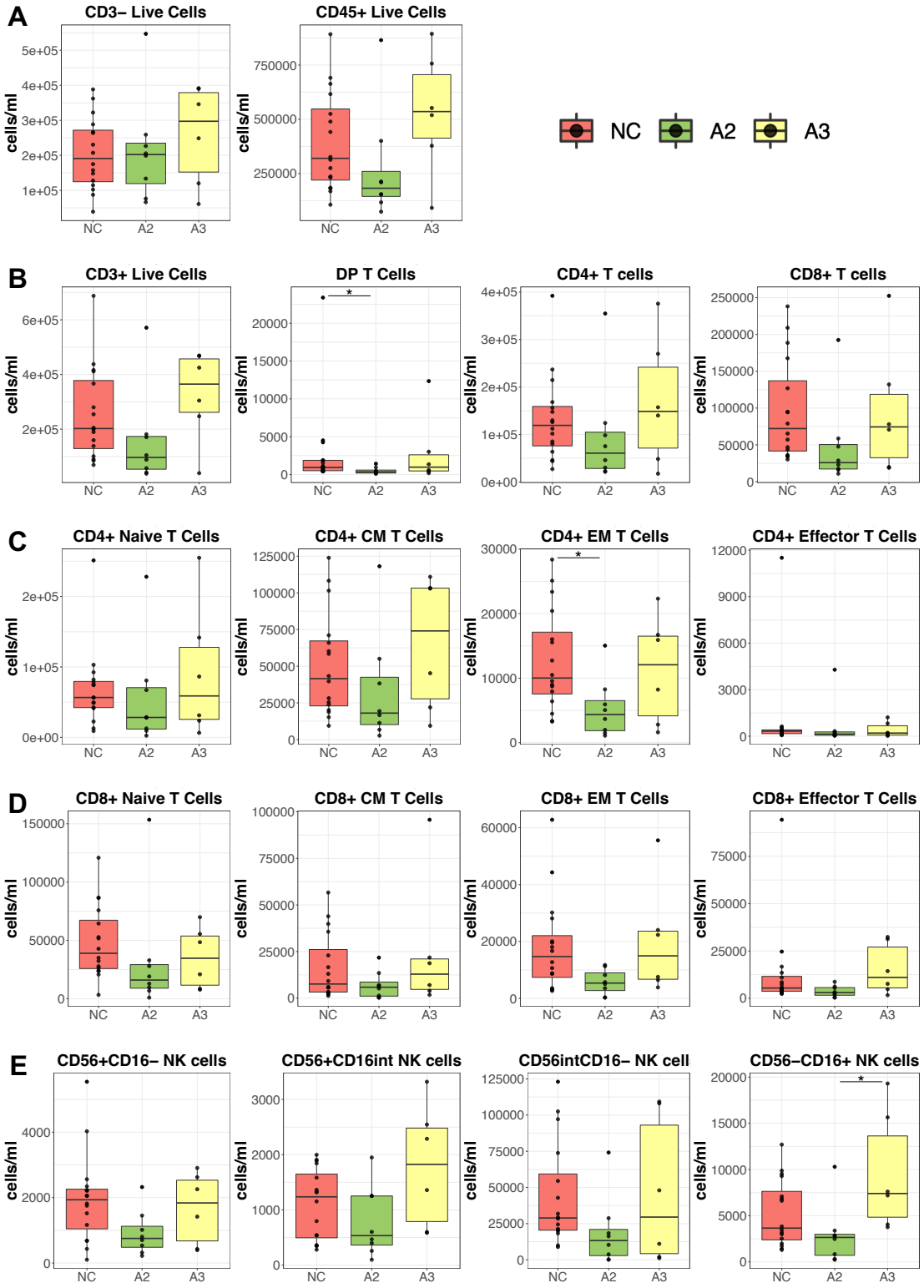


Figure 3.11: ***Comparison of absolute numbers of PBMC subpopulations non-IBD controls (NC) with CD patients categorised based on their age of disease onset.***

PBMCs of 14 CD patients and 16 non-IBD controls were analysed using flow cytometry. CD patients were categorised based on the age of onset of their disease into categories A2 (17-40 years of age) and A3 (>40 years of age). Flow cytometry data were quantified in FlowJo™ (Version 10.8 and below). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below). *p<0.05.

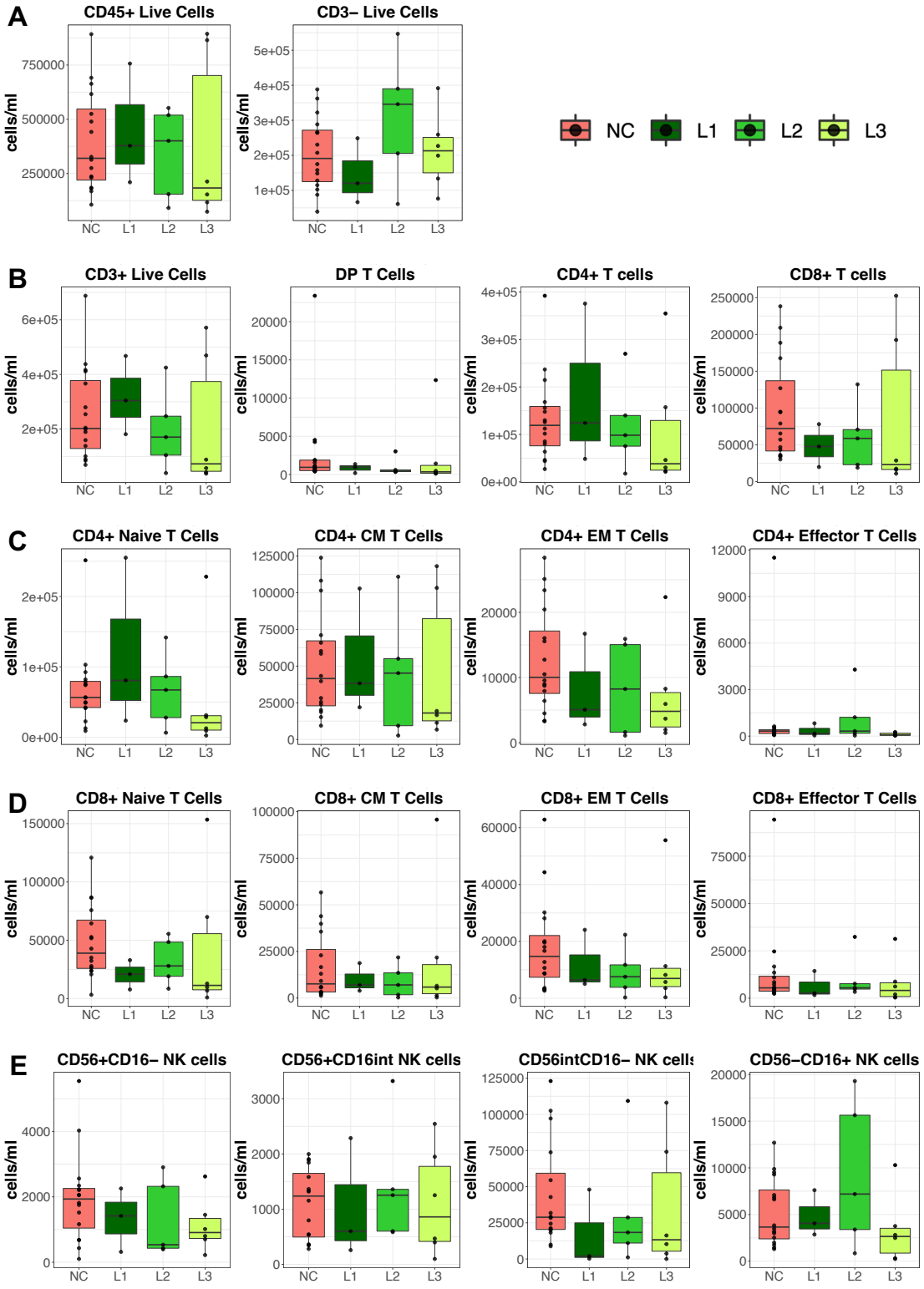


Figure 3.12: ***Comparison of absolute numbers of PBMC subpopulations non-IBD controls (NC) with CD patients categorised based on their disease location.***

PBMCs of 14 CD patients and 16 non-IBD controls were analysed using flow cytometry. CD patients were categorised based on the location of their disease into categories L1 (ileal), L2 (colonic) and L3 (ileocolonic). Flow cytometry data were quantified in FlowJo™ (Version 10.8 and below). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).

3.5.2 Summary of multiple comparison analyses

Stratification of patients based on disease activity, location and previous steroid exposure did not elucidate any differences between the CD subsets. Patients with a younger age of onset (category A2) exhibited a decrease in NK cells with high cytotoxicity compared to patients that presented with disease >40 years of age (category A3) and fewer $CD4^+$ EM T and DP T cells than the non-IBD controls.

3.5.3 Logistic subset regression models

Stratification based on treatment and disease behaviour resulted in sample groups with n <3 and were therefore excluded from multiple comparison analyses. I used subset logistic regression modelling to test potential links between these clinical variables and PBMC populations. Logistic subset regression modelling is a valuable tool to test the strength and type of relationship between multiple X (or independent variables) with a single Y (or dependent variable). In this case, the independent variables were the absolute counts of 29 PBMC subpopulations which were compared to a single dependent variable, current treatment, or disease behaviour. To ensure the most robust model with the best fit and lowest error rate was selected, logistic subset regression models were evaluated using the Akaike Information Criterion (AIC) statistic. The AIC determines the model requiring the least independent variables explaining the most variation within the dependent variable.

Logistic regression models with the lowest AIC were selected using the *bestglm()* function in *R*. Due to limitations of this function, only up to n =15 PBMC subpopulations could be modelled simultaneously. To overcome this issue, I ran three different analyses per dependent variable: T cell populations (naïve, effector, CM and EM $CD4^+$ and $CD8^+$ T cells), $CD25^+$ T cell populations ($CD25^+$ naïve, effector, CM and EM $CD4^+$ and CD8+ T cells) and NK cells ($CD56^+CD16^-$, $CD56^+CD16^{int}$, $CD56^-CD16^+$ and $CD56^-CD16^{int}$ NK cells). The logistic regression outputs did not show any significant associations between current treatment status or disease behaviour (representative data shown in *Table 3.2*).

	Model variables *dependent variable* ~ independent variable(s)	Intercept	AIC	p
5-ASA				
T Cell Populations	*5-ASA* ~ $CD4^+CD45RA^+CD62L^+$ T cells	-3963	4	0.998
NK Cell Populations	*5-ASA* ~ $CD56^-CD16^+$ NK cells + $CD56^{int}CD16^+$ NK cells + $CD56^+CD16^{int}$ NK cells	-677.11	8	0.999
Thiopurines				
T Cell Populations	*Thiopurines* ~ Live $CD3^+$cells + $CD4^+CD45RA^+CD62L^+$ T cells + DP T cells	2203.5	8	0.998
NK Cell Populations	*Thiopurines* ~ $CD56^{int}CD16^-$ NK cells + $CD56^+CD16^{int}$ NK cells	-1989.8	6	0.999
MTX				
T Cell Populations	*MTX* ~ $CD8^+$ T cells	955.0	4	0.999
NK Cell Populations	*MTX* ~ $CD56^+CD16^-$ NK cells + $CD56^{int}CD16^+$ NK cells + $CD56^+CD16^{int}$ NK cells	1820	8	0.996
Adalimumab				
T Cell Populations	*Adalimumab* ~ $CD4^+CD45RA^+CD62L^+$ T cells + $CD4^+CD45RA^+CD62L^-$ T cells	-3799.6	8	0.998
NK Cell Populations	*Adalimumab* ~ $CD56^{int}CD16^-$ NK cells	-1141.3	4	0.999
Vedolizumab				
T Cell Populations	*Vedolizumab* ~ $CD4^+CD45RA^+CD62L^-$ T cells	-241.57	4	0.999
NK Cell Populations	*Vedolizumab* ~ $CD56^+CD16^{int}$ NK cells + $CD56^{int}CD16^-$ NK cells	-216	6	0.999
Combination				
T Cell Populations	*Combination* ~ $CD4^+CD45RA^-CD62L^+$ T cells + $CD4^+CD45RA^-CD62L^-$ T cells	-95.96	6	0.999
NK Cell Populations	*Combination* ~ $CD56^+CD16^{int}$ NK cells + $CD56^+CD16^-$ NK cells	-18.36	6	0.999

Table 3.2: ***Results data of logistic regression models showing the relationship between PBMCs numbers and treatment groups.***

The dependent variable (y) was defined as current treatment groups: 5-ASA (aminosalicylates), thiopurines, MTX (methotrexate), adalimumab (anti-TNF), vedolizumab (anti-α4β7) and combination therapy (characterised in *Table 3.1*). Independent variables (x) were defined as PBMC subpopulations. Models with the lowest Akaike Information Criterion (AIC) for T and NK cell populations were selected using bestglm() function. The intercept refers to the model estimate. All statistical analysis was performed in R studio (R version 4.0.2 and below).

3.5.4 Linear regression models

In addition to HBI, total WBC, lymphocyte, and neutrophil count, as well as CRP levels, were measured in the blood of CD patients in an independent clinical laboratory on the same day as the PBMC isolation. Since there were no associations between HBI and any PBMC subpopulations, I performed linear regression analyses to test whether any of these additional markers showed any correlations with the absolute cell counts. I performed the analyses in R studio using the *lm()* function

from the R base package. *Figure 3.13* summarises the level of significance (p-value) and correlation (R^2-value). All correlations between the inflammatory markers and PBMC populations were positive unless stated otherwise.

Total WBC count had significant correlations with four PBMC subpopulations. The association with the highest significance was between total WBC and $CD56^-CD16^+$ NK cells ($p < 0.0001$). Two additional NK cell populations correlated with total WBC: the $CD56^+CD16^{int}$ ($p = 0.0024$) and $CD56^+CD16^-$ ($p = 0.0301$) phenotypes. $CD8^+$ T cell population correlated with WBC count as the only T cell subset ($p = 0.0275$). Based on these findings, an increase in cytokine-producing and cytotoxic NK cells was linked to an increase in overall cell count in the blood of CD patients together with $CD8^+$ T cells. However, it should be noted that the WBC count did not associate with a particular $CD8^+$ T cell subset. The WBC counts did not correlate with the live $CD45^+$ population, which would have been expected since CD45 is expressed on all WBCs.

CRP correlated with 9 PBMC subpopulations. It had the strongest correlation with $CD45^+$ ($p < 0.0001$) and $CD3^-$ cells ($p < 0.0001$). Interestingly both CM $CD4^+$ ($p = 0.0005$) and CM $CD8^+$ T cells ($p < 0.0001$) were associated with CRP. Other T cell populations that significantly correlated with CRP were effector $CD8^+$ T cells ($p = 0.0096$), overall and naïve $CD4^+$ T cells ($p = 0.0158$) as well as DP T cells ($p < 0.0001$). $CD56^+CD16^-$ NK cells also showed a significant correlation with CRP ($p = 0.0333$). Therefore, CRP levels appear to increase with the levels of T cells, CM T cells in particular, and cytokine-producing NK cells ($CD56^+CD16^-$) in PB.

Lymphocyte count correlated significantly with 5 out of the 9 population that were also linked to CRP: CM $CD8^+$ T cells ($p = 0.0024$), naïve $CD4^+$ T cells ($p = 0.0055$), $CD4^+$ T cells ($p = 0.0291$), effector $CD8^+$ T cells ($p = 0.0467$), and live $CD3^-$ cells ($p = 0.0474$). Most of these populations are T cells suggesting that an increase in T cell subpopulations causes an increase in lymphocyte count, which was expected. The positive correlation between lymphocyte count and live $CD3^-$ cells could refer to an association with $CD3^-$ B lymphocytes, which were not quantified in this dataset, and lymphocyte count.

Neutrophil count was significantly associated with 10 PBMC populations. It correlated with live $CD45^+$ ($p = 0.0022$) and $CD3^-$ ($p = 0.0003$) cells, the effector and

CM phenotypes of both, $CD4^+$ (p_{CM} =0.0136, $p_{Effector}$ =0.0287) and $CD8^+$ (p_{CM} <0.0001, $p_{Effector}$ =0.0041) T cells, overall $CD4^+$ (p =0.0081) and naïve $CD4^+$ T cells (p =0.0007), DP T cells (p =0.0018) as well as $CD56^+CD16^-$ (p =0.0252). Interestingly, lymphocyte count shared all and CRP all but one significant correlation (apart from effector $CD4^+$ T cells) with the neutrophil count. The overall increase in neutrophils in PB correlated with an increase in 7 out of 11T cell populations but also with live $CD3^-$ cells suggesting an association with non-T cell PBMC phenotypes. In addition to $CD56^+CD16^-$ NK cells, this could relate to PBMC subsets not quantified in this study, such as B cells or monocytes.

Finally, I also ran the regression models against HBI as the dependent variable to see whether the results reflected the multiple comparison analysis. As expected, there were no significant correlations between HBI and any PBMC subpopulations.

3.5.5 Summary regression models

The neutrophil count showed the most significant associations with the PBMC subpopulations, followed by CRP levels and lymphocyte count. All three markers exhibited a similar correlation pattern dominated by positive regression with T cell subpopulations. In contrast, WBC count correlated with 3 out of the 4 NK cell subsets and only one T cell population.

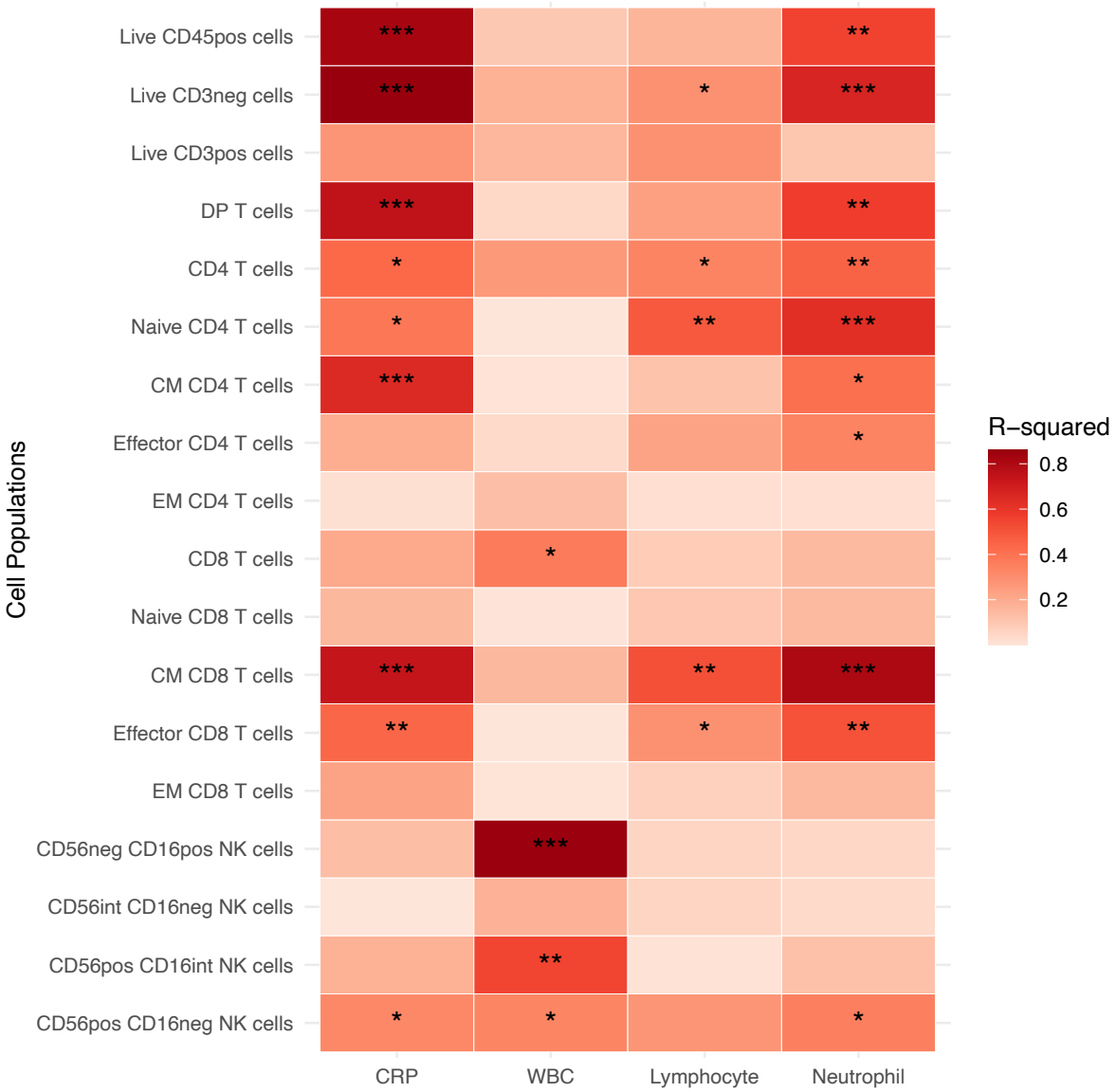


Figure 3.13: ***Linear regression model predicting association between PBMC populations and peripheral blood inflammatory markers.***

Dependent variables (y) are total WBC, lymphocyte count, neutrophil count or CRP. Independent variables (x) were defined as PBMC subpopulations. Models with the lowest cross-validation error (leave-one-out-cross validation). Negative regression is displayed in blue, and positive regression is shown in red. Gradient displays the R2- value. All statistical analysis was performed in R studio (R version 4.0.2 and below). *p<0.05, **p<0.01, ***p<0.001.

3.6 The difference in peripheral blood cytokine levels between CD and NC

Cytokine profiling is a valuable method to study inflammatory status and the type of inflammation. I, therefore, isolated plasma from peripheral blood samples on the same day as the PBMCs (n_{CD}=14, n_{NC}=14). Two non-IBD controls did not provide enough fresh blood to isolate PBMC and plasma; therefore, this analysis only includes 14 NC samples. Using the magnetic bead-based multiplex ELISA Luminex technology 12 cytokine targets were quantified, including IL-1ß, IL-6, CCL2, IL-21, IFN-γ, IL-10, IL-23, TNF-α, IL-17/17A, IL-12p70 subunit, IL-31 and IL-4. The pro-inflammatory cytokines IL-1ß, IL-6 and TNF-α, are upregulated during a wide selection of immune responses. They were included in the multiplex because of their prominent role in IBD inflammation (Muzes *et al.*, 2012). IL-17/17A and IL23 are associated with Th17 cells, immune cells that drive inflammation in both CD and UC. To identify a Th1 phenotype, IFN-γ was included. The IL12p40 subunit is part of IL-12 and IL-23, two cytokines driving Th1 and Th17 development, as well as IFN-γ secretion. IL-4, IL-21 and IL-31 represented Th2 immunity. The immunoregulatory cytokine IL-10 was part of this panel to identify patients with a potential ongoing anti-inflammatory response. CCL2 is an inflammation-associated chemokine that acts as a chemoattractant predominantly for monocytes and macrophages (Deshmane *et al.*, 2009) but has also been shown to attract lymphocytes (Carr *et al.*, 1994).

3.6.1 Comparison of cytokine levels between non-IBD controls and CD patients

Figure 3.14 shows the differences in cytokine profile between non-IBD controls and CD patients. Only seven targets (IL-6, IFN-γ, TNF-α, IL-21, IL-23 and CCL2 and IL-10) were detectable in the multiplex analysis and were included in the following analyses. To ensure that experimental error was not the reason for the low protein measurements, a standard curve, and positive and negative control were run alongside the samples as instructed by the supplier and showed the expected signals. Overall, PB cytokine expression was deficient in both IBD patients and non-IBD controls, with the exception of a few individuals. No significant differences were found between the two sample groups. Interestingly, the one

patient and one control who displayed high IFN-γ expression also exhibited high levels of IL-21. Additionally, one of these patients (1617_19) had proportionally very high expression of IFN-γ, IL-21, IL-6, and TNF-α. Patient 2448_19 expressed over 4-fold more IL-10 than any other patient or non-IBD control; similarly, patient 1930_19 had approximately 3-fold more IL-23 than the rest of the cohort.

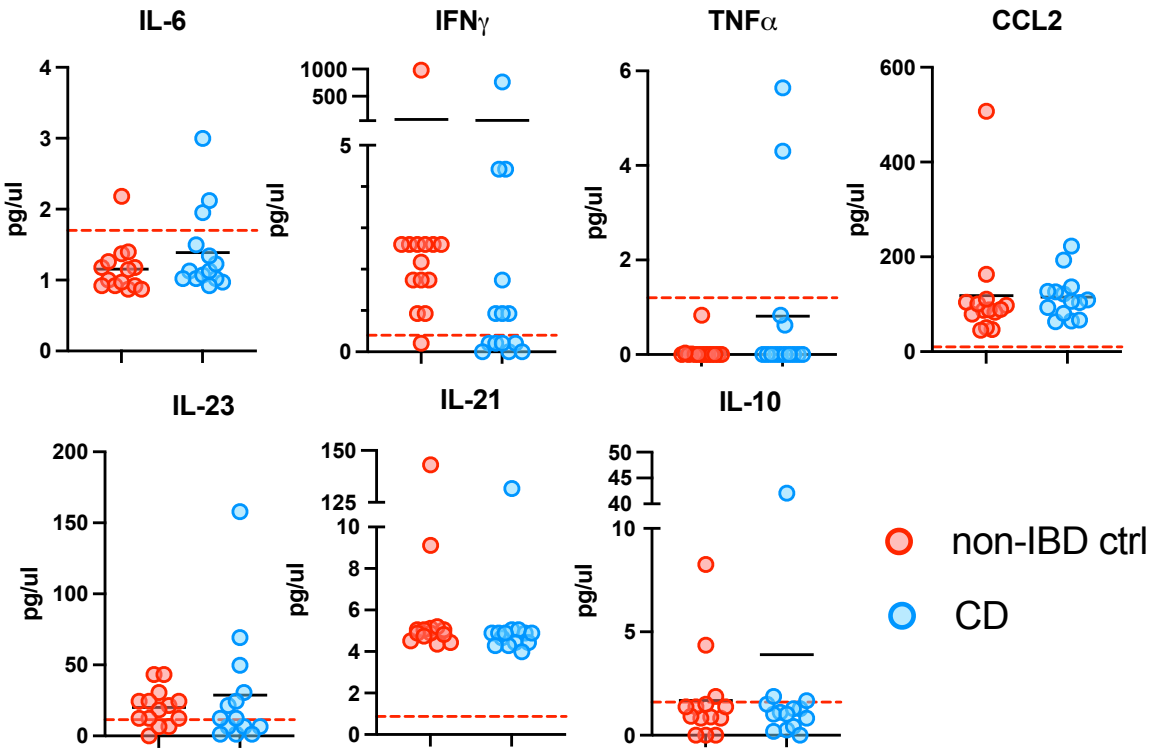


Figure 3.14: ***Comparison of peripheral blood cytokine levels between non-IBD controls and IBD patients.***

The plasma of 14 CD patients and 14 non-IBD controls were analysed using a 12-panel multiplex cytokine assay. The targets of the assay included: IL-1ß, IL-6, CCL2, IL-21, IFN-γ, IL-10, IL-23, TNF α, IL-17/17A, IL-12p70 subunit, IL-31 and IL-4. Only cytokines with detectable levels are included in the analysis. Concentrations were extracted from standard curves. Black lines represent the mean. Dotted red lines represent the sensitivity cut-off as determined by the supplier. An unpaired T-Test or Man-Whitney was performed depending on the distribution of the data. Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below).

3.6.2 Stratification of CD patient PB cytokine levels based on their disease activity, current treatment, and Montreal classification

CD patients were stratified based on their inflammatory status, treatment, and disease phenotype for multiple comparison analysis, as previously shown for PBCMs in *3.5.1*. Only one significant difference was recorded, showing

upregulation of IFN-γ in the control group compared to patients with an age of onset over 40 (category A3, p =0.033) (*Figure 3.15*). It should be noted that only two plasma samples had a detectable level of IFN-γ, one in the NC and one in the A2 group. Therefore, although there is a statistically significant difference between A3 patients and the control group, most patients and controls still had the same level of IFN-γ apart from the one positive individual. These data suggest that CD patients had similar cytokine activity to non-IBD controls in their PB.

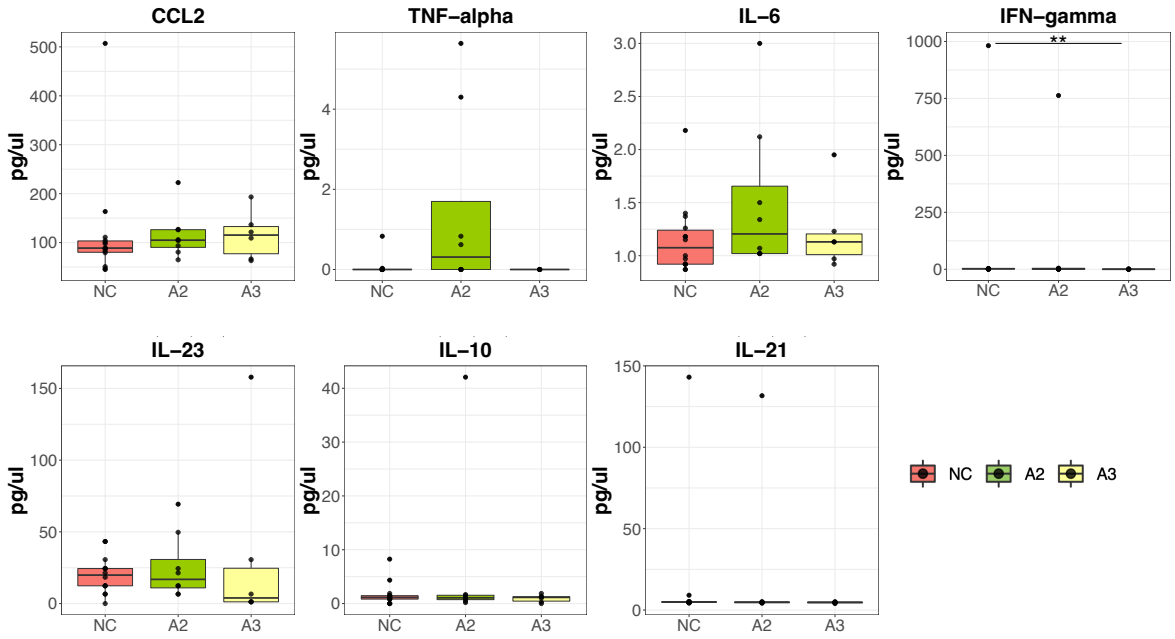


Figure 3.15: ***ANOVA comparing cytokine levels of non-IBD controls (NC) with CD patients categorised based on their age of disease onset.***

CD patients were categorised based on the age of onset of their disease into categories A2 (17-40 years of age) and A3 (>40 years of age). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below). **p<0.01.

3.7 Integration of cytokine and flow cytometry data

Data integration is a valuable tool to combine different types of datasets, in this case, multiplex magnetic bead cytokine assay and flow cytometry outputs. The combination of these data could elucidate the behaviours of T cell types such as Th1, Th2 and Th17 due to associations between PBMCs and the cytokines measured. Furthermore, it may cluster patients and controls based on the expression pattern of peripheral blood cytokine and PBMC levels and identify correlations between all variables. The mixOmics package is a very useful tool for data integration in *R studio*. Dr Umer Zeeshan Ijaz kindly provided the

methodology used for the following analysis. The type of analysis applied is called "Data integration Analysis for Biomarker discovery using Latent variable approaches for Omics studies" (short DIABLO).

The DIABLO analysis initially identifies variables with the highest discriminatory power between two or more sample groups (in this case, CD vs non-IBD controls) and subsequently generates unsupervised hierarchical clustering of the expression pattern and calculates the correlations between all the inputs based on Euclidean distance. The variables that explained the most variation between the two sample groups for both cells and cytokines were visualised using a heatmap with hierarchical clustering (*Figure 3.16*). The cytokines selected by the analyses were IL-21, IFN-γ, IL-23, IL-10, CCL2 and IL-6. The DIABLO considered all PBMC subpopulations for the analysis. The cytokine expression did not show a clear pattern in distinguishing CD and non-IBD controls. Clustering with the cytokines were $CD25^+$ CM and EM $CD8^+$ T cells as the only cell types. The rest of the $CD25^+$ T cells formed one cluster, and so did effector and naïve $CD4^+$ and $CD8^+$ T cells. NK cells clustered separately from all T cells. Overall, the DIABLO analysis of the PBMCs divided the samples into three groups with normal, medium, and low cell numbers. Interestingly, the group with the low PBMC levels was solely comprised of CD patients. NK cells and activated CM and EM $CD8^+$ T cells did not appear to contribute to the signature as much as the rest of the PBMC subpopulations. The analysis did not detect any significant correlations between PBMCs and cytokines.

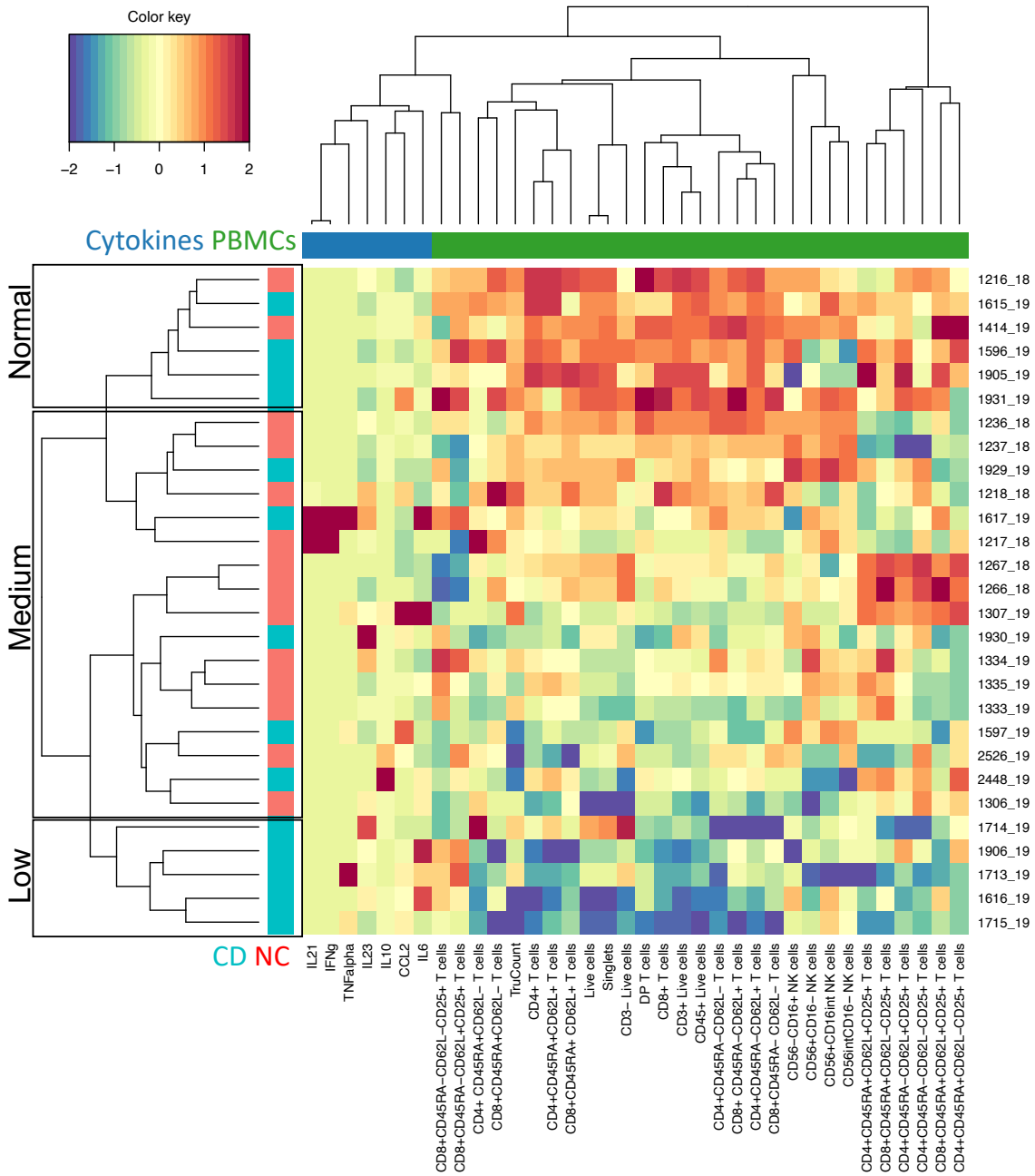


Figure 3.16: ***DIABLO output showing hierarchical clustering of variables (cytokines and PBMCs) and samples (non-IBD controls and CD).***

Peripheral blood was sampled from 14 CD patients and 14 non-IBD controls. Plasma and PBMCs were isolated. Data integration was performed using plasma cytokine levels (in pg/ul), and PBMC subpopulation counts (cells/ml) as measured by multiplex ELISA and flow cytometry. DIABLO analysis was performed in R studio using the *block.plsda()* function of the *mixOmics* package. All statistical analysis was performed in R studio (R version 4.0.2 and below).

In summary, cytokine levels did not exhibit a clear pattern distinguishing CD patients and non-IBD controls. However, PBMC subpopulation levels based on unsupervised clustering separated CD patients into three subsets.

3.8 Discussion

Peripheral blood inflammatory markers are used to monitor disease activity in CD. However, often these do not accurately identify the intestinal inflammation status of patients (Schoepfer *et al.*, 2010b; Papay *et al.*, 2013). This chapter aimed to investigate differences in the peripheral blood signature between CD and non-IBD controls and within the CD cohort to understand differences in their circulating immune cells better. Subsequently, the data were linked to clinical parameters to elucidate potential associations between distinct CD phenotypes with specific peripheral immunopathotypes.

This study's limitations must be considered when interpreting the results in this chapter. This cross-sectional study lacked sample size and, therefore, statistical power. It should therefore be treated as an exploratory study with the potential to guide further research. To confirm the results, larger sample size is required to reach a statistical power of 0.8, especially when stratifying patients into subsets based on their clinical features. Additionally, the analysis in this study failed to accurately distinguish between Tregs and activated T cells due to the shared expression of the CD25 surface receptor. Intracellular staining with the TF FOXP3 and the expression of this surface marker would allow for differentiating activated from regulatory cells. Staining for other TFs upregulated in distinct Th subsets, such as T-bet in Th1, GATA3 in Th2, and RORγt in Th17 cells would also inform about the differentiation phenotype of the Th T cell populations.

No significant differences were found between CD and non-IBD patients in absolute numbers of PBMCs (*Figures 3.3-3.7*). Even though the frequency of the activated CM $CD8^+$ T cells was increased in CD, this observation was not supported by changes in absolute numbers. CD patients displayed a more extensive spread in cell numbers and frequencies for most of the PBMC subpopulations. Many autoimmune and inflammatory conditions show changes in the size of leukocyte populations and levels of inflammation-associated proteins. For example, the activity of RA can be measured using CRP (Shrivastava *et al.*, 2015) and neutrophil-to-lymphocyte ratio (Targońska-Stępniak *et al.*, 2020). Additionally, an increase in PB neutrophils is linked to severe disease and a decrease in lymphocyte count

with upregulation of type I interferon activity in the autoimmune condition SLE (Han *et al.*, 2020).

Based on the location of the inflammation, FC is routinely used in CD disease monitoring in addition to serological markers. When comparing PB inflammatory markers and FC with endoscopic activity, FC showed the strongest correlation (Schoepfer *et al.*, 2010b). Although accurate for some, this suggests that CD inflammation does not necessarily manifest changes in the peripheral blood. This was reflected in our cohort. Stratification of patients based on their HBI did not show any significant differences between the CD subsets when compared with each other and the NC control group by multiple comparison analysis (*Figure 3.8*). This result was supported using linear regression modelling comparing HBI and PBMC subpopulations (data not shown). Interestingly, other inflammatory markers such as CRP levels, WBC, neutrophil and lymphocyte count positively correlated with several PBMC subpopulations (*Figure 3.13*). Studies comparing current methodologies to monitor inflammation in CD have shown that the CDAIS, a more detailed version of the HBI, only shows a loose correlation with changes in PB inflammatory markers such as CRP (Vieira *et al.*, 2009). As previously described, the inflammatory markers are often not more reliable in understanding the current disease state in CD. My data are consistent with these previous studies, showing significant heterogeneity in IBD patients and few correlations between blood biomarkers and clinical disease scores. The lack of differences in PBMC populations and PB cytokine levels could also indicate general low disease activity, which is possible given that the HBI was used to identify inflamed individuals. An endoscopic procedure and/or stool sample analysis would be required to confirm the inflammation state of the patients and should be included in future studies.

Stratification of patients based on the age of disease onset of CD patients had the only significant results in the multiple comparison analyses. Patients with disease onset between 17 and 40 years had a decrease in cytokine-producing NK cells ($C56^{+}CD16$- phenotype) compared to patients who exhibited disease older than 40 years for the first time (*Figure 3.11*) and a decrease in DP T cells and $CD4^{+}$ EM T cells compared to non-IBD controls. Patients with earlier disease onset tend to display a higher genetic predisposition to the disease and, therefore, often experience more severe disease and more complications throughout disease progression (Okou and Kugathasan, 2014). For example, paediatric CD more

frequently leads to hospitalisation and surgery (Duricova *et al.*, 2014). Additionally, the longer the inflammation is ongoing, the more changes the immune system may experience. Therefore, it is perhaps not surprising that patients diagnosed between 17 and 40 years of age display changes when compared to patients with altered disease onset (>40 years of age) and controls. To test this hypothesis, patients could be stratified based on their disease duration in addition to age of onset. The role of PB DP T cells in CD pathology is currently not very well understood. Recent scRNAseq analysis of cynomolgus monkey PB $CD4^+CD8^+$ T cells by Choi *et al.* (2021) suggested that they have diverse expression patterns. 10 T cell subsets were identified, which included *GATA3*-expressing Th2-like, *RORC*-expressing Th17-like, *FOXP3*-expressing Treg-like and granzyme gene-expressing CTL-like cells. In RA, they have been previously linked with erosive disease behaviour ($p < 0.001$) and joint damage ($p = 0.008$) (Nguyen *et al.*, 2022), supporting the hypothesis that they are involved in aberrant immune responses during chronic inflammation.

NK cells were predominantly linked to changes in total WBC count of CD patients. Mucosal NK cells have been shown to drive both inflammation and epithelial cell repair in intestinal inflammation, depending on their state of maturity. Most intestinal NK cells are mature $CD56^{bright}$ cells with high levels of cytokine production, particularly TNF-α and IFN-γ but low cytotoxicity (Leon *et al.*, 2003). Additionally, IFN-γ-producing NK cells were increased in intestinal tissue of CD patients compared to non-IBD controls (Takayama *et al.*, 2010). However, the regression analyses in *Figure 3.13* indicated a correlation of not only cytokine-producing but also cytotoxic NK cells with total WBC numbers. Since a reduction in WBC has been associated with intestinal inflammation in CD and NK cells can be recruited to the gut to promote immune responses against commensals (Lima *et al.*, 2015; Lanier *et al.*, 2022), both NK cell types might be recruited from the periphery during the inflammatory response to enhance IFN-γ and cytotoxic immune responses in the intestine. Surprisingly, WBC count did not significantly correlate with general PBMC populations, such as live $CD45^+$, $CD3^-$ and $CD3^+$ cells. This could suggest that changes in the total WBC count are due to a cell type not included in the PBMC population. During PBMC isolation, granulocytes cannot be extracted from the blood due to their larger size and granularity (Kleiveland, 2015). Differences in the total WBC count might be due to a granulocyte

subpopulation, which are not included in the $CD45^+$ cells of my analysis. Furthermore, a PBMC population that was not quantified, such as B cells or monocytes, may have reduced or increased in CD patients and affected the total WBC count.

In addition to WBC count, lymphocyte and neutrophil counts as well as CRP levels were measured in the PB of CD patients and correlated with the PBMC subpopulations (*Figure 3.13*). Different from the total WBC count, the other three measures showed similar correlation patterns. 5 PBMC subpopulations showed significant positive regressions with all three markers: live $CD3^-$ negative cells, $CD4^+$, naïve $CD4^+$, CM $CD8^+$, and effector $CD8^+$ T cells. Interestingly, the neutrophil count showed the most associations with the PBMC populations, even though granulocytes were not isolated. This indicates that PB neutrophils have similar behaviour to the T cell subpopulations. A study by Therrien *et al.* (2019) collected matched blood and biopsies of 73 CD patients and evaluated their inflammatory status using HBI and SES-CD. Neutrophil infiltration was observed in the colonic mucosa of inflamed patients as defined by SES-CD. Additionally, patients with inactive disease (SES-CD $\leq$1) had higher expression of $CD64^+$ circulating neutrophils than inflamed patients, suggesting neutrophils leave the PB during inflammation and potentially migrate to the intestines. The mechanism of action for the front-line therapeutic vedolizumab is to target the migration of T cells into the intestines. vedolizumab has recently been shown to have a clinical response rate of 47% a year after treatment initiation measured by endoscopic evaluation in a cohort of 117 CD patients (Zingone *et al.*, 2020). The success of the T cell migration-inhibiting therapeutic supports the hypothesis that PB T cells may exhibit a similar migratory pattern to neutrophils, explaining their correlations.

Another interesting observation in the linear regression analyses was the similar correlation pattern between CRP and lymphocyte/neutrophil counts. CRP is an acute phase protein produced by hepatocytes in response to IL-6 and IL-1ß (Hurlimann, Thorbecke and Hochwald, 1966; Morley and Kushner, 1982). It has been found to be the most reliable PB biomarker of inflammation in CD (Yazar *et al.*, 2020b). It cannot be concluded from the data whether the changes in the PBMC numbers are causative for CRP production. However, it can be suggested that the same stimulus driving the release of CRP potentially also causes changes in the PBMC subpopulations that correlate with the protein.

The cytokine profile of CD patients and NC groups did not differ. One statistical difference was observed when patients were stratified based on the age of onset and showed a decrease in patients with late disease onset (>40 years of age) of IFN-γ when compared to non-IBD controls (*Figure 3.15*). Even though this finding might relate to the differences in cytokine-producing NK cells found between the patient subsets A2 and A3 in the PBMC analysis, most samples had non-detectable IFN-γ levels, making it difficult to draw any conclusions from the results. When integrating flow cytometry and multiplex ELISA data, it becomes evident that cytokines did not segregate CD patients from NC controls (*Figure 3.16*). However, a subset of CD patients was identified, which displayed reduced levels of T cells in comparison to the rest of the cohort, including non-IBD controls. This finding was very interesting and supported my hypothesis that only a subpopulation of CD patients may exhibit changes in their circulatory immune cells depending on their disease immunopathotypes or disease state.

3.9 Conclusions

From the work in this chapter, it can be concluded that there were no consistent differences between CD patients and non-IBD controls in terms of PBMCs and cytokines levels. However, unsupervised clustering of patients and controls elucidated three patient subsets that can be distinguished based on their PBMC counts but not cytokine levels.

Chapter 4: Identification of a CD patient subset with reduced PB T cells and enhanced Th1 and $CD8^+$ intestinal T cell immune responses

4.1 Introduction

In *chapter 3*, a subset of CD patients was identified, characterised by a reduction of T cells in their circulation (*Figure 3.16*). Stratification of patients based on their T cell profile has proven successful in other inflammatory diseases, such as RA (Bader *et al.*, 2019; Frederique Ponchel *et al.*, 2020). Based on PB regulatory and naïve T cell levels, predictive models could estimate whether at-risk individuals progressed to inflammatory arthritis or not (F Ponchel *et al.*, 2020). Furthermore, these blood parameters could predict response to the front-line RA therapeutic methotrexate, demonstrating that PB T cell profiling is a valuable tool for disease stratification in inflammatory diseases and can potentially improve current treatment strategies. Therefore, further understanding of disease mechanisms in CD patient subsets with distinct PB signatures may help to elucidate whether reduced T cells in circulation could be employed as a biomarker for inflammation or to determine new potential treatment targets.

Aims:

In this chapter, I aimed to investigate further the characteristics of CD patients with PB T cell low levels identified in *chapter 3*. A publicly deposited whole blood transcriptomic dataset with a larger sample size was studied to validate the findings of the small cohort analysed with flow cytometry. A T cell gene signature was generated to quantify T cell levels in the transcriptomic dataset. Subsequently, an additional microarray dataset with matched blood and biopsy samples was studied to identify whether low PB T cell levels have the potential to function as a biomarker for distinct intestinal immune signatures.

4.2 PB Flow cytometry: Stratification of CD patients based on their circulating T cell profile

Previous DIABLO analysis showed a difference in T cell populations between CD patients (*Figure 3.16*). Most patients in the "low" subset had reduced levels of all T cell populations apart from $CD25^+$ T cells. Using multiple comparison analyses, the T cell levels of non-IBD controls (NC), CD patients with normal PB T cells (NT) and CD patients with low PB T cells (LT) were compared, as described previously in *chapter 3*. Based on the DIABLO analysis in *Figure 3.16*, cytokine levels did not contribute to the clustering; therefore, I decided to focus on PBMCs only.

4.2.1 Multiple Comparison of general PBMC populations, T and NK cell levels between CD subsets and non-IBD controls

The absolute number of cells, generated using percentage parent populations and total cell count/ml post-PBMC isolation, was compared between the three different groups. As expected, $CD45^+$ and $CD3^+$ live cells, $CD4^+$ and $CD8^+$ T cells were significantly reduced in LT patients compared to non-IBD controls and the NT patient group (*Figure 4.1, Table 4.1*). Interestingly, $CD3^-$ cells were not different. DP T cells were fewer in LT patients when compared to the control but not the NT patient group.

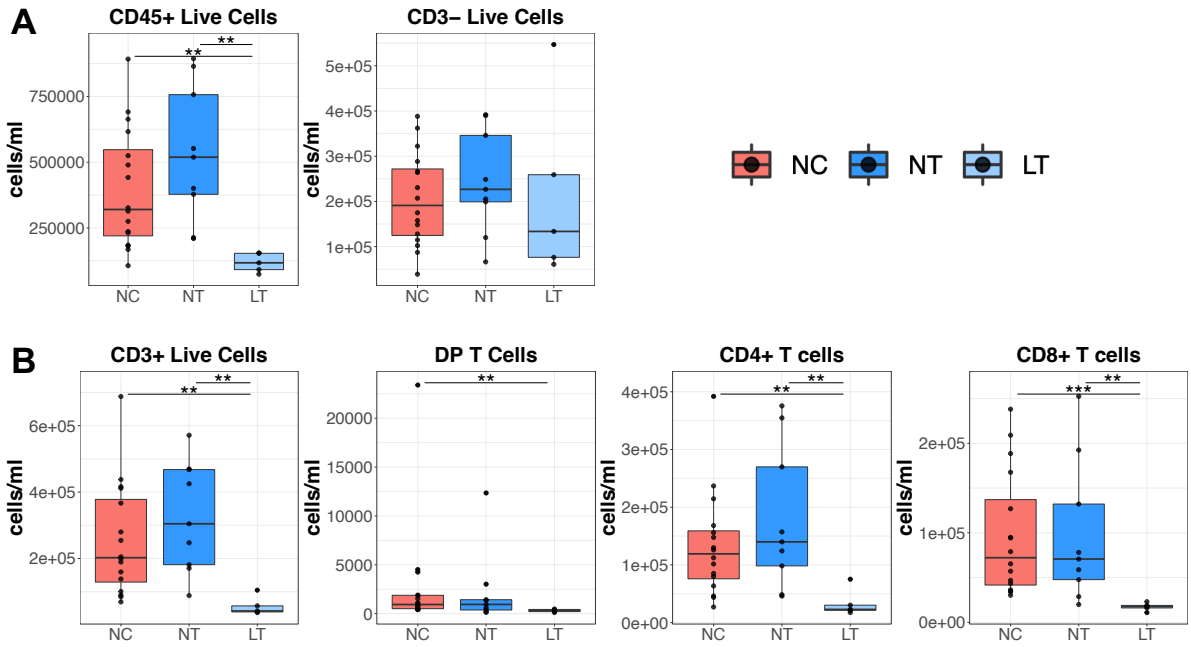


Figure 4.1: ***Comparison of absolute numbers of CD45⁺ and CD3⁻ cell and T cell levels between non-IBD controls and CD patients categorised based on their PB T cell profile.***

PBMCs of 14 CD patients and 16 non-IBD controls (NC) were analysed using flow cytometry. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =5) and T cell normal (NT, *n* =9). A) Multiple comparison of live $CD45^+$ and $CD3^-$ absolute numbers (cells/ml). B) Multiple comparison of live $CD3^+$, DP, $CD4^+$ and $CD8^+$ T cells absolute numbers (cells/ml). Flow cytometry data were quantified in FlowJo™ (Version 10.8 and below). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below). **p<0.01, ***p<0.001.

Feature	Up	Down	p
$CD45^+$ live cells	NC	LT	0.0021
$CD45^+$ live cells	NT	LT	0.0021
$CD3^+$ live cells	NC	LT	0.0035
$CD3^+$ live cells	NT	LT	0.004
DP T cells	NC	LT	0.0085
$CD4^+$ Live cells	NC	LT	0.0056
$CD4^+$ Live cells	NT	LT	0.008
$CD8^+$ Live cells	NC	LT	0.0003
$CD8^+$ Live cells	NT	LT	0.004

Table 4.1: ***P-values associated with comparison of absolute numbers of $CD45^+$ and $CD3^-$ cell and T cell levels between non-IBD controls and CD patients categorised based on their PB T cell profile.***

PBMCs of 14 CD patients and 16 non-IBD controls (NC) were analysed using flow cytometry. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =5) and T cell normal (NT, *n* =9). A) Multiple comparison of live $CD45^+$ and $CD3^-$ absolute numbers (cells/ml). B) Multiple comparison of live $CD3^+$, DP, $CD4^+$ and $CD8^+$ T cells absolute numbers (cells/ml). Flow cytometry data were quantified in FlowJo™ (Version 10.8 and below). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).

$CD4^+$ and $CD8^+$ T cell naïve, CM, effector and EM subpopulations were studied next *(Figure 4.2A/B, Table 4.2*). Most of the T cell subsets displayed a significant decrease in the LT subset when compared to the NT and non-IBD cohorts, supporting the findings in *chapter 3*. No difference was found in effector $CD4^+$ and CM $CD8^+$ T cells. In the case of effector $CD4^+$, this finding might be due to one LT patient exhibiting very high levels of this cell phenotype. NT patients did not differ from the control group with respect to their PB T cell profile.

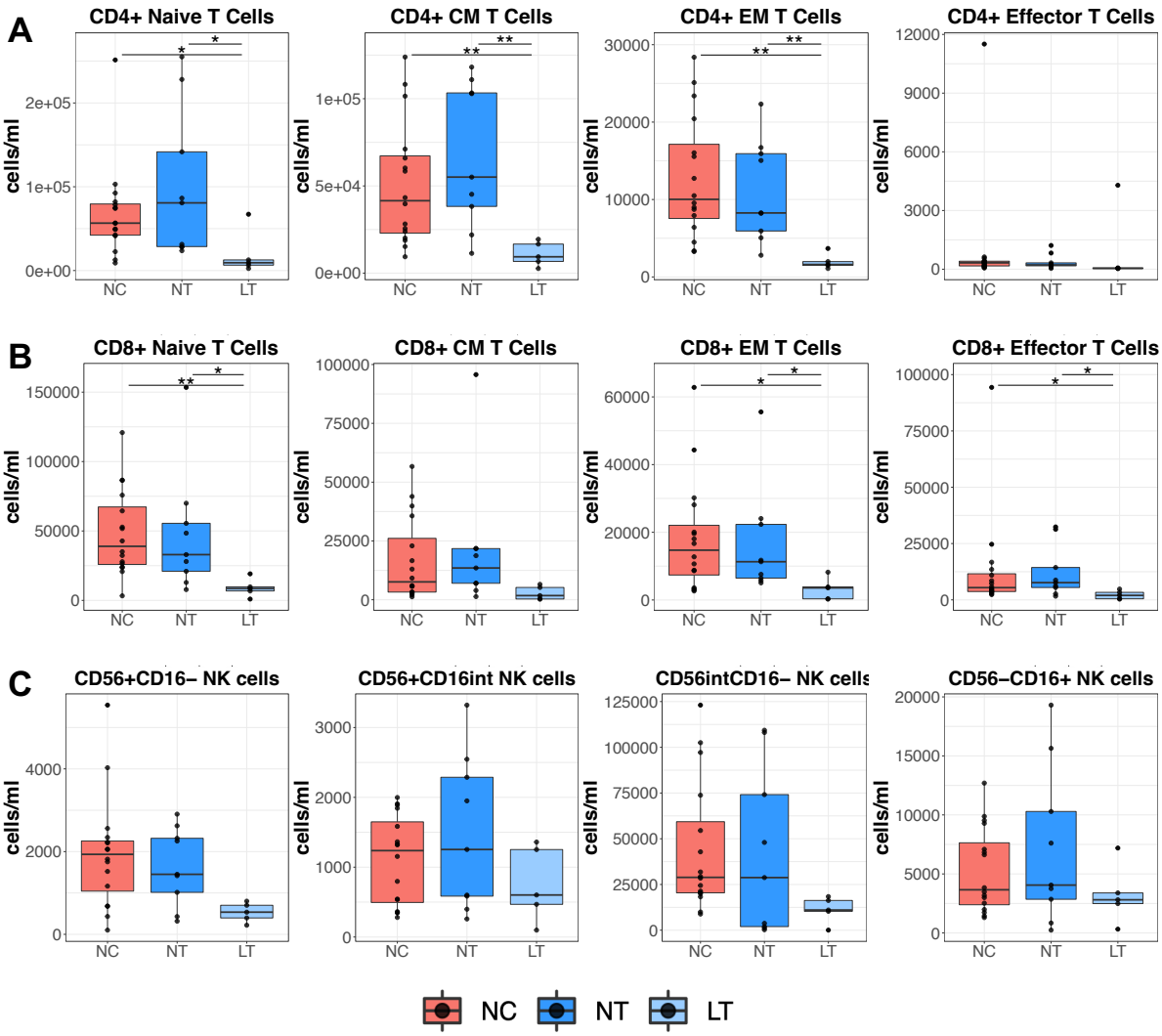


Figure 4.2: ***Comparison of absolute numbers of CD4⁺ and CD8⁺ T cell and NK cell subpopulation levels between non-IBD controls and CD patients categorised based on their PB T cell profile.***

PBMCs of 14 CD patients and 16 non-IBD controls were analysed using flow cytometry. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =5) and T cell normal (NT, *n* =9). Observed values were defined as absolute number of PBMC subpopulations. A) Absolute numbers (cells/ml) show the levels of $CD4^+$ T cell subpopulations: naïve T cells ($CD4^+CD45RA^+CD62L^+$), central memory (CM, $CD4^+CD45RA^-CD62L^+$), effector ($CD4^+CD45RA^+CD62L^-$) and effector memory (EM, $CD4^+CD45RA^-CD62L^-$). B) Absolute numbers (cells/ml) show the levels of $CD8^+$ T cell subpopulations: naïve T cells ($CD8^+CD45RA^+CD62L^+$), central memory (CM, $CD8^+CD45RA^-CD62L^+$), effector ($CD8^+CD45RA^+CD62L^-$) and effector memory (EM, $CD8^+CD45RA^-CD62L^-$). C) Absolute numbers (cells/ml) show the levels of NK cell subpopulations. Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below). *p<0.05, **p<0.01.

Feature	Up	Down	p
$CD4^+$ CM T cells	NC	LT	0.0056
$CD4^+$ CM T cells	NT	LT	0.008
$CD4^+$ Naïve T cells	NC	LT	0.036
$CD4^+$ Naïve T cells	NT	LT	0.036
$CD4^+$ EM T cells	NC	LT	0.0012
$CD4^+$ EM T cells	NT	LT	0.004
$CD8^+$ Naïve T cells	NC	LT	0.0035
$CD8^+$ Naïve T cells	NT	LT	0.024
$CD8^+$ EM T cells	NC	LT	0.036
$CD8^+$ EM T cells	NT	LT	0.036
$CD8^+$ Effector T cells	NC	LT	0.0333
$CD8^+$ Effector T cells	NT	LT	0.038

Table 4.2: ***P-values associated with comparison of absolute numbers of $CD4^+$ and $CD8^+$ T cell subpopulation levels between non-IBD controls and CD patients categorised based on their PB T cell profile.***

PBMCs of 14 CD patients and 16 non-IBD controls were analysed using flow cytometry. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =5) and T cell normal (NT, *n* =9). Observed values were defined as absolute number of PBMC subpopulations. A) Absolute numbers (cells/ml) show the levels of $CD4^+$ T cell subpopulations: naïve T cells ($CD4^+CD45RA^+CD62L^+$), central memory (CM, $CD4^+CD45RA^-CD62L^+$), effector ($CD4^+CD45RA^+CD62L^-$) and effector memory (EM, $CD4^+CD45RA^-CD62L^-$). B) Absolute numbers (cells/ml) show the levels of $CD8^+$ T cell subpopulations: naïve T cells ($CD8^+CD45RA^+CD62L^+$), central memory (CM, $CD8^+CD45RA^-CD62L^+$), effector ($CD8^+CD45RA^+CD62L^-$) and effector memory (EM, $CD8^+CD45RA^-CD62L^-$). C) Absolute numbers (cells/ml) show the levels of NK cell subpopulations. Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).

To test whether the changes in the PB profile of LT patients were solely related to T cells, NK cell levels were also stratified (*Figure 4.2C, Table 4.2*). No significant changes were seen in the multiple comparison analysis, supporting the hypothesis that the patient subset is defined by reduced levels of PB T cells.

4.2.2 Summary of multiple comparison analysis

The stratification of patients based on their circulating immune cell levels by multiple comparison analysis exhibited that LT CD patients had fewer T cells, as was shown by the DIABLO analysis. $CD45^+$ live cells, comprising all leukocytes, were reduced, but there was no difference in $CD3^-$ cand NK cell numbers, which indicated that this loss was due to lower levels of T cells in the PB. Effector $CD4^+$

T cells did not differ, indicating that not all but only specific subpopulations of T helper cells may be involved in this signature.

4.3 Whole Blood Transcriptomics: Validation of T cell low PB CD subset in a microarray dataset

Disease stratification is often limited due to the sample size within a patient cohort. To investigate whether the findings of the flow cytometry analyses were reproducible, a publicly available microarray dataset, deposited with the NCBI under the accession number GSE94648 by Planell *et al*. (2017), was examined. cDNA from fresh whole blood of 48 CD patients and 22 non-IBD controls was hybridised and subsequently normalised and corrected for batch effect by the authors of this study as described in *chapter 2*.

4.3.1 Summary of GSE94648 patient cohorts

Patient characteristics are summarised in *Table 4.3*. Disease activity was measured using the CDEIS, a method scoring the extent and number of ulcerations in the intestine of patients during endoscopy. 81.25% of patients were inflamed, and only 9 patients did not show disease activity in the gut. The majority of patients were on monotherapies (58.33%), such as general immunosuppressants (thiopurines and methotrexate, *n* =13), anti-TNF biologics (*n* =8), 5-ASA drugs (*n* =5) or steroids (*n* =2). A quarter of patients were not on any treatment at the time of sampling, and 16.67% received combination therapies.

In terms of disease phenotype, data for two out of the three Montreal classification parameters (location and behaviour) were available. Most patients had ileocolonic (category L3, 56.25%) disease, followed by individuals with inflammation restricted to the ileum (category L1, 31.25%). Only 6 patients were in category L2 representing the colonic disease location. The inflammatory disease behaviour (category B1) classified 60.42% of the cohort. 27.08% displayed stricturing disease (category B2), and 5 patients had penetrating disease behaviour (category B3). One patient had both stricturing and penetrating CD and which was described as category B2+B3.

	CD (*n= 48*)	**NC *(n= 22)***
Mean Age, years (Range)	37.292 (18-67)	44.591 (23-65)
Standard Deviation Age, years	12.697	9.359
F:M ratio	1.136	0.692
Mean CDEIS (Range)	7.5375(0-36)	N/A
Endoscopic Activity *Inflamed* *Non-inflamed*	 39 (81.25%) 9 (18.75%)	N/A
Treatment **Monotherapies** *Steroids* *Immunosuppressants* *5-ASA* *Anti-TNF* **Combination Therapies** *Immunosuppressants + Anti-TNF* *Immunosuppressants + Steroids* *Anti-TNF + Steroids* **No treatment**	 **28 (58.33%)** 2 13 5 8 **8 (16.67%)** **4** 3 1 **12 (25.00%)**	N/A
Location (%) L1: *Ileal* L2: *Colonic* L3: *Ileocolonic* Behaviour (%) B1: *Inflammatory* B2: *Stricturing* B3: *Penetrating* B2+B3: *Stricturing/Penetrating*	 15 (31.25%) 6 (12.50%) 27 (56.25%) 29 (60.42%) 13 (27.08%) 5 (10.41%) 1 (2.08%)	N/A

Table 4.3: ***Summary of cohort information of GSE94648 microarray dataset.***

48 CD patient samples and 22 non-IBD volunteers (NC) were recruited by Planell *et al.* (2017). Markers of inflammation (CRP, percentage neutrophils/lymphocytes/monocytes, Crohn's disease endoscopy index of severity (CDEIS)), current treatments and Montreal classifications (disease behaviour and location) were included for CD patients. Immunosuppressants comprise thiopurines and methotrexate. N/A: not applicable. CRP: C-reactive protein. 5-ASA: aminosalicylates.

4.3.2 Identification of T cell-associated genes

Since microarray studies report levels of specific sequences, identified by their hybridisation to chip-bound probes, a list of T cell-associated gene probes was required to identify patients in the study that might be defined by the LT phenotype. This signature was generated by identifying genes that correlated ($p < 0.05$, $R^2 > 0.5$) with the three genes encoding chains of the T cell marker CD3 (*CD3D*, *CD3E* and *CD3G*) in the *GSE94648* dataset. 32 genes fulfilled the correlation criteria (*Table 4.4, Figure 4.3*). Among these was the gene encoding the CD3 gamma chain of the T cell receptor (*CD3G*) and zeta chain of TCR-associated

protein kinase 70 (*ZAP70*). Also on the list were the co-stimulatory *CD28* (cluster of differentiation 28) gene, found on the surface 80-90% of $CD4^+$ and around half of $CD8^+$ T cell populations (Acuto and Michel, 2003; Esensten *et al.*, 2016), as well as *CD27* (cluster of differentiation 27), highly expressed on naïve PB T cells (Hamann *et al.*, 1997; Hendriks *et al.*, 2000).

Gene	CD3D	CD3E	CD3G
TLE5	0.705418605811962	0.622792139225478	0.562533948851299
ANXA6	0.523322207376257	0.664396832164825	0.594029223881922
CAMK4	0.518018131505005	0.618664131451207	0.652769878072822
CD3G	0.506110044188576	0.549063518659444	1
CD27	0.510919274524925	0.696173259739704	0.630283847050108
CD28	0.510149401819454	0.568897389847317	0.590511770303637
GPX7	0.673250654533225	0.528617177692144	0.625245390936192
IMPDH2	0.543376443259284	0.710854679218225	0.624754295914872
LCK	0.601325243953915	0.86805438180028	0.677122602912474
LDHB	0.60836911272632	0.726908374131293	0.626747982710452
LY9	0.660592257282905	0.7690121001907	0.604243843140608
SLAMF1	0.548940351409418	0.576345019264593	0.504704526578804
TLE2	0.651362178625753	0.603177037662042	0.519617805217422
ZAP70	0.636614589607434	0.746109372799902	0.505932820191847
HSD17B8	0.57876858898041	0.502412159047092	0.502790945054762
SKAP1	0.629567165380729	0.726709742673779	0.528742230157564
DHRS3	0.52556630502561	0.692373315292349	0.602442300900339
TESPA1	0.556572854374804	0.690566419774166	0.554549857251067
ZBTB40	0.506665549698441	0.512465185662872	0.537269347892329
ABHD14A	0.568118196356032	0.645618635987348	0.587916575122029
LAT	0.752421140253703	0.733918873880484	0.52127235655223
SIT 1.00	0.532879694297351	0.762575838740822	0.621764321518266
CUTA	0.703277042956196	0.578108173302959	0.587093686905399
UBASH3A	0.640774493124492	0.646473044000381	0.520575112165194
PJA1	0.55996413655866	0.650612005477586	0.573270742145359

POLR1E	0.653683350080017	0.551957186813377	0.618521699057621
TMEM116	0.669582371426608	0.515390509326787	0.528084457259691
C12orf65	0.517119488458397	0.63475829094779	0.652924868315472
THEM4	0.512709491004936	0.650141866226768	0.51087286720443
OCIAD2	0.582544582956141	0.624575032041996	0.678920915705708
ATG9B	0.560289070824005	0.586396493229498	0.564831853808389
SBK1	0.50231867764311	0.716769272102349	0.513191467677327

Table 4.4: ***PB T cell-associated gene signature.***

32 signature genes and their R^2-value of correlations with *cd3d*, *cd3e* and *cd3g* expression. Genes were selected based on significance ($p < 0.05$) and correlation ($R^2 > 0.5$). Correlation was performed using the *rcorr()* function of the Hmisc package. Data used are deposited on NCBI under GSE94648 Planell *et al.* (2017). All analysis was performed in R studio (R version 4.0.2 and below).

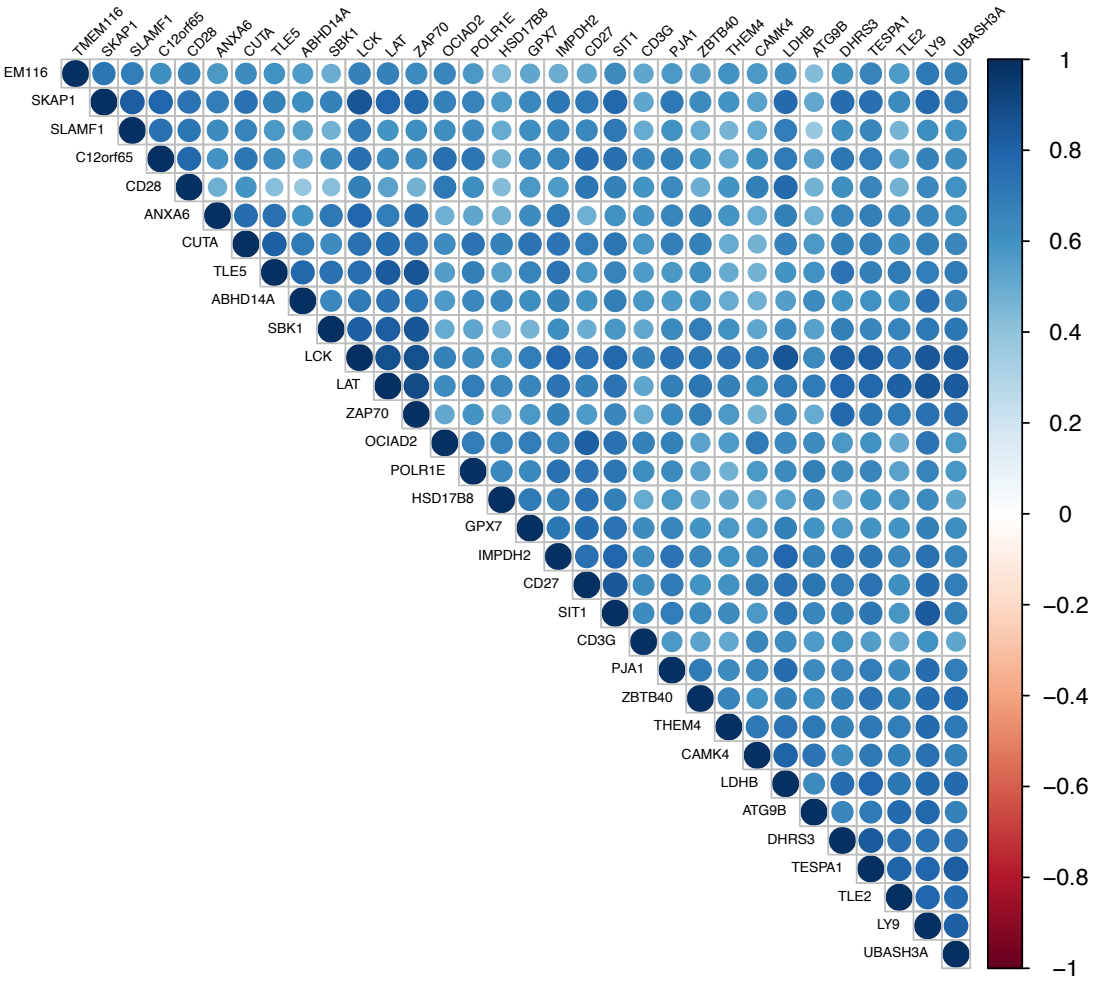


Figure 4.3: ***Correlation plot showing PB T cell-associated gene signature.***

32 signature genes and their R^2-value of correlations with *cd3d*, *cd3e* and *cd3g* expression. Genes were selected based on significance ($p < 0.05$) and correlation ($R^2 > 0.5$). Correlation was performed using the *rcorr()* function of the Hmisc package. The correlation plot was generated using the corrplot package. Data used are deposited on NCBI under GSE94648 by Planell *et al.* (2017). All analysis was performed in R studio (R version 4.0.2 and below).

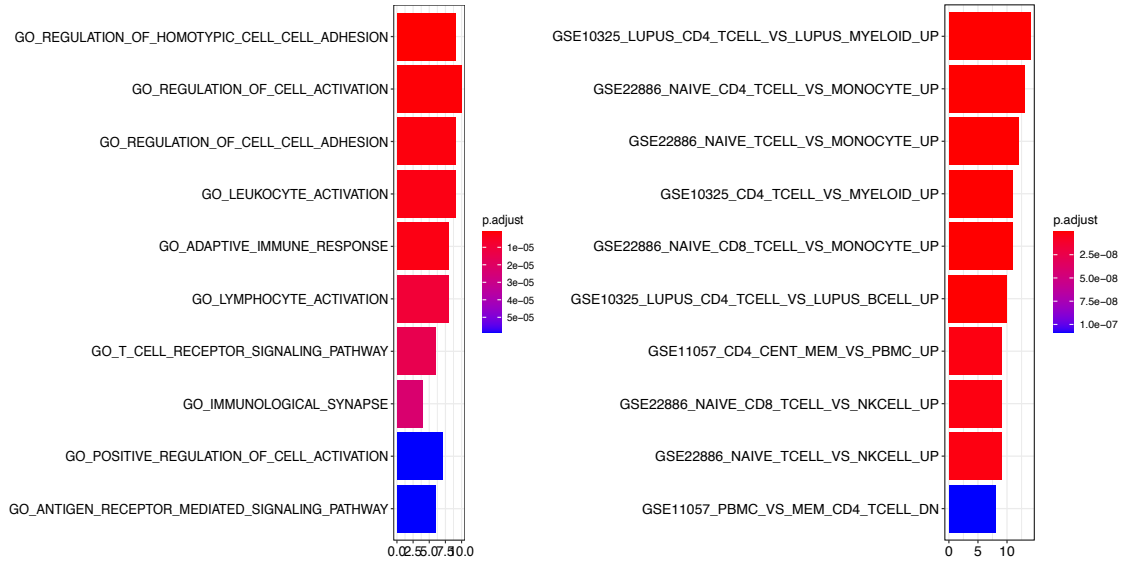


Figure 4.4: ***Gene set enrichment analysis of PB T cell-associated gene signature.***

32 signature genes were tested for significant associations with gene ontologies (GOs) and medical subheadings (MeSH). Molecular signature database (MsigDB) was applied as reference for GO analysis using the *enricher()* function of the msigdbr package. MeSH analysis was performed with the *enrichMESH()* function from the meshes package accessing the pub2pubmed database curated by the NCBI. Data used are deposited on NCBI under GSE94648. All analysis was performed in R studio (R version 4.0.2 and below) and required the clusterProfiler package.

To confirm that the signature was, in fact, T cell-related, I performed gene ontology (GO) and medical subheading (MeSH) analyses, two methods showing enrichment of gene sets (*Figure 4.4*). The enriched categories included the GOs: *adaptive immune response*, *lymphocyte activation* and *T cell receptor signalling pathway*. The MeSH analysis compares the gene expression between immune cells based on publicly deposited datasets on NCBI. For example, GSE22886_NAIVE_TCELL_V_MONOCYTE_UP states that the genes in the signature are upregulated in naïve T cells compared to myeloid cells in the dataset deposited under NCBI accession GSE22886. The 10 MeSH displaying the most significant association with the gene signature showed upregulation of T cell populations compared to other leukocyte phenotypes. These results validated that a blood T cell-associated signature had been generated.

4.3.3 sparse Least-Squares Discriminant Analysis (sPLS-DA): Stratification of CD patients and non-IBD controls based on the expression of PB T cell-associated genes in adult CD

After generating and validating a blood T cell-associated gene signature, the 32 target genes were used to study differences among PB T cell levels in CD patients.

Discrimination analysis was applied to stratify CD patients based on their PB T cell levels, called sparse least-squares discriminant analysis (sPLS-DA). The sPLS-DA was performed in R studio using the *splsda()* function of the mixOmics package (Rohart *et al.*, 2017). Similar to the DIABLO, this analysis allows the identification of variables (in this case T cell signature gene expression levels) that have the highest discriminating power between two or more sample groups. Like the DIABLO, it subsequently clusters samples and variables using the Euclidean distance. Patients were stratified based on their inflammation profile, as defined by endoscopic activity, into active disease (CD) and disease in remission (CDR). The two CD subsets were compared to the non-IBD controls (NC). To correct for batch effect, Planell *et al*. included technical replicates of a selection of samples as batch controls. These controls were included in this analysis to ensure that the batch effect was corrected successfully and were labelled in *Figure 4.5* using diamond shapes.

The sPLS-DA loaded all genes from the T cell-associated signature, indicating that all genes contribute to the variation between the sample group. Unsupervised clustering split the samples into two distinct clusters, one characterised by low gene expression (*Figure 4.5*). 10 CD samples were in this group, made up of 8 patients and 2 batch controls, as well as one non-IBD control. The rest of the CD patients (*n* =38) were not distinguished from the controls and displayed a gradient of T cell gene expression. None of the endoscopically inactive patients displayed the low PB T cell phenotype and mixed in with the remaining CD cohort and controls. Consistent with the previous flow cytometry findings, this analysis demonstrated that T cell-associated gene expression also identified a subset of CD patients with potentially fewer circulating T cells.

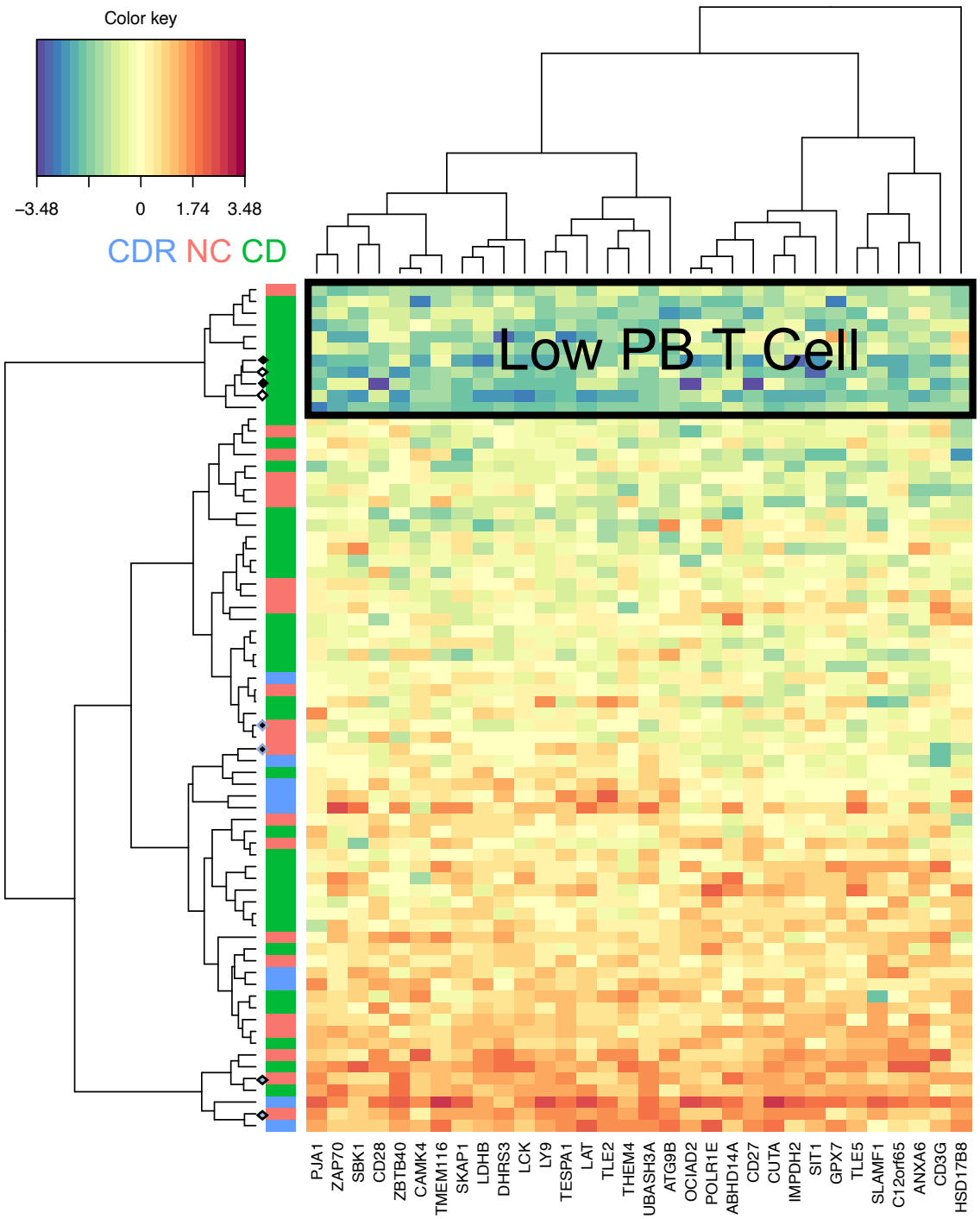


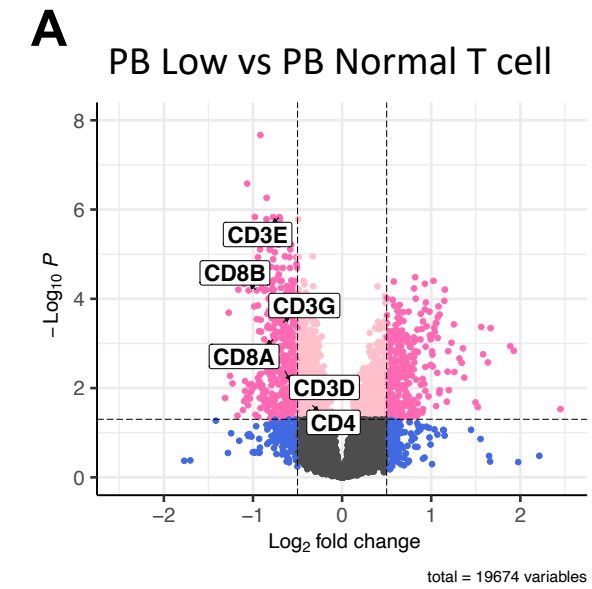
Figure 4.5: ***Sparse least-squares discriminant analysis (sPLS-DA) output comparing PB T cell-associated gene expression levels between active CD, CD in remission and non-IBD controls.***

sPLS-DA analysis elucidates variables (PB T cell-associated genes) with the highest discriminant power to distinguish three sample groups. Sample groups included 22 non-IBD controls (NC, red), 39 CD patients with active disease (CD, green) and 9 patients in remission (CDR, blue) as defined by CDEIS. Data analysed were in form of normalised probe intensity values. Diamond shapes on the left-hand side of the sample panel are indicators for matching batch controls. The *splsda()* function from the mixomics package was used for analysis. Data used is deposited on NCBI under GSE94648 by Planell *et al.* (2017). All analysis was performed in R studio (R version 4.0.2 and below).

4.3.4 Differential Gene Expression analysis: Normal PB T cell vs Low PB T cell patients in adult CD

Subsequent to stratifying the CD cohort into normal (NT) and low PB T cells (LT), I used the GEO2R, a tool developed by the NCBI, to perform a differential gene expression analysis and identify significant differences between these CD subsets (*Figure 4.6A*). 2077 differentially expressed genes (DEGs) were identified, and the top 10 DEGs are displayed in *Figure 4.6B*. Among these were *CD3E,* encoding the CD3 epsilon chain of the TCR, the IL-21 receptor gene *IL21R*, highly expressed on activated T cells (Wu *et al.*, 2005), and the lymphocyte-specific protein tyrosine kinase (*LCK*), involved in TCR signalling (Rossy, Williamson and Gaus, 2012), all of which were down-regulated in the LT group.

Using WikiPathway analysis (Wu *et al.*, 2012, 2021), over-represented pathways within the DEGs were studied (*Figure 4.6C*). The pathways *T cell receptor (TCR) and co-stimulatory signaling, TCR pathway during Staphylococcus aureus infection* and *TCR signaling pathway* were the top three significantly enriched pathways with most genes being reduced in LT CD patients. Additionally, changes in energy metabolism and B cell receptor signaling pathway were linked to the DEGs. The T and B cell receptor pathways had many overlapping genes, whereas the genes in the energy metabolism pathway were distinct. Changes in energy metabolism could be a sign of an altered activation status of T cells. Overall, the results of the sPLS-DA, DEG and pathway enrichment analyses indicated that a subset of CD patients with fewer PB T cells was also present in this patient cohort.



B

Gene	logFC	p.adj
SPOCK2	-0.9166209	0.0000000214
IL24	-1.064971	0.000000262
LCK	-0.8462436	0.000000547
CD3E	-0.7029855	0.00000147
IL21R	-0.7731831	0.00000147
FCMR	-0.9788793	0.00000147
USP11	-0.6966697	0.00000164
LY9	-0.8472964	0.00000164
FAM102A	-0.9178745	0.00000238
TSPYL2	-0.8320792	0.00000578

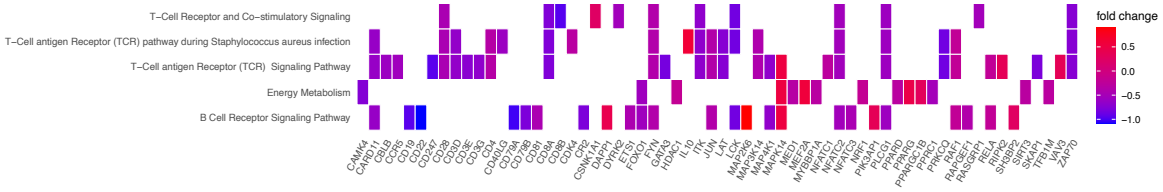


Figure 4.6: ***Differential gene expression analysis comparing PB low and normal T cell CD groups.***

A) Volcano plot showing differential gene expression, highlighting DEGs (*adj.p* <0.05) in rose and genes with -0.5< log2 fold change >0.5 in blue. Genes with -0.5< log2 fold change >0.5 and p <0.05 were shown in pink. B) Table with top 10 DEGs according to adjusted p-values. C) Pathways significantly associated with DEGs. The analysis was performed with the *enricher()* function using the WikiPathway wpid2gene database (Version 20180810). Data used is deposited on NCBI under GSE94648 by Planell *et al*. (2017). Differential gene expression analysis was performed using the GEO2R tool developed by NCBI. Pathway analysis was performed in R studio (R version 4.0.2 and below) and required the clusterProfiler package.

4.3.5 Multiple comparison analysis of PB T cell-associated gene expression between adult CD subsets and non-IBD controls

To further characterise the LT patient subset, multiple comparison analysis was used to study T cell subpopulation genes. The expression level of genes encoding the T cell co-receptors CD3 (*CD3D*, *CD3E* and *CD3G*), CD4 (*CD4*) and CD8 alpha and beta chains (*CD8A* and *CD8B*) were investigated and compared between controls and the CD groups (*Figure 4.7*, *Table 4.5*). Apart from *CD3D*, all genes in this ANOVA model were significantly reduced in LT patients in comparison to NT patients and controls. Interestingly, the expression of *CD3D* was higher in NT. Overall, these data suggested that both $CD4^+$ and $CD8^+$ T cells are decreased in

the low T cell subset compared to controls and patients with normal T cell gene expression levels.

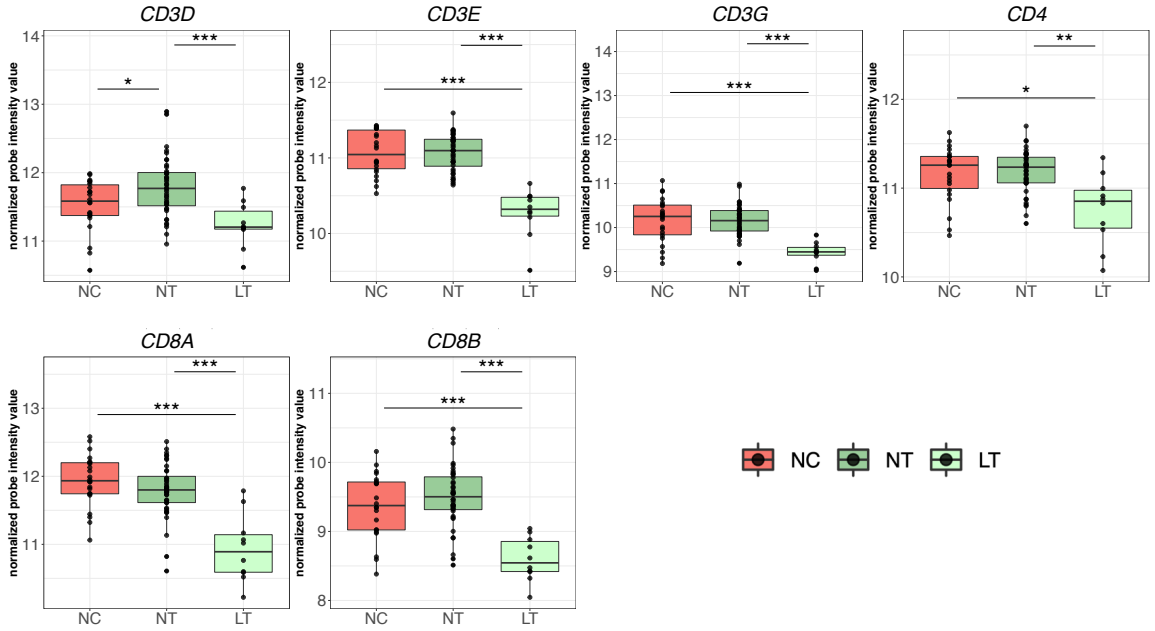


Figure 4.7: ***Comparison of the expression of genes encoding CD3, CD4 and CD8 receptors between non-IBD controls and CD patients categorised based on their PB T cell profile.***

RNA derived from the whole blood of 48 CD patients and 22 non-IBD controls (NC) was analysed using microarray hybridisation, and normalised probe intensity values were used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =8) and T cell normal (NT, *n* =40). Data used are deposited on NCBI under GSE94648 by Planell *et al*. (2017). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below). *p<0.05, **p<0.01, ***p<0.001.

Feature	Up	Down	p
CD3D	NT	NC	0.0407
CD3D	NT	LT	0.0007
CD3E	NC	LT	<0.0001
CD3E	NT	LT	<0.0001
CD3G	NC	LT	<0.0001
CD3G	NT	LT	<0.0001
CD4	NC	LT	0.0163
CD4	NT	LT	0.0037
CD8A	NC	LT	<0.0001
CD8A	NT	LT	<0.0001
CD8B	NC	LT	<0.0001
CD8B	NT	LT	<0.0001

Table 4.5: ***P-values associated with comparison of the expression of genes encoding CD3, CD4 and CD8 receptors between non-IBD controls and adult CD patients categorised based on their PB T cell profile.***

RNA derived from the whole blood of 48 CD patients and 22 non-IBD controls (NC) was analysed using microarray hybridisation, and normalised probe intensity values were used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =8) and T cell normal (NT, *n* =40). Data used are deposited on NCBI under GSE94648 by Planell *et al*. (2017). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).

Percentages of immune cells can be quantified in bulk transcriptomics using digital cytometry tools such as CIBERSORTx (Newman *et al.*, 2019b). This method deconvolutes the microarray data using a signature matrix with the estimated gene expression of isolated immune cell populations. This type of analysis is useful to validate changes in general $CD4^+$ and $CD8^+$ T cells but also to identify differences in Th subsets. The CIBERSORTx output quantified $CD8^+$ and overall $CD4^+$ T cells as well as Th1 and Th2 polarised cells post *in vitro* stimulation (*Figure 4.8, Table 4.6*). $CD8^+$ T cells were less abundant in LT patients compared to NT and NC groups, reflecting the results of the gene expression pattern. The changes within the $CD4^+$ populations differed from the findings in *Figure 4.7*, which showed a decrease in *CD4* expression. However, overall $CD4^+$ T cells and Th2 cells showed no changes between the sample groups. Interestingly, Th1 polarised cells appeared to be fewer in LT than in NT CD subsets, suggesting that the reduction in *CD4* gene expression might relate to a reduction in Th1 T cells. Additionally, it

should be noted that even though CIBERSORTx digital cytometry approach has been validated extensively in different tissues also, including PB, it only generates an estimate of immune cell proportion.

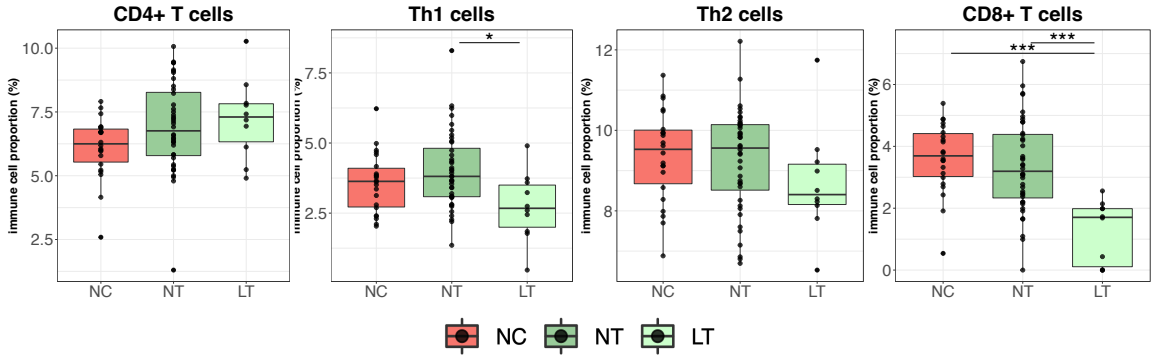


Figure 4.8: ***Comparison of CD4⁺ and CD8⁺ T cell percentages per sample as calculated by digital cytometry between non-IBD controls and adult CD patients categorised based on their PB T cell profile.***

Proportions of T cell subsets in each sample were compared between 48 CD patients and 22 non-IBD controls (NC). CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =8) and T cell normal (NT, *n* =40). The digital cytometry tool CIBERSORTx was used to deconvolute microarray analysis-derived normalised probe intensity values and quantify the frequencies of immune cell populations in each sample. Data used are deposited on NCBI under GSE94648 by Planell *et al*. (2017). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below). *p<0.05, ***p<0.001.

Feature	Up	Down	p
CD8⁺ T cells	NC	LT	<0.0001
CD8⁺ T cells	NT	LT	<0.0001
Th1 cells	NT	LT	0.014

Table 4.6: ***P-values associated with the comparison of CD4 and CD8 T cell percentages per sample as calculated by digital cytometry between non-IBD controls and adult CD patients categorised based on their PB T cell profile.***

Proportions of T cell subsets in each sample were compared between 48 CD patients and 22 non-IBD controls (NC). CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =8) and T cell normal (NT, *n* =40). The digital cytometry tool CIBERSORTx was used to deconvolute microarray analysis-derived normalised probe intensity values and quantify the frequencies of immune cell populations in each sample. Data used are deposited on NCBI under GSE94648 by Planell *et al*. (2017). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).

Cytokines and surface receptors are useful markers to distinguish Th1 and Th2 cells, and differences in their expression were used to test the results of the CIBERSORTx output (*Figure 4.9, Table 4.7*). The genes encoding the cytokine IFNγ and the receptors IL-12 receptor beta subunit 1 (*IL12RB1*) and 2 (*IL12RB2*) were

included due to their expression in Th1 cells. *IL4*, *IL5* and *IL13* genes were used to characterise Th2 cells. Th17 cells were not part of the CIBERSORTx analysis output but are known to be linked to CD pathology. Therefore, the multiple comparison analysis also contained the genes encoding for the Th17 cytokines *IL17A*, *IL21*, *IL22*, and IL-23 subunit alpha (*IL23A*). IL-1ß, TNF-α and IL-6 were included as hallmarks of inflammation and IL-10 as a marker for homeostasis and immune regulation. Actin beta *(ACTB*) is commonly used as a housekeeper gene during PCR analysis of human whole blood derived cDNA. This gene was used as a control in this model to ensure that changes are based on actual differences between the groups and not due to differences in cell-derived RNA quantified within the samples.

IFNG expression was significantly decreased in the LT group compared to the rest of the CD cohort, and *IL12RB2* was lower in the LT patient subset than in the control group (*Figure 4.9A*). Both the cytokine and the receptor are highly expressed by Th1 cells. Additionally, *IL23A*, which drives Th17 proliferation, was reduced in LT patients compared to the other sample groups (*Figure 4.9C*), and the immunoregulatory cytokine *IL10* was significantly increased (*Figure 4.9D*). No changes were found in relation to Th2-associated cytokines, confirming that patients within the LT CD group displayed a reduction of Th1 rather than Th2-polarised cells. Finally, no significant differences were found in *ACTB* gene expression between the sample groups, suggesting that the changes are based on actual differences and not due to differences in cell-derived RNA quantity.

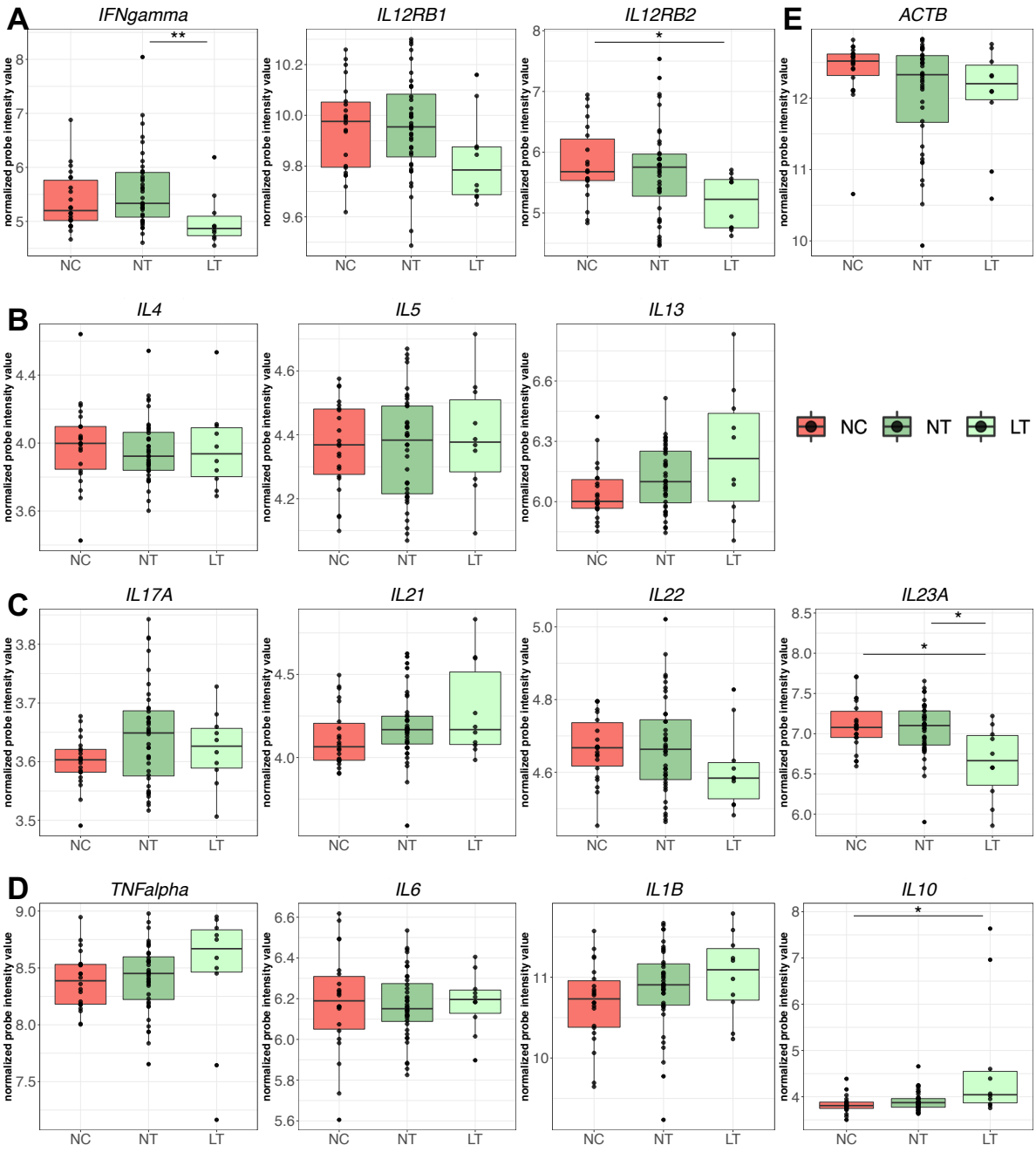


Figure 4.9: ***Comparison of the expression of genes encoding Th1, Th2 and Th17 markers between non-IBD controls and adult CD patients categorised based on their PB T cell profile.***

RNA derived from the whole blood of 48 CD patients and 22 non-IBD controls (NC) was analysed using microarray hybridisation, and normalised probe intensity values were used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =8) and T cell normal (NT, *n* =40). A) Gene expression of Th1-associated cytokines. B) Gene expression of Th2-associated cytokines. C) Gene expression of Th17-associated cytokines. D) Gene expression of inflammation and immunoregulation-associated cytokines. E) Gene expression of *ACTB*. Data used are deposited on NCBI under GSE94648 by Planell *et al*. (2017). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below). *p<0.05, **p<0.01.

Feature	Up	Down	p
IL12RB2	NC	LT	0.0605
IFNG	NT	LT	0.0098
IL23A	NC	LT	0.0194
IL23A	NT	LT	0.0194
IL10	LT	NC	0.0605

Table 4.7: ***Summary of p-values in ANOVA comparing the expression of genes encoding Th1 and Th2 markers between non-IBD controls and adult CD patients categorised based on their PB T cell profile.***

RNA derived from whole blood of 48 CD patients and 22 non-IBD controls (NC) was analysed using microarray hybridisation and normalised probe intensity values used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =8) and T cell normal (NT, *n* =40). A) Gene expression of Th1-associated cytokines. B) Gene expression of Th2-associated cytokines. C) Gene expression of Th17-associated cytokines. D) Gene expression of inflammation and immunoregulation-associated cytokines. E) Gene expression of *ACTB*. Data used are deposited on NCBI under GSE94648 Planell *et al*. (2017). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below)

4.4 Matched PB and intestinal transcriptomics

Since changes in the PB can be an indication of ongoing disease in the intestine of CD patients, I next aimed to study the link between low PB T cells and intestinal inflammation. For this analysis, data from matched whole blood and gut tissue biopsies were required. The collection of matched samples was not possible at the GRI due to limited time in the clinic. After an extensive search on NCBI, I only identified one suitable dataset. A paediatric matched blood and intestinal microarray dataset, which was deposited under GSE126124 by Palmer *et al*. (2019). Ideally, the analysis would have been conducted in an adult CD patient cohort, but data availability limitations did not make this possible. In the paediatric study whole blood of 37 CD patients and 32 non-IBD controls aged between 8 and 18 years were sampled. Matched intestinal biopsies were collected for a subset of paediatric patients and controls (n_{CD} =36 and n_{NC} =18 controls). Additional patient and control information was summarised in *Table 4.8*.

	CD		NC	
	Blood (*n* = 37)	**Tissue (*n* = 36)**	**Blood (*n* = 32)**	**Tissue (*n* = 18)**
Mean Age, years (Range)	12.805 (8.2-18.7)	12.736 (8.2-18.7)	13.55 (8.2-18.5)	13.029 (8.2-18.5)
Standard Deviation Age, years	2.844	2.914	3.346	3.268
F:M ratio	0.378	0.351	0.468	0.474

Table 4.8: ***Summary of cohort information of GSE126124 microarray dataset.***

Whole blood was sampled from 37 paediatric CD patient samples and 32 age and sex-matched non-IBD controls (NC) by Palmer *et al*. (2019). Matched tissue biopsies were collected for 36 patients and 18 controls.

4.4.1 sPLS-DA: Stratification of CD patients and non-IBD controls based on expression of PB T cell-associated genes in paediatric CD

In the analyses of the GSE94648 dataset, sPLS-DA proved to be a useful tool to stratify CD patients based on their T cell-associated gene expression. Therefore, the same analysis was performed on the whole blood derived microarray data of the paediatric 36 CD patients and 32 non-IBD controls (*Figure 4.10*). The sPLS-DA only loaded 16 out of the 32 T cell-associated genes. Additional genes were excluded due to a lack of discriminant power between the sample based on the sPLS-DA. Similar sample clustering was observed in the blood of children compared to adult patients and controls. Based on the sPLS-DA a small group of paediatric CD patients (*n* =10) and a single non-IBD control formed a cluster with distinct lower expression of T cell-associated genes. Unlike the analysis results from adult data, not all the other paediatric patients mixed in with controls and formed two additional subsets: medium and normal T cell gene expression. Overall, the expression of T cell-associated genes showed a gradient in the patient cohort, which reflects the behaviour of adult patients.

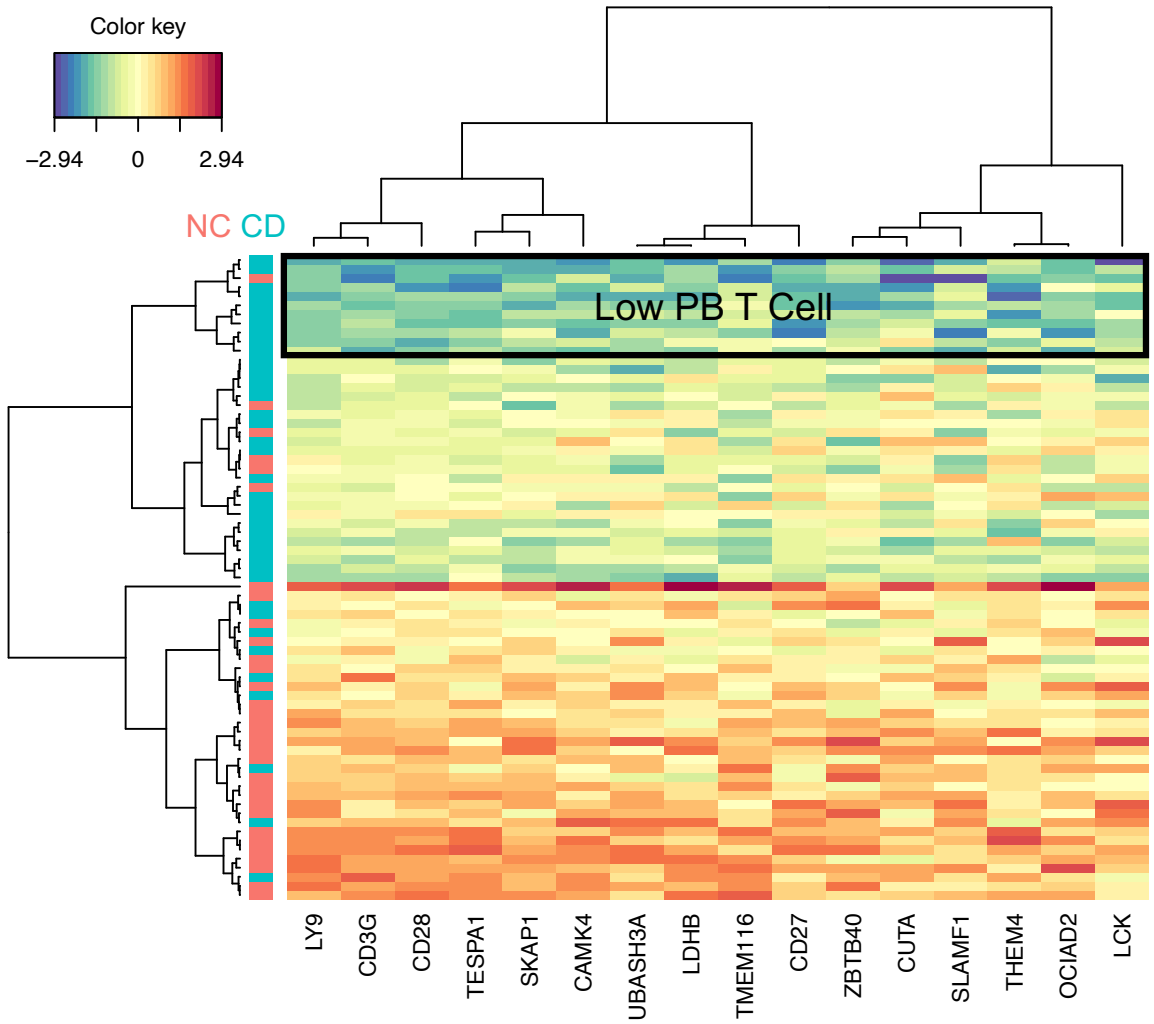


Figure 4.10: ***Sparse least-squares discriminant analysis (sPLS-DA) output comparing PB T cell-associated gene expression levels between paediatric CD and non-IBD controls.***

sPLS-DA analysis elucidates variables (PB T cell-associated genes) with the highest discriminant power to distinguish two sample groups. Sample groups included 22 non-IBD controls (NC, red) and 37 CD patients (CD, blue). Data analysed were in form of normalised probe intensity values. The *splsda()* function from the mixomics package was used for analysis. Data used is deposited on NCBI under GSE126124 by Palmer *et al.* (2019). All analysis was performed in R studio (R version 4.0.2 and below).

4.4.2 Multiple comparison analysis of PB T cell-associated gene expression between paediatric CD subsets and non-IBD controls

Based on the sPLS-DA, CD patients were again stratified into LT and NT subsets. The expression of *CD3D*, *CD3E*, *CD3G*, *CD4*, *CD8A* and *CD8B* encoding genes were compared between patient groups and non-IBD controls (*Figure 4.11*, *Table 4.9*). The expression of all CD3 and CD8 genes were significantly decreased in the LT subset in comparison to both NT patients and non-IBD controls. *CD4* was also expressed less in LT than in non-IBD controls. NT patients showed a reduction in

CD3E, *CD3G* and *CD8B* gene expression when compared to non-IBD controls, supporting the findings of the sPLS-DA, which described NT patients to also be distinct from NCs.

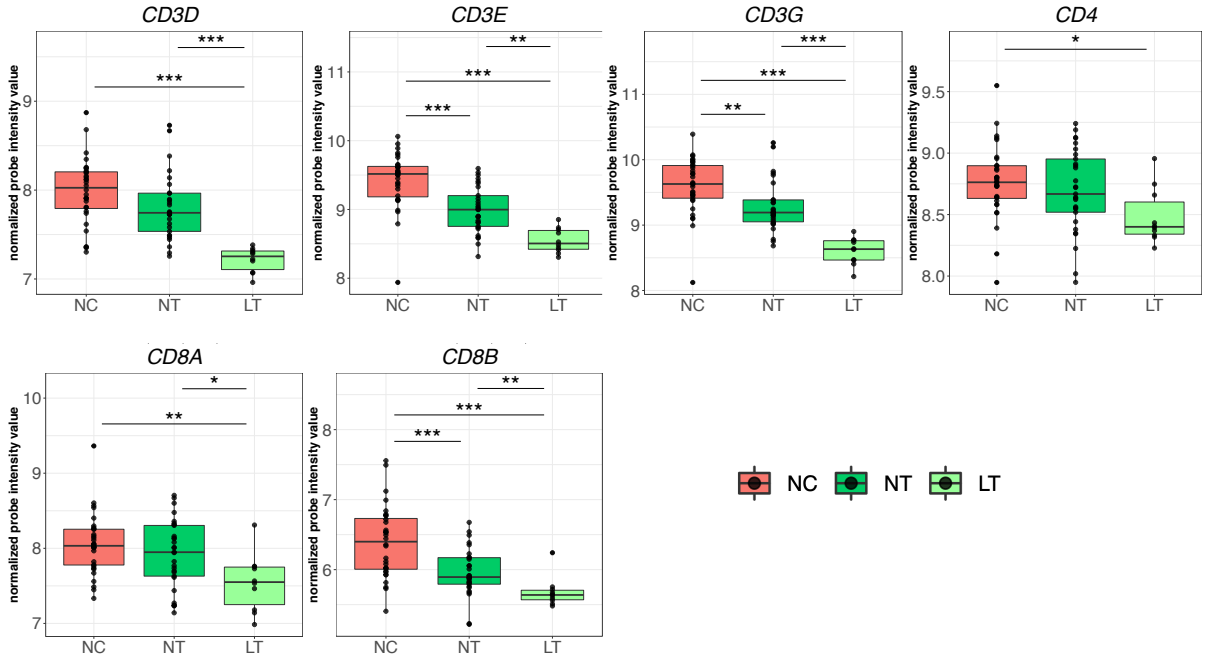


Figure 4.11: ***Comparison of the expression of genes encoding CD3, CD4 and CD8 receptors between non-IBD controls and paediatric CD patients categorised based on their PB T cell profile.***

RNA derived from the whole blood of 37 CD patients and 32 non-IBD controls (NC) were analysed using microarray hybridisation, and normalised probe intensity values were used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =10) and T cell normal (NT, *n* =27). Data used are deposited on NCBI under GSE126124 by Palmer *et al*. (2019). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).*p<0.05, **p<0.01, ***p<0.001.

Feature	Up	Down	p
CD8A	NC	LT	0.0039
CD8A	NT	LT	0.0428
CD8B	NC	LT	<0.0001
CD8B	NC	NT	0.0006
CD8B	NT	LT	0.0017
CD4	NC	LT	0.0353
CD3D	NC	LT	<0.0001
CD3D	NT	LT	<0.0001
CD3E	NC	LT	<0.0001
CD3E	NC	NT	<0.0001
CD3E	NT	LT	0.0028
CD3G	NC	LT	<0.0001
CD3G	NC	NT	0.0057
CD3G	NT	LT	<0.0001

Table 4.9: ***P-values associated with the comparison of the expression of genes encoding CD3, CD4 and CD8 receptors between non-IBD controls and paediatric CD patients categorised based on their PB T cell profile.***

RNA derived from the whole blood of 37 CD patients and 32 non-IBD controls (NC) were analysed using microarray hybridisation, and normalised probe intensity values were used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =10) and T cell normal (NT, *n* =27). Data used are deposited on NCBI under GSE126124 by Palmer *et al*. (2019). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).

For further characterisation of the T cell phenotypes absent from the PB in paediatric LT patients, the multiple comparison analysis shown in *Figure 4.9* was repeated. The expression of *IFNG*, *IL10, IL12RB1, IL12RB2, IL13, IL17D, IL21, IL22, IL23A, IL1B, IL21, IL4, IL5, IL6* and *TNF* analysed in CD subsets and non-IBD controls (*Figure 4.12*, *Table 4.10*). *ACTB* was again included as the housekeeper gene. Unlike adult patients, only the NT subsets showed a decrease in the Th1-associated cytokines IFNG and IL12RB2 compared to controls. The LT subset had significant increases in *IL17D* when compared to the NC group, in *IL21* when compared to the NT group and in *IL22* compared to both. Both patient subsets exhibited less gene expression of *IL23A* than controls. Only LT patients had elevated expression of *IL1B* in comparison to the control group. These results suggested that even though the PB of LT paediatric patients expressed lower levels of PB T cell-associated genes, they did not display a decrease in Th1-associated genes when compared to

the NT phenotype as it was observed in adults. However, the LT patient subset appeared to have more expression of genes encoding for cytokines encoding for Th17 cells, suggesting a higher frequency of this Th phenotype in their PB.

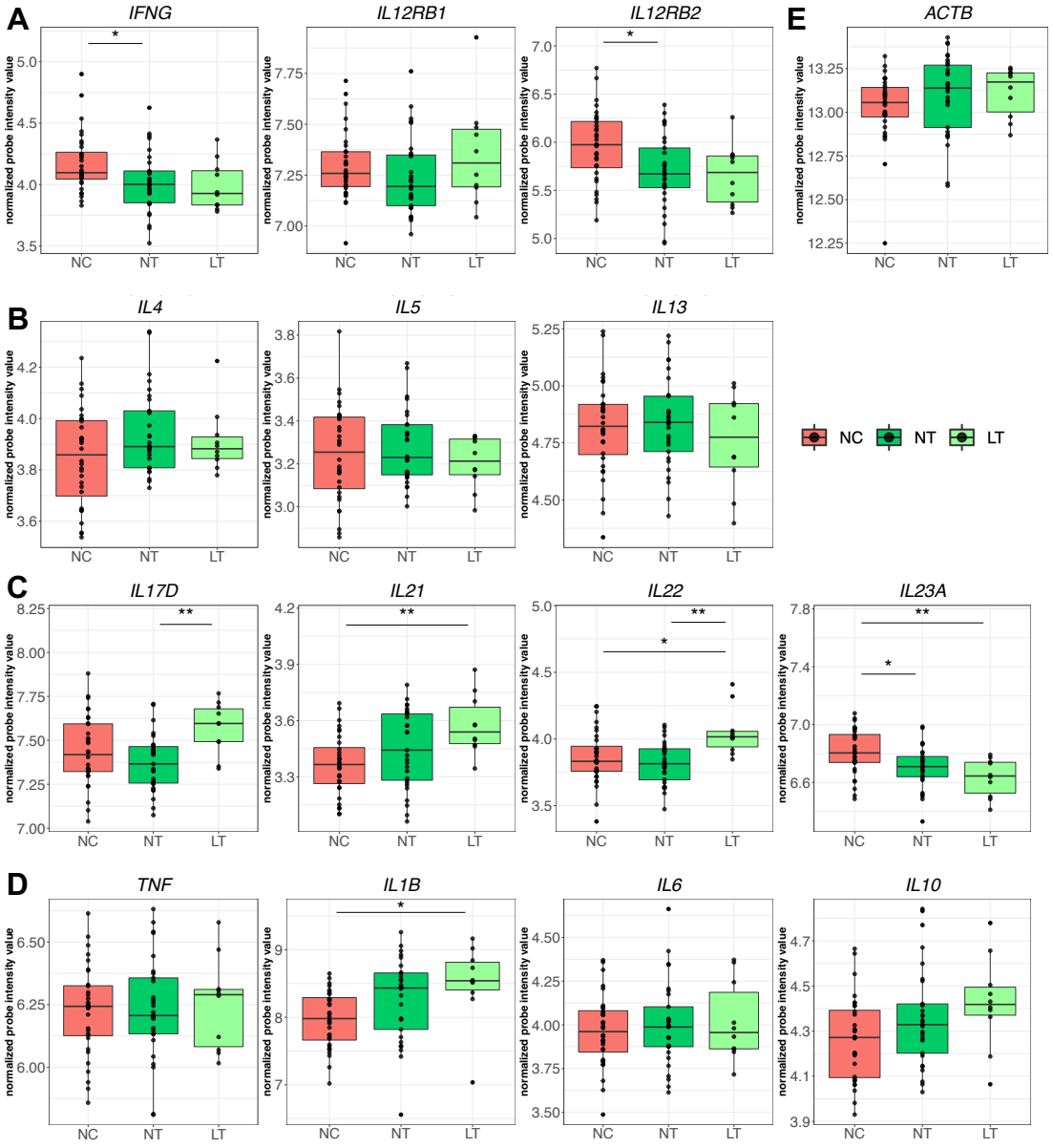


Figure 4.12: ***Comparison of the expression of genes encoding Th1, Th2 and Th17 markers between non-IBD controls and paediatric CD patients categorised based on their PB T cell profile.***

RNA derived from the whole blood of 37 CD patients and 32 non-IBD controls (NC) were analysed using microarray hybridisation, and normalised probe intensity values were used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =10) and T cell normal (NT, *n* =27). Data used are deposited on NCBI under GSE126124 by Palmer *et al*. (2019). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).*p<0.05, **p<0.01.

Feature	from	to	p
IFNG	NC	NT	0.0435
IL12RB2	NC	NT	0.0161
IL17D	LT	NT	0.0088
IL21	NC	LT	0.0056
IL22	LT	NC	0.0113
IL22	LT	NT	0.0021
IL23A	NC	NT	0.0211
IL23A	NC	LT	0.0042
IL1B	LT	NC	0.0162

Table 4.10: ***P-values associated with the comparison of the expression of genes encoding Th1 and Th2 markers between non-IBD controls and paediatric CD patients categorised based on their PB T cell profile.***

RNA derived from the whole blood of 37 CD patients and 32 non-IBD controls (NC) were analysed using microarray hybridisation, and normalised probe intensity values were used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =10) and T cell normal (NT, *n* =27). Data used are deposited on NCBI under GSE126124 by Palmer *et al*. (2019). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).

4.4.3 Multiple comparison analysis of T cell-associated genes between LT and NT CD subsets and non-IBD controls in intestines of children

The main aim of analysing the GSE126124 data was to investigate changes in the intestines of CD patients with distinct PB T cell profiles. Patients were again categorised into LT and NT, and the intestinal expression of the genes encoding CD3, CD4 and CD8 was compared between paediatric CD patients and non-IBD controls (*Figure 4.13*). No differences were detected apart from an increase in *CD4* expression in NT patients compared to controls (p =0.0353). Additionally, differential gene expression analysis did not show any DEGs between the two CD subsets. The intestine harbours a large selection of immune cells, including tissue-resident T cells, keeping commensal bacteria in check constantly. I, therefore, hypothesised that this ongoing activity might mask the differences in T cell genes in the intestines between the CD subsets.

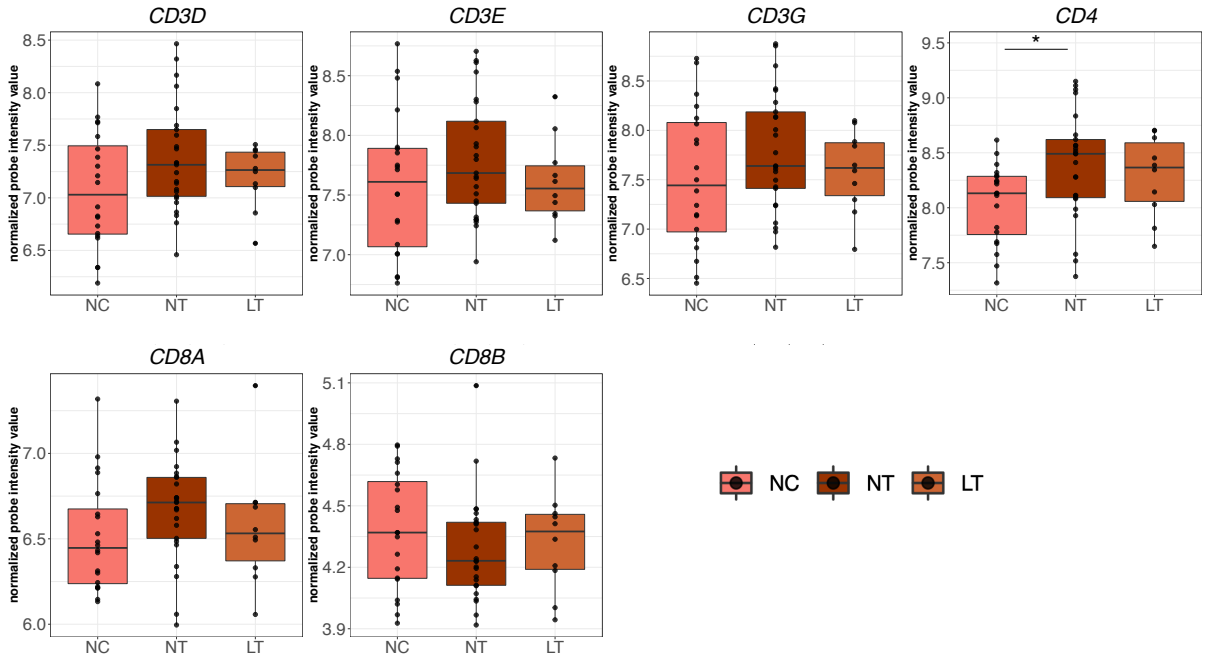


Figure 4.13: ***Comparison of the intestinal expression of genes encoding CD3, CD4 and CD8 receptors between non-IBD controls and paediatric CD patients categorised based on their PB T cell profile.***

RNA derived from intestinal biopsies of 36 CD patients and 18 non-IBD controls (NC) were analysed using microarray hybridisation, and normalised probe intensity values were used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =10) and T cell normal (NT, *n* =26). Data used are deposited on NCBI under GSE126124 by Palmer *et al*. (2019). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below). *p<0.05.

4.4.4 sPLS-DA: Stratification of paediatric patients based on the expression of intestinal CD-specific genes

In order to distinguish genes that were specific to CD from those associated with general immunity in the gut, I performed differential gene expression analysis between controls and patients in the tissue dataset. 6674 DEGs were identified and subsequently used to load a sPLS-DA (*Figure 4.14*). Since the genes significantly different between CD and NC were used as discrimination variables, only patients were included in the analysis. 58 DEGs were selected by the sPLS-DA algorithm to distinguish between the sample groups. Hierarchical clustering identified two distinct patient subsets, named CD1 and CD2. CD1 included all 9 LT patients that were sampled during endoscopy and 10 NT patients. The other 18 individuals from the NT cohort were part of the CD2 subset.

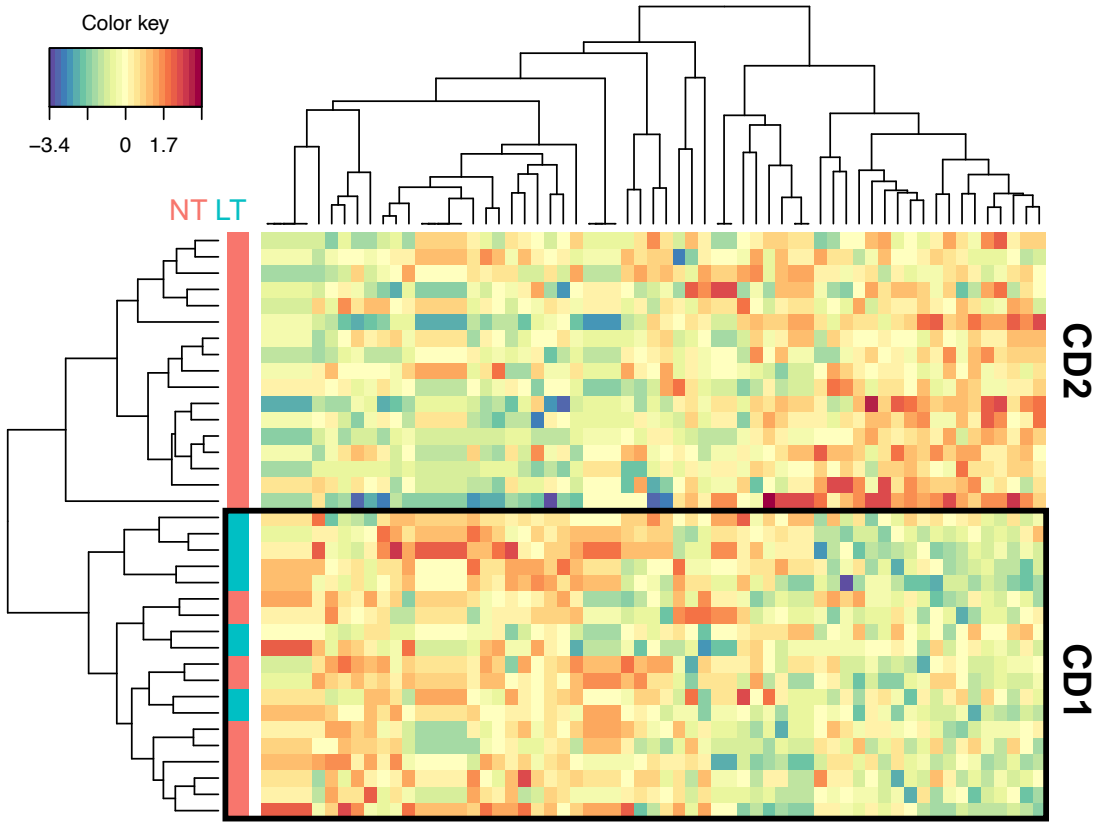


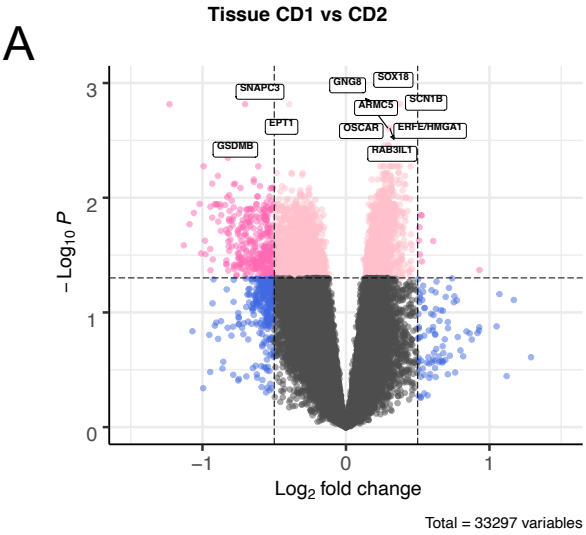
Figure 4.14: ***Sparse least-squares discriminant analysis (sPLS-DA) output CD-associated gene expression levels between low and normal PB T cell subsets in intestinal tissue.***

sPLS-DA analysis elucidates variables (DEGs between CD and NC) with highest discriminant power to distinguish two sample groups. Sample groups included 27 normal (NT, red) and 9 low PB T cell patients (LT, blue). Data analysed were in form of normalised probe intensity values. The *splsda()* function from the mixomics package was used for analysis. Data used is deposited on NCBI under GSE126124 by Palmer *et al*. (2019). All analysis was performed in R studio (R version 4.0.2 and below).

4.4.5 Intestinal differential gene expression analysis: CD1 vs CD2 paediatric patients

All LT patients belonged to the same CD subset, suggesting that their PB T cell profile was associated with the intestinal immunopathotype of group CD1. To understand whether CD1 and CD2 patients showed differences in their intestinal signature, I performed a differential gene expression analysis, which identified 3514 DEGs (*Figure 4.15*).

A



B

Gene	adj.P.Val	logFC
SNAPC3	0.00153	-3.95E-01
ARMC5	0.00153	2.56E-01
SOX18	0.00153	3.53E-01
SCN1B	0.00153	3.78E-01
EPT1	0.00231	-4.56E-01
ERFE/HMGA1	0.00244	3.04E-01
GNG8	0.00251	2.96E-01
RAB3IL1	0.00348	2.97E-01
GSDMB	0.00348	-7.61E-01
OSCAR	0.00348	2.95E-01

Figure 4.15: ***Differential gene expression analysis comparing intestinal gene expression of CD1 and CD2 subsets.***

Volcano plot showing differential gene expression, highlighting DEGs (*adj.p* <0.05) in rose and genes with -0.5< log2 fold change >0.5 in blue. Genes with -0.5< log2 fold change >0.5 and p <0.05 were shown in pink. B) Table with top 10 DEGs according to the adjusted p-value. Differential gene expression analysis was performed using the GEO2R tool developed by NCBI. Data used is deposited on NCBI under GSE126124 by Palmer *et al*. (2019). All analysis was performed in R studio (R version 4.0.2 and below) and required the clusterProfiler package. *SNAPC3:* Small nuclear RNA activating complex polypeptide 3. *ARMC5:* Armadillo repeat containing 5. *SOX18:* SRY-box transcription factor 18. *SCN1B:* Sodium voltage-gated channel beta subunit 1. *EPT1:* Selenoprotein I. *ERFE/HMGA1:* Erythroferrone. *GNG8:* G protein subunit gamma 8. *RAB3IL1:* RAB3A interacting protein like 1. *GSDMB*: Gasdermin B. *OSCAR:* Osteoclast associated Ig-like receptor.

Since I was only interested in understanding differences in intestinal immune responses, a cluster of genes related to the GO immune system process i.e., all genes related to the immune system and its development (EMBL-EBI, 2018), were selected, resulting in a list of 365 genes. GO enrichment analysis (Yu G, Wang L, Han Y, 2012) was completed on this gene set (Figure 4.16A/B). The top five GOs enriched were: T cell activation, mononuclear cell differentiation, regulation of hemopoiesis, myeloid cell differentiation and regulation of immune effector process. 22 genes within the T cell activation GO were downregulated in CD1 (AP3B1, CASP3, CEACAM1, CHD7, D2G1, DROSHA, HMGB1, HSPD1, IL7, KIF13B, LGALS9, LGALS9B, LGALS9C, MTOR, MYB, PAK1, PRKDC, RC3H1, SOCS6, SOD1, TOX, TP35) and 25 genes upregulated (BATF, CD7, CD8B, CDBPB, CLEC4G, EBI4, EOMES, FOXP3, GON4L, HLX, ICAM1, LAG3, LEP, LMO1, MAPK8IP1, NLRC3, NOD2, PAX1, PRR7, RELB, RUNX3, SART1, TNFRSF4, TNFSF14, ZFPM1) compared to CD2.


A
GOBP_T_CELL_ACTIVATION
GOBP_MONONUCLEAR_CELL_DIFFERENTIATION
GOBP_REGULATION_OF_HEMOPOIESIS
GOBP_MYELOID_CELL_DIFFERENTIATION
GOBP_REGULATION_OF_IMMUNE_EFFECTOR_PROCESS
GOBP_NEGATIVE_REGULATION_OF_IMMUNE_SYSTEM_PROCESS
GOBP_REGULATION_OF_RESPONSE_TO_BIOTIC_STIMULUS
GOBP_REGULATION_OF_LEUKOCYTE_DIFFERENTIATION
GOBP_LYMPHOCYTE_ACTIVATION_INVOLVED_IN_IMMUNE_RESPONSE
GOBP_SOMATIC_DIVERSIFICATION_OF_IMMUNE_RECEPTORS
20
30
40
Count
GeneRatio
0.075
0.100
0.125
p.adjust
1e−13
2e−13
3e−13
4e−13
B
fold change
0.0
−0.5
C
Tuberculosis
Proteoglycans in cancer
Neurotrophin signaling pathway
Natural killer cell mediated cytotoxicity
MAPK signaling pathway
Lipid and atherosclerosis
Human papillomavirus infection
Fluid shear stress and atherosclerosis
Epstein–Barr virus infection
Antigen processing and presentation
fold change
0.4
0.2
0.0
−0.2
−0.4
−0.6
ACVR2A
ANGPT4
APAF1
ATP6V0A1
BMPR1A
CACNA1C
CACNA1F
CALML5
CAMK2D
CAMX
CARD9
CASP3
CD8B
CD8B2
CEBPB
COL2A1
CYP27B1
DLG1
DROSHA
EZR
FOXO3
FZD8
HDAC1
HES5
HSP90AB1
HSPD1
ICAM1
IRF7
ITGA6
JAG1
KIR2DL1
KIR2DL3
KIR2DS4
KIR3DL1
KITLG
KLF2
MAPK1
MAPK8IP1
MECOM
MTOR
NCF1
NCR1
NF1
NFATC1
NOD2
PAK1
PDIA3
PGF
PIK3CB
PRKACG
PSMD1
PSMD11
PTK2
RAB5B
RELB
RHOA
RPS6KA5
RUNX3
SH2B2
SKP2
SPHK1
TLR3
TP53
TP73
TRAF2
TRPV4
UBR4
VEGFB
WNT10B

152

Figure 4.16: ***Gene set enrichment analysis of DEGs related to GO immune system process.***

A+B) Gene ontology analysis was performed with the *enrichGO()* function using the Ensembl database as reference. A) Barplot showing the number of genes in each GO, their p-value and GeneRatio. B) Heatplot showing the fold change of genes in the GO analysis. Upregulation and downregulation refer to CD1 subset. C) Heatplot showing the fold change of genes in enriched KEGG pathway using the *enrichKEGG()* function. Data used is deposited on NCBI under GSE126124 by Palmer *et al.* (2019). All analysis was performed in R studio (R version 4.0.2 and below) and required the clusterProfiler package.

Differences in T cell activation between the CD1 and CD2 subsets were studied by looking at the genes enriched in the distinct subsets. The gene responsible for transcription for the CD8 co-receptor beta chain (*CD8B*) was amongst the upregulated T cell activation genes in the CD1 subset. Additionally, the transcription factor EOMES, which has been linked to memory T cell differentiation of $CD8^+$ T cell activation and differentiation in response to viral infection, had elevated gene expression in CD1 patients (Pearce *et al.*, 2003), as did *FOXP3*, commonly known to be highly expressed in T cells with regulatory/immunosuppressive function. The Treg population may expand to counteract the aberrant inflammation. The nod-like receptor family CARD domain containing 3 (NLRC3) has immunosuppressive capacities for T cells, specifically in suppressing autoreactive T cells (Hu *et al.*, 2018; Fu *et al.*, 2019) and genetic CD risk factor *NOD2*, known to play a major role in antibacterial protection, were also increased in the CD1 group. In terms of cell migration, CD1 patients had more *ICAM1* expression, a receptor involved in leukocyte recruitment to the gut (Reinisch *et al.*, 2018; Kuhbandner *et al.*, 2019). The kinase, mammalian target of rapamycin (*MTOR*), activated during T cell activation and differentiation, was one of the 22 downregulated genes in the CD1 subset compared to CD2. The *MTOR signalling* cascade can be induced by TCR and cytokines such as IL-7, which is responsible for maintaining T cell survival (Rathmell *et al.*, 2001) and displayed lower gene expression in CD1 compared to CD2 patients as well.

After identifying enriched GOs, I performed KEGG analysis to link pathways to the gene set (*Figure 4.16C*). Eight pathways were significantly associated fitting into the following categories: atherosclerosis *(lipid and atherosclerosis* and *fluid shear stress and atherosclerosis*), Th1/Th17 immunity (*Tuberculosis* and *Epstein-Barr virus infection*), general pro-inflammatory immune response (*antigen processing and presentation* and *MAPK signaling pathway*) and *NK cell mediated cytotoxicity*. The gene encoding for Gasdermin-B (*GSDMB*), which has been linked to cytotoxic activity of lymphocytes in cancer (Zhou *et al.*, 2020), was one of the top 10

upregulated genes in the CD1 patient subset (*Figure 4.15*). Overall, these findings suggested an increase in cytotoxic and Th1 effector functions in CD1 patients compared to CD2, potentially supported by $CD8^+$ T and $CD4^+$ Th1 cell migration from PB to the intestine. This infiltration may be facilitated *via* the adhesion molecule ICAM1.

4.5 Association of reduced PB T cell CD subset with clinical parameters

Investigating differences in the intestine of patients with low PB T cells was an essential step to understanding the contribution of this blood phenotype to disease. However, peripheral blood biomarkers for clinical characteristics, such as inflammation, treatment response and phenotype, are currently missing in CD. Therefore, potential links between CD patient PB and clinical data, such as inflammatory markers, current treatment, and disease phenotype, were studied.

4.5.1 Differences in blood inflammatory markers and CD disease severity scores between adult LT and NT CD patients

As shown in *Table 3.1,* patients sampled for flow cytometry were simultaneously screened for inflammatory blood markers, including total WBC, neutrophil and lymphocyte count, as well as CRP levels. Additionally, CRP and percentages of neutrophils, monocytes and lymphocytes were measured in 30 of the 48 patients in the GSE94648 adult CD gene expression cohort. HBI was used to measured disease severity in the flow cytometry cohort and CDEIS in the GSE94648 dataset.

When looking at the flow cytometry cohort, the comparison of inflammatory blood markers between patients with NT and LT patients unsurprisingly displayed a significant decrease in the lymphocyte count in the latter subset *(*$p = 0.0198$, *Figure 4.17*). No changes were seen in any of the other clinical parameters. When looking at the GSE94648 data, additional differences were detected. LT patients had an increase in percentage neutrophils ($p = 0.0001$), as well as CRP ($p = 0.0002$) and CDEIS ($p = 0.0024$, *Figure 4.18*), suggesting higher disease activity.

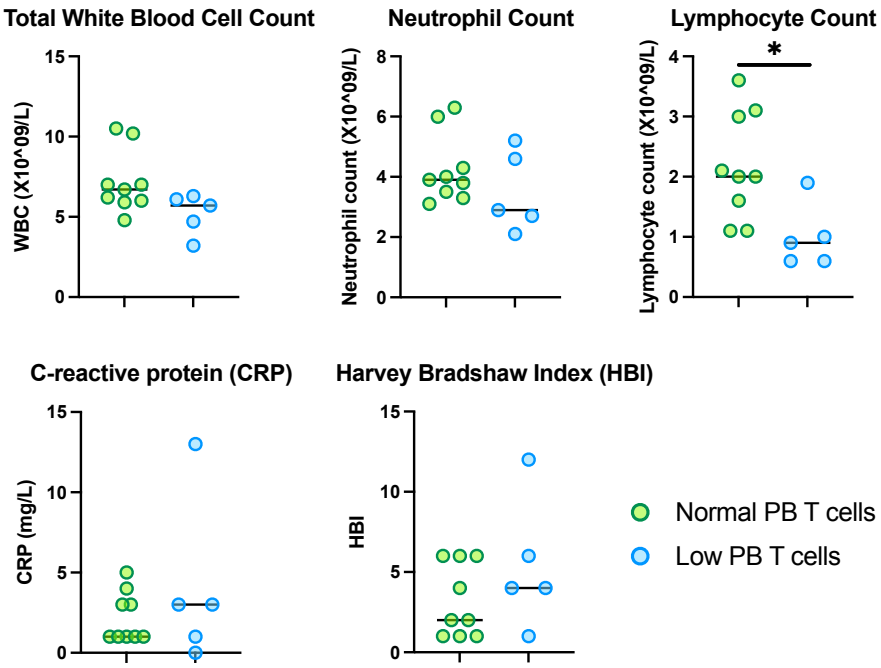


Figure 4.17: ***Comparison of inflammatory markers between normal and low PB T cell CD patients in flow cytometry cohort.***

The total white blood, neutrophil, and lymphocyte count, CRP and HBI were compared between 9 normal (green) and 5 low (blue) PB T cell patients. The bar represents the mean. Unpaired T-Test or Mann-Whitney was performed depending on the distribution of the data. Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below). *p<0.05.

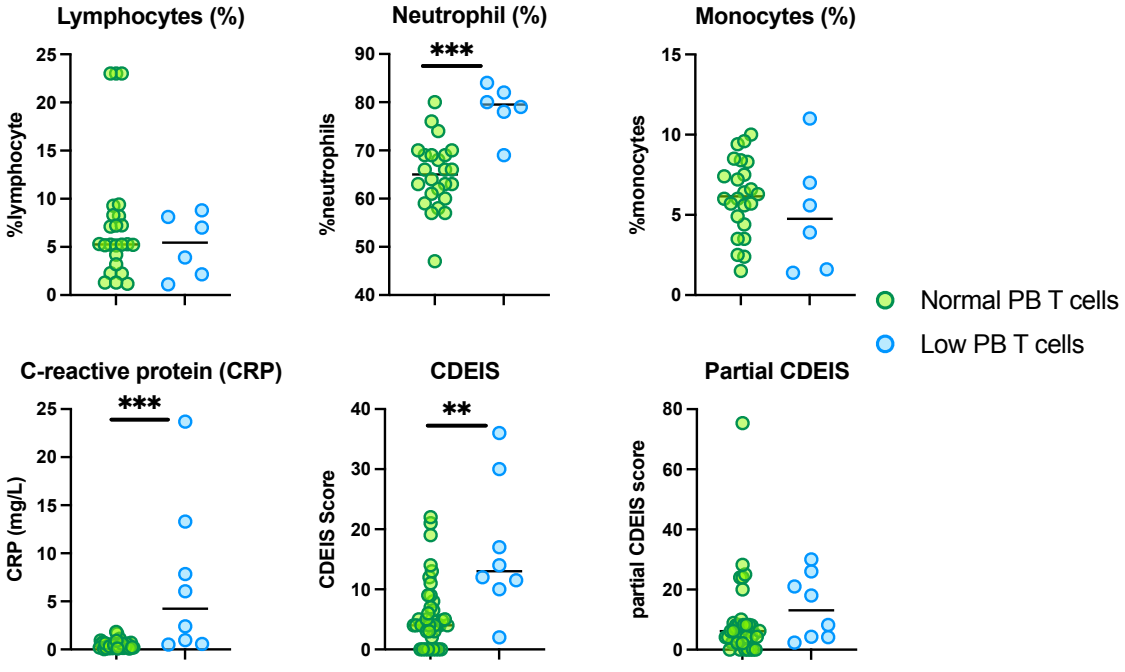


Figure 4.18: ***Comparison of inflammatory markers between normal and low PB T cell CD patients in GSE94648 dataset.***

Percentage lymphocytes, neutrophil and monocytes, CRP, CDEIS and partial CDEIS were compared between 40 normal (green) and 8 low (blue) PB T cell patients. The bar represents the mean. Unpaired T-Test or Mann-Whitney was performed depending on the distribution of the data. Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below). **p<0.01, ***p<0.001.

Figure 4.19 and *Figure 4.20* summarise the current treatment and disease phenotype categorised by Montreal classification in the flow cytometry and GSE94648 cohorts, respectively. The two patient subsets received a diverse range of immune inhibiting medications. However, anti-TNF and 5-ASA drugs were only taken by NT patients in both datasets. Regarding CD phenotype, GSE94648 patient information only included the Montreal classifications for disease behaviour and location but not the age of onset. The percentages of patients with inflammatory, stricturing or penetrating disease were similar between NT and LT in the microarray cohort. The disease behaviour in the flow cytometry patient cohort was less diverse, with 12 out of 14 exhibiting solely inflammatory CD. Therefore, it was unsurprising that all LT patients belonged to this category. When comparing disease location, colonic patients (category L2) were more frequent in LT for both datasets than NT.

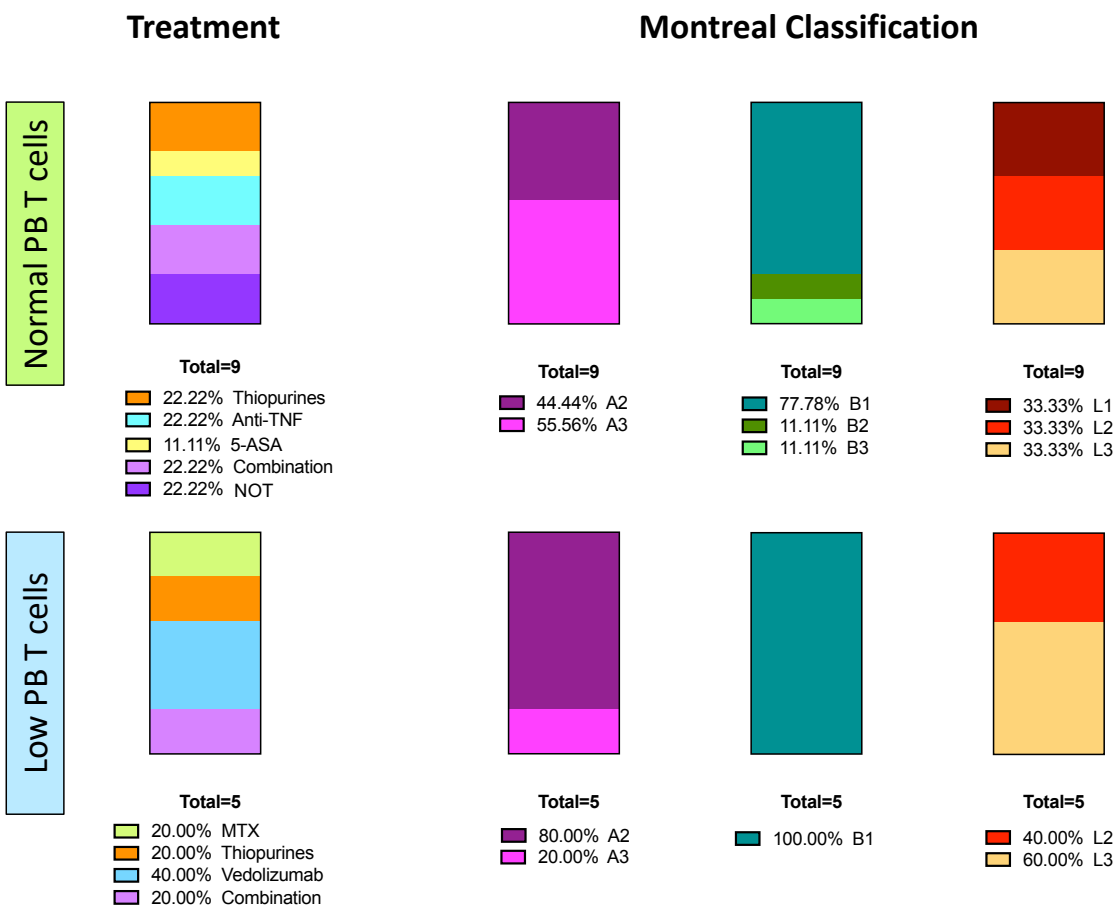


Figure 4.19: ***Summary of current treatment and Montreal classification in flow cytometry cohort.***

All categories were displayed as percentages in normal (green) and low (blue) PB T cell subsets. Treatment categories included were thiopurines, anti-TNF, 5-ASA, combination therapies *i.e., more than one therapy administered), methotrexate (MTX), vedolizumab and no treatment (NOT). Montreal classification included age-of-onset (A2: 17-40 years, A3: 41-70 years), disease behaviour (B1: inflammatory, B2: stricturing, B3: penetrating) and disease location (L1: ileal, L2: colonic, L3: ileocolonic). Data was visualised in GraphPad Prism (Version 9.3.1 and below).

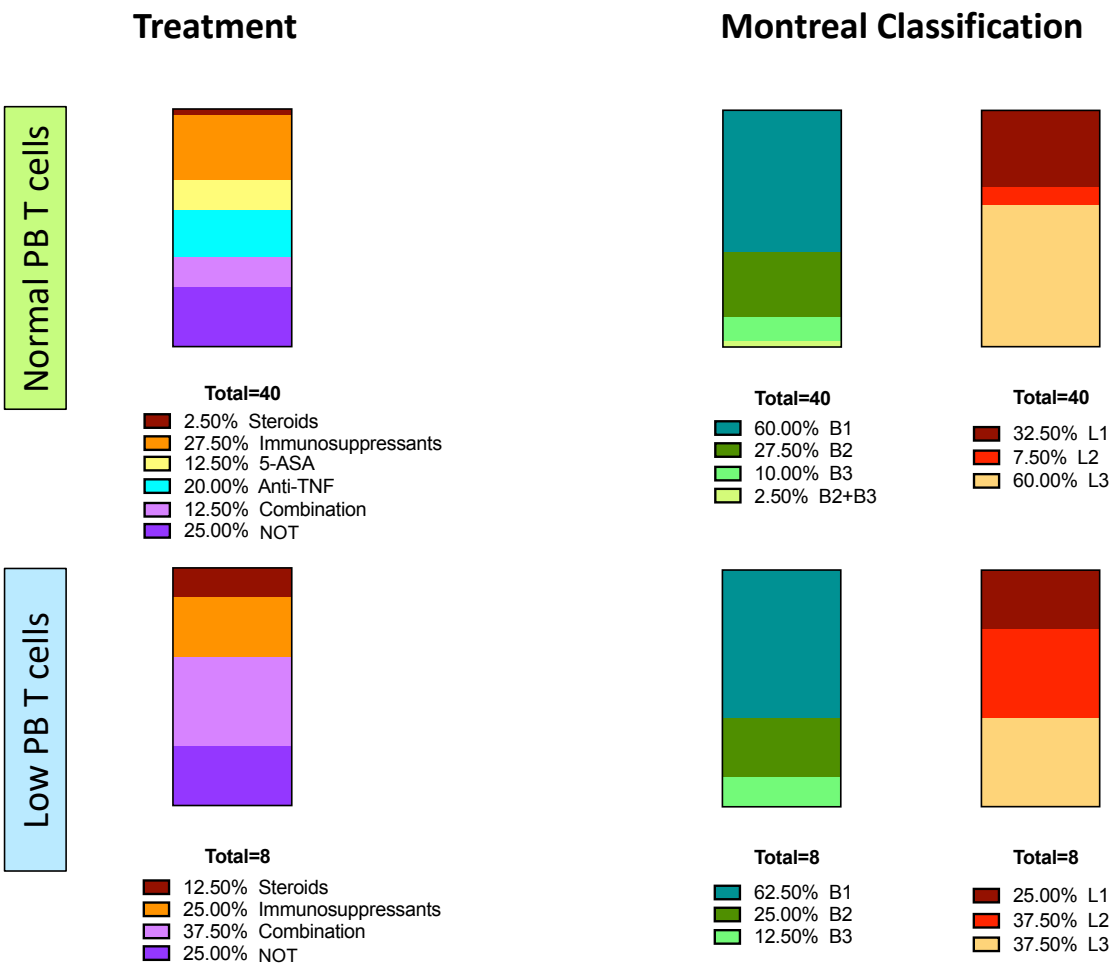


Figure 4.20: ***Summary of current treatment and Montreal classification in GSE94648 dataset.***

All categories were displayed as percentages in normal (green) and low (blue) PB T cell subsets. Treatment categories included were steroids, immunosuppressants, 5-ASA, anti-TNF, combination i.e., more than one therapy administered) and no treatment (NOT). Montreal classification included disease behaviour (B1: inflammatory, B2: stricturing, B3: penetrating, B2+B3: stricturing and penetrating) and disease location (L1: ileal, L2: colonic, L3: ileocolonic). Data was visualised in GraphPad Prism (Version 9.3.1 and below).

4.6 Migration potential of PB T cells in CD patients

Based on the findings in this chapter, I formulated the hypothesis that Th1 and $CD8^+$ T cells leave the PB and travel to the intestine to contribute to inflammation, a process enhanced in LT patients. To understand the mechanism T cells use to travel from the blood to the intestine, the gene expression of chemokines and their receptors in whole blood of both adult (GSE94648 data) and paediatric (GSE12426) CD patients were compared between the LT and NT subsets.

4.6.1 Multiple comparison analysis of gut-homing marker gene expression between LT and NT CD subsets and non-IBD controls in the blood of adult and paediatric patients

The expression of the genes encoding the gut-homing markers chemokine receptor type 6 (*CCR6*), chemokine receptor type 9 (*CCR9*) and α4β7 integrin (*ITGA4* and *ITGB7*) was compared between non-IBD controls, LT, and NT CD patient subsets. The skin-homing marker Cutaneous Lymphocyte-Associated Associated Antigen (*SELPLG*) was used as a control of general migratory potential. Since *ICAM1* expression was increased in CD1 patients, its ligands integrin subunit beta 2 (*ITGB2*), lymphocyte function-associated antigen 1 (*ITGAL*) and macrophage-antigen 1 (*ITGAM*) were also analysed.

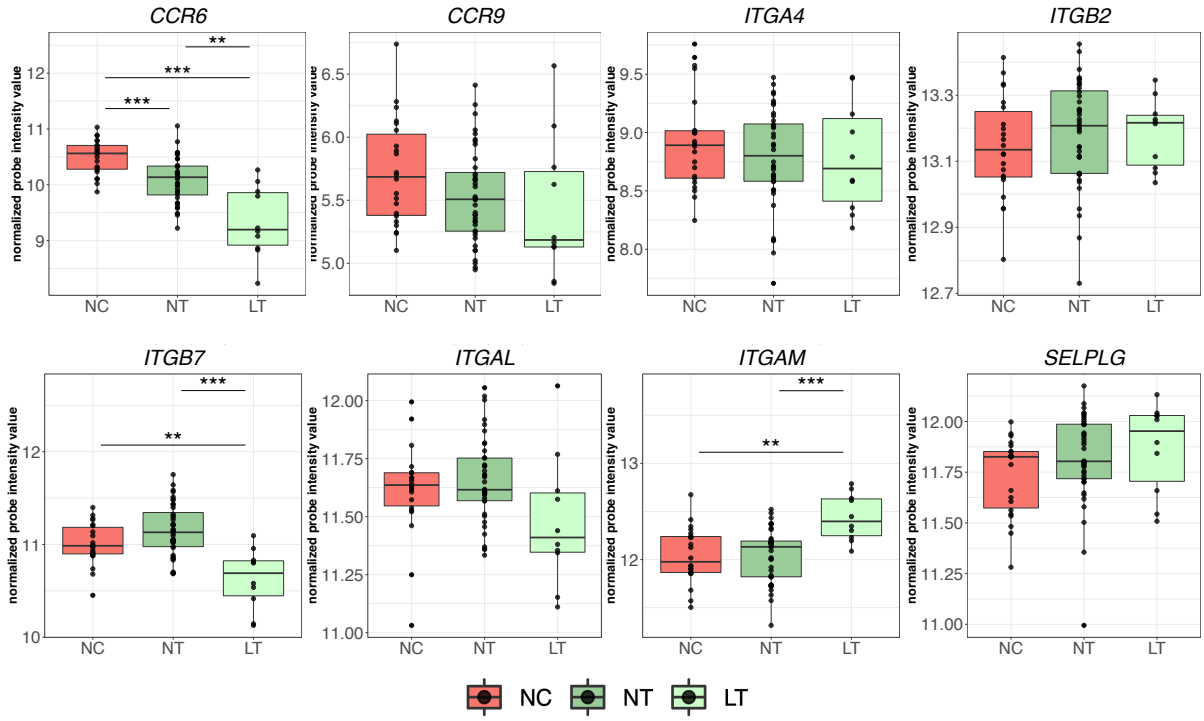


Figure 4.21: ***Comparison of the expression of genes encoding homing markers between non-IBD controls and adult CD patients categorised based on their PB T cell profile.***

RNA derived from the whole blood of 48 CD patients and 22 non-IBD controls (NC) was analysed using microarray hybridisation, and normalised probe intensity values were used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =8) and T cell normal (NT, *n* =40). Data used are deposited on NCBI under GSE94648 by Planell *et al*. (2017). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below). **p<0.01, ***p<0.001.

Feature	Up	Down	p
ITGB7	NC	LT	0.0016
ITGB7	NT	LT	<0.0001
CCR6	NC	LT	<0.0001
CCR6	NC	NT	0.0018
CCR6	NT	LT	0.001
ITGAM	LT	NC	0.001
ITGAM	LT	NT	0.0005

Table 4.11: ***P-value associated with the comparison of the expression of genes encoding homing markers between non-IBD controls and CD patients categorised based on their PB T cell profile.***

RNA derived from the whole blood of 48 CD patients and 22 non-IBD controls (NC) was analysed using microarray hybridisation, and normalised probe intensity values were used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =8) and T cell normal (NT, *n* =40). Data used are deposited on NCBI under GSE94648 by Planell *et al*. (2017). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).

In adult patients of the GSE94648 dataset, *CCR6*, *ITGB7* and *ITGAL* were significantly decreased in the PB of LT compared to NT and non-IBD controls (*Figure 4.23, Table 4.11*). Paediatric patients showed the same changes in chemokine expression and an additional reduction in the expression of *ITGA4* (*Figure 4.24, Table 4.12*). *ITGAM* was increased in LT CD patients compared to controls and the NT subset. Both paediatric CD subsets displayed an increase in skin homing *SELPLG*. Paediatric NT patients had reduced expression levels of *CCR9* when compared to non-IBD controls. Based on these findings, I hypothesised that CCR6 and α4β7 integrin-mediated migration might contribute to the migration of PB cells to the intestine in patients with low peripheral blood T cells.

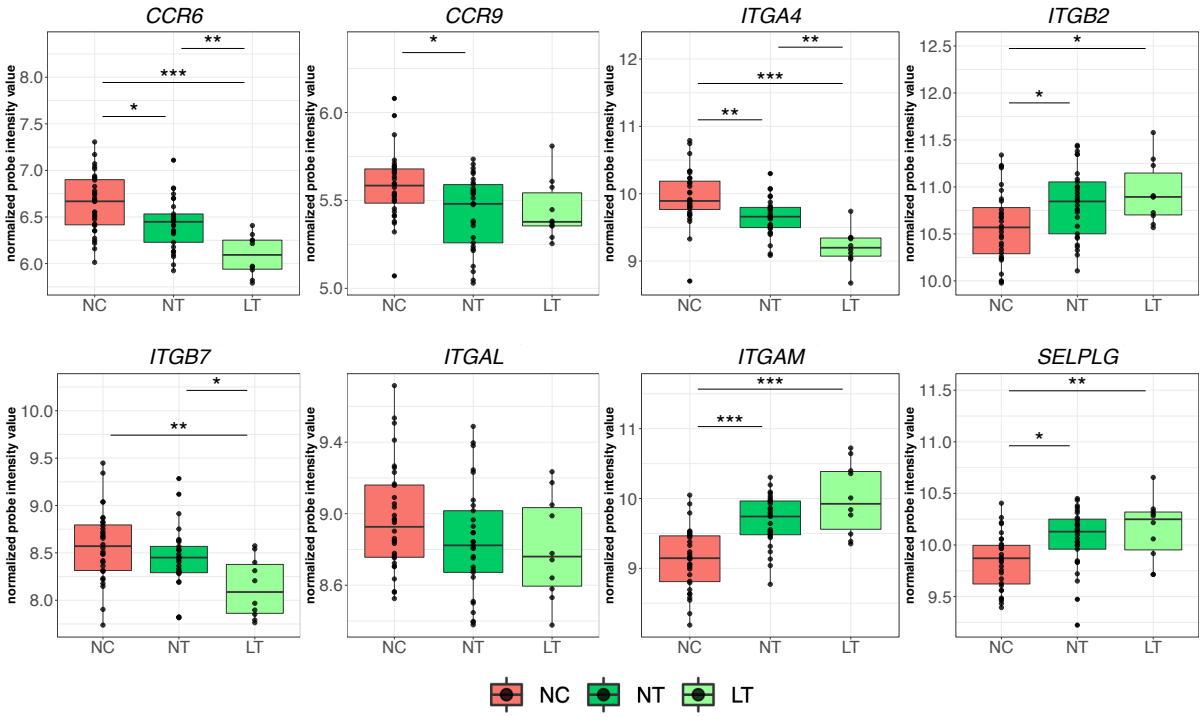


Figure 4.22: ***Comparison of the expression of genes encoding homing markers between paediatric non-IBD controls and paediatric CD patients categorised based on their PB T cell profile.***

RNA derived from whole blood of 36 CD patients and 18 non-IBD controls (NC) were analysed using microarray hybridisation and normalised counts used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =10) and T cell normal (NT, *n* =26). Data used are deposited on NCBI under GSE126124 by Palmer *et al*. (2019). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below). *p<0.05, **p<0.01, ***p<0.001.

Feature	from	to	p
ITGAM	NC	LT	6.3E-06
ITGAM	NC	NT	2.4E-06
ITGAM	LT	NT	0.039
ITGA4	NC	LT	<0.0001
ITGA4	NC	NT	0.0027
ITGA4	NT	LT	0.0017
ITGB7	NC	LT	0.0015
ITGB7	LT	NT	0.013
CCR6	NC	LT	0.0064
CCR6	NT	LT	<0.0001
CCR9	NC	NT	0.0118
SELPLG	LT	NC	0.0091
SELPLG	NT	NC	0.0101
ITGB2	LT	NC	0.0239
ITGB2	NT	NC	0.0304

Table 4.12: ***P-value associated with the comparison of the expression of genes encoding homing markers between paediatric non-IBD controls and CD patients categorised based on their PB T cell profile.***

RNA derived from whole blood of 48 CD patients and 22 non-IBD controls (NC) was analysed using microarray hybridisation and normalised counts used for analysis. CD patients were categorised based on PB T cell levels into T cell low (LT, *n* =8) and T cell normal (NT, *n* =40). Data used are deposited on NCBI under GSE126124 by Palmer *et al*. (2019). Data were tested for normal distribution and equal variance and subsequently analysed using ANOVA, Welsh ANOVA or Kruskal-Wallis test. Multiple comparison was performed with Tukey's, Benjamini-Hochberg, or Wilcoxon rank sum post-tests. All statistical analysis was performed in *R studio* (R version 4.0.2 and below).

4.7 Discussion

In this chapter, I aimed to further characterise the CD patients with low circulating T cells. I investigated their specific T cell phenotypes, intestinal immunopathotypes, clinical characteristics and gut-homing marker gene expression profile.

Analysis of the flow cytometry cohort after stratifying CD patients into LT and NT subsets showed that most PB T cells were significantly fewer in the LT patient subset (*Figure 4.1*). All leukocytes (i.e., $CD45^+$ cells) and $CD3^+$ cells were decreased in LT compared to NT patients and controls. $CD3^-$ live cells, comprising most non-T cell populations apart from NKT cells, did not significantly differ.

These findings reflect the results of the DIABLO analysis in *Figure 3.16*, which found a reduction of a majority of PB T cells within the LT patient subset. The only T cell subpopulations in the multiple comparison analysis that did not display a difference were effector $CD4^+$ and CM $CD8^+$ T cells (*Figure 4.1*). Naïve $CD4^+$ T cells have yet to be polarised, whereas their effector counterparts already have a distinct Th phenotype. Not every Th cell subset might contribute to CD patients' intestinal inflammation. It is, therefore, possible that only a subpopulation of $CD4^+$ effector T cells was leaving the PB; however, the flow cytometry panel used for this analysis did not allow to distinguish between Th subsets.

The CD phenotype with the LT phenotype was also found in an additional whole blood-derived transcriptomic dataset. CD3, CD4 and CD8-encoding gene expression were reduced in LT compared to NT patients (*Figure 4.7*). Similarly, multiple comparison analyses of the CIBERSORTx outputs, which estimated fractions of immune cells per sample, showed a decrease in $CD8^+$ and Th1-polarised $CD4^+$ T cells (*Figure 4.10*). No difference between the sample groups was found in overall $CD4^+$ T cells and Th2 cells. Further analysis showed a decrease in gene expression of the Th1-associated markers IFN-γ and IL-12RB2 (*Figure 4.9*). Infiltration of the gut mucosa by Th1 and $CD8^+$ T cells is a common feature of CD (Fuss *et al.*, 1996b). Activated $CD4^+$ and $CD8^+$ T cells have been detected in the PB IBD patients but not in the circulation of non-IBD controls (Funderburg *et al.*, 2013), suggesting intestinal inflammation can affect circulating T cell phenotypes. In the same study, $CD8^+$ T cells expressed high levels of the mucosal marker CD38 in patient samples, which might indicate future gut-homing. In terms of Th1 and $CD8^+$ T cell interaction, IL-2 secretion from Th1 cells has been found to promote IFN-γ-producing $CD8^+$ T cells and aid them in differentiating into memory cells (Huang *et al.*, 2007). Deficient $CD8^+$ T cell responses towards commensal bacteria have been reported in the diseased gut mucosa, involving less $CD8^+$ T cell memory pool expansion (Noble *et al.*, 2020). Even though further research is needed to confirm these findings, my analysis indicates that $CD8^+$ T cells may leave the PB and home to the intestine, where they contribute to CD immunopathology in a subset of patients. Th1 cells also appear to be recruited to promote IFN-γ production and aid $CD8^+$ memory T cell differentiation.

Paediatric patients with low PB T cells also showed a decrease in CD3 and CD8-encoding genes in the blood compared to controls and NT patients. No difference in *CD4* was seen between the patient subsets, thus, supporting the hypothesis that only specific Th phenotypes were reduced in their PB. Interestingly, when looking at Th1 markers, both LT and NT patient groups had reduced gene expression compared to non-IBD controls. Paediatric and adult-onset CD harbour many differences. Disease in children has been found to be more extensive than in adult CD, often affecting both the ileum and colon (Freeman, 2004). For example, the surgical rate is significantly higher in paediatric patients compared to adults (Kurowski *et al.*, 2021), as is the percentage of patients with EIMs (Jang, Kang and Choe, 2019). Paediatric patients with IBD also have higher levels of inflammatory markers prior to the administration of biologics, including plasma CRP levels and FC (Bouazzi *et al.*, 2022). Weiser *et al.* (2016) have compared the intestinal gene expression between adult and paediatric CD and showed that although expression patterns were partially comparable, overall, the two sample groups were distinguished based on age. These findings combined suggested that paediatric patients exhibit a different or more severe disease phenotype than adults with CD, which potentially explains why stratification of the PB T cell profile differed between children and adults.

Stratification of paediatric patients based on their CD-specific intestinal gene expression profile (i.e., genes differentially expressed between disease and non-diseased tissue) highlighted two distinct patient subsets, CD1 and CD2. Interestingly, the CD1 subset included all the PB T cell low patients. When looking at differences in T cell activation and leukocyte differentiation genes enriched in CD1 compared to CD2 patients, most of them appeared to relate to $CD8^+$ T cell activation/memory and $CD8^+$ T cell exhaustion. *CD8B* (encoding the CD8 co-receptor beta chain) and the transcription factor *EOMES* were upregulated compared to the CD2 subset. The CD8B protein is part of the CD8 receptor complex expressed on cytotoxic T cells. EOMES is a marker for memory $CD8^+$ T cells (Banerjee *et al.*, 2010) and has also been found to be required for their activation and differentiation in response to viral infection (Pearce *et al.*, 2003). Additionally, it has been shown to be associated with $CD8^+$ T cell exhaustion during anti-tumour responses (Li *et al.*, 2018). The lymphocyte activating gene 3 (LAG3/CD223) functions as T cell activation marker (Huard *et al.*, 1994). However,

it has also been related to $CD8^+$ T cell exhaustion during chronic viral infection (Blackburn *et al.*, 2008). The gene encoding for the basic leucine zipper activating transcription factor (ATF)-like transcription factor (BATF) was also upregulated in the CD1 subset. Loss of BATF has been shown to impair $CD8^+$ T cell differentiation and effector functions (Kurachi *et al.*, 2014; Topchyan *et al.*, 2021). CD7 is expressed on PB T cells and can function as marker to subset $CD8^+$ T cells. An increase in CD7 expression is an indicator for memory and naïve $CD8^+$ T cell infiltration (Aandahl *et al.*, 2003). Field Noble et al. (2020b) tested the functionality of $CD8^+$ tissue resident memory T cells (Trms) generated from PBMCs and found a dampened anti-bacterial response to commensals when compared to non-IBD controls. This suggests that naïve and memory T cells leave the PB to aid the dysfunctional Trms in the intestines in LT patients. Epstein-Barr virus-inducible protein 3 (EBI3) is part of the IL-27 cytokine protein complex, which has been found to induce self-renewal in $CD8^+$ T cells during chronic viral infection (Huang *et al.*, 2019) as well as promote their effector functions *in vitro* (Morishima *et al.*, 2005) and memory expansion in response to certain vaccines (Liu *et al.*, 2013, 2014). The KEGG pathway analysis showed associations between the CD1 subset and anti-viral and bacterial immune responses, supporting previous findings of differences in Th1/Th17-associated immunity and NK cell cytotoxicity. Since there were no changes in PB NK cells between NT and LT patients in the flow cytometry data, the latter might relate to a general increase of cytotoxic activity in the intestines. Overall, the findings in this support the hypothesis that $CD8^+$ T cells (potentially together with Th1 and Th17 cells) of LT patients are migrating from PB to the intestine to contribute to intestinal inflammation.

In addition to the CD8-associated genes, immunoregulatory *FOXP3* and *NLCR3* were also increased in CD1 patients. *FOXP3* expression, a marker for Treg cells, could potentially be linked to activated non-suppressive T cell subsets. Activated Th1, as well as Th17 phenotypes, have been found to express *FOXP3; however*, Th17 cells had more stable and higher levels (McMurchy *et al.*, 2013). However, it appears more likely that the increase could also be due to a general higher level of T cell activity, with immunosuppressive Tregs cells attempting to counteract the strong inflammatory response. In terms of migration potential, *ICAM1* was upregulated, which has been suggested as a promising target in IBD therapy to reduce leukocyte migration to the gut (Reinisch *et al.*, 2018). However, the genes

encoding its ligands *ITGAL and ITGAM* did not show decreased expression in the PB of LT patients when compared to controls and NT patients.

The link between clinical parameters such as disease phenotype, treatment and PB inflammatory markers and the LT phenotype was studied in the flow cytometry and adult transcriptomic datasets. Total lymphocyte count was decreased in the PB low T cell subset of the flow cytometry cohort (*Figure 4.19*). The percentage lymphocyte of all PB WBCs was also measured in GSE94648 patients but did not show any difference between the CD subsets. The flow cytometry data showed a reduction in all circulating $CD45^+$ cells i.e., all WBCs, in the low PB T cell patient subset due to the loss of PB T cells. Assuming that this is also the case for GSE948648, lymphocyte, as well as WBC count, would be reduced in the low PB T cell patient subset, cancelling out any changes when looking at lymphocyte frequencies. Additionally, a study in colorectal cancer has found that lymphocyte count and percentage had distinct predictive characteristics for disease prognosis, suggesting that the markers were not representative of one another (Zhao *et al.*, 2017).

CRP and CDEIS were significantly increased in LT patients of GSE94648, suggesting that patients with low PB T cells had more endoscopic activity and intestinal damage. The latter might result from the damage induced by cytotoxic T cells. No significant differences were found in CRP and HBI between LT and NT patients of the flow cytometry dataset. A reason for this might be the smaller sample size, particularly in the case of CRP, and the lower sensitivity in measuring intestinal inflammation of HBI compared to CDEIS (Schoepfer *et al.*, 2010b; Vermeire *et al.*, 2010). Overall, these results indicate that patients in the LT subset have more inflammation in their intestines and potentially increased CRP levels in their blood than the rest of the CD patient cohort. The data in this chapter suggests that they exhibit a strong intestinal cytotoxic immune response. To my knowledge, a cytotoxic T and NK cell response in the intestines have not been correlated with endoscopic activity in CD so far. However, cytotoxicity has been shown to induce damage in a selection of epithelial cells at mucosal sites, including intestinal epithelial cells (Gerloff *et al.*, 2013; Van De Sandt *et al.*, 2017; Li *et al.*, 2022). Based on these findings, I formulated the working hypothesis that patients with

the low PB T cell phenotype might potentially have an increase in endoscopic activity due to enhanced intestinal epithelial damage.

I hypothesised that the leukocytes leaving the blood of CD patients in the LT subset are migrating to the intestines. To further test this hypothesis, the gene expression level of gut-homing markers was compared in the PB of paediatric and adult patients. The data suggested a decrease in leukocytes expressing α4ß7 integrin and CCR6 in LT patients compared to the rest of the CD cohort and controls. The anti-α4ß7 therapeutic vedolizumab is now commonly used in CD maintenance and has been shown to induce long-term remission (Sandborn *et al.*, 2013a). CCR6 and its ligand CCL20 are increased in colonic tissue derived from CD patients with active disease (Skovdahl *et al.*, 2015). Moreover, CCR6 is part of the IBD genetic risk factor gene group (Graham and Xavier, 2013). Drugs targeting the CCR6-CCL20 axis have been developed but not introduced into routine clinical due to adverse side effects (Ranasinghe and Eri, 2018). Based on these findings, patients with the LT phenotype might be prime candidates for the T cell migration attenuating therapeutic vedolizumab.

4.8 Conclusion

This chapter identified a subset of patients with reduced levels of circulating T cells in three distinct CD cohorts and in both adult and paediatric disease. Adult patients showed that lymphocytes associated with type 1 immunity, Th1 $CD4^+$ and $CD8^+$ T cells, were leaving the blood. The differences between CD subsets were less clear in paediatric disease. The low PB T cell phenotype was representative of a patient group with increased cytotoxic CD8+ T cell activity in the intestine of paediatric patients. When looking at clinical data, the patients with low PB T cells appeared to have more endoscopic activity, i.e., they exhibited more damage in their intestinal lumen. The α4ß7 integrin and CCR6 were reduced in the blood of LT patients, suggesting that cells expressing these gut-homing markers are leaving the blood.

Chapter 5: Identification of an intestinal CD immunopathotype characterised by high expression in CHI3L1 and IL-1β, granuloma formation and anti-TNF non-response

5.1 Introduction

Up to 46% of CD patients still require surgical interventions even though a wide selection of front-line therapeutics is available (Frolkis *et al.*, 2013). Anti-TNF blocking treatments revolutionised clinical care for many inflammatory diseases, including CD, due to their efficacy, fast-acting nature and reduced side effects compared to general immunosuppressants. However, only approximately 20% of patients maintain long-term remission when administered with these highly specific drugs (Hanauer *et al.*, 2002; Gisbert and Panés, 2009). Additionally, previous anti-TNF exposure has been linked to an increased risk of disease refraction in response to follow-up treatments such as vedolizumab (Sands *et al.*, 2017; Verstockt *et al.*, 2020). Changes in immune system-associated gene expression have been linked to distinct treatment responses in CD. Martin *et al.* (2019) identified a complex signature in the ileum of patients, which might be linked to anti-TNF non-response. It included a network of stromal cells, activated lymphocytes, and pro-inflammatory MNPs. Friedrich *et al.* (2021) have found an increase in neutrophil and MNP activation and stroma-driven infiltration in patients with non-response to anti-TNF compared to the responders. Despite this progress in understanding pathways driving anti-TNF non-response over recent years, there is no biomarker for treatment outcome prediction in clinical use to date. Therefore, further research into the mechanisms driving the failure of anti-TNF therapy is required. Based on the evidence in the literature, I hypothesised that differences in mucosal immune responses of CD patients drive distinct immunopathotypes which underlie the range of treatment responses.

Aims:

In this chapter, I aimed to study differences in colonic gene expression between CD and normal colonic tissue using bulk RNAseq data and identify changes in

distinct immune pathways within the patient population using clustering and gene overrepresentation analyses. The findings were validated in a publicly deposited microarray gene expression dataset derived from colonic RNA with a larger sample size and access to the anti-TNF response status of patients.

5.2 Bulk RNA sequencing: Differences in intestinal gene expression of CD patients and non-IBD controls

RNA was extracted from a small cross-sectional cohort of CD patients (n=4) and non-IBD controls (n=6) and analysed by bulk RNAseq. The reads from the sequencing analysis were aligned and normalised by John Cole as previously published (Cole *et al.*, 2021).

5.2.1 Quality controls of bulk RNA sequencing data

Prior to any analyses, the quality of the RNA sequencing was tested in various ways. *Figure 5.1* shows density plots of the normalised counts of non-IBD controls and CD patients. This quality control step is important to determine if all samples have a similar distribution of genes with similar counts and to judge their comparability. All samples showed the highest density of genes at around 1000 counts. Additionally, all samples apart from NC_2 and NC_4 had a second and third peak with gene counts at approximately 7 and just over 0. Discussion around whether these samples should be excluded resulted in the conclusion that most genes with low counts are classified as background noise, and the region of interest lies within the second peak. All samples have most of their reads in a similar count region, suggesting that the data are comparable.

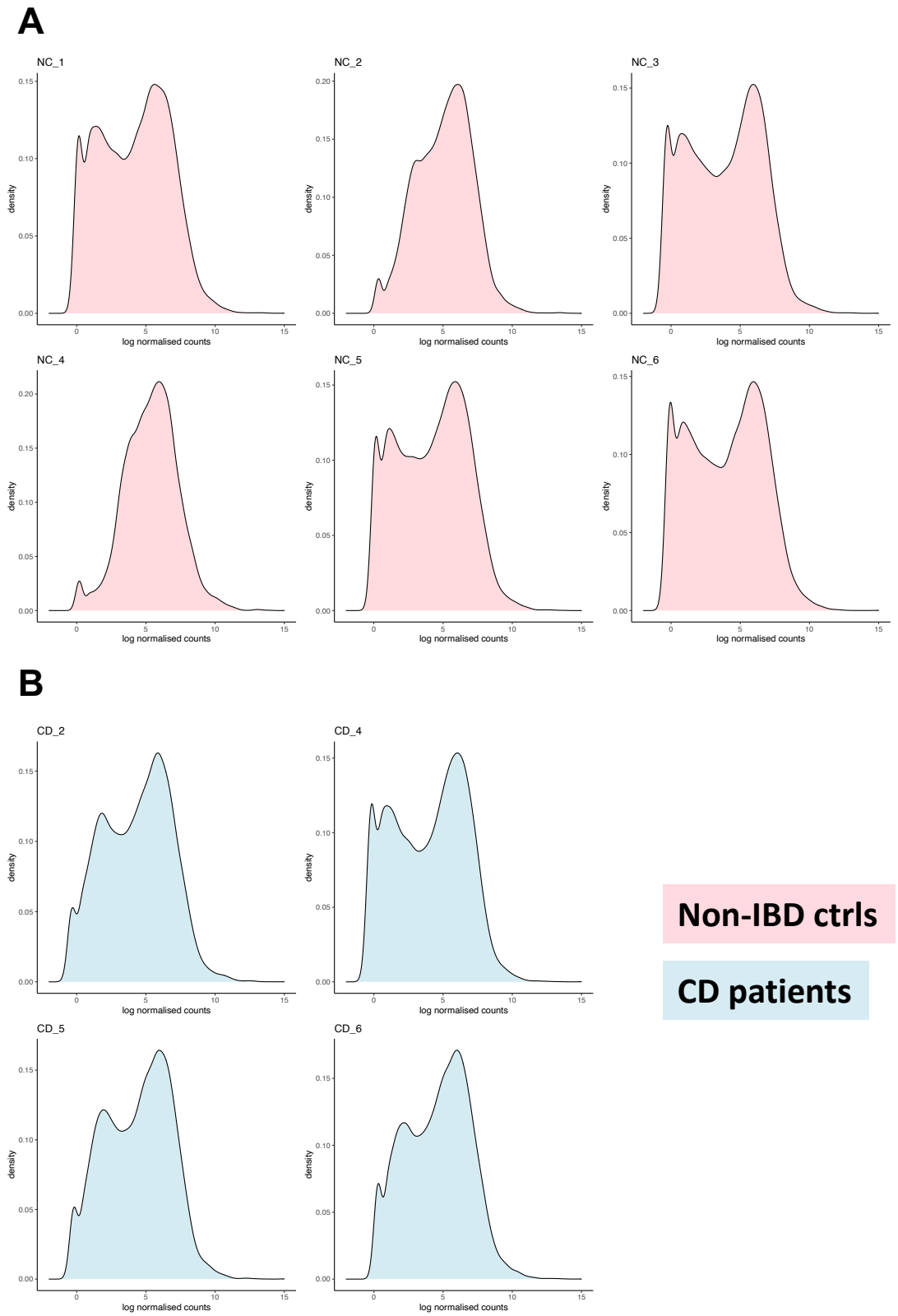


Figure 5.1: ***Gene count distribution in bulk RNAseq samples.***

RNA derived from colonic tissue of A) 6 non-IBD controls and B) 4 CD patients was analysed using bulk RNA sequencing. The density of counts is displayed as fraction of 1 and were visualised using the *geom_density()* function of the ggplot2 package. Normalised counts are shown in log transformation. All analysis was performed on R studio (R version 4.0.2 and below).

The sample distance matrix (*Figure 5.2A*) and PCA (*Figure 5.2B*) give an indication of the similarity between samples based on the sample-to-sample distance based on all gene expressions and principal components. The latter shows the amount of variation explained within the dataset using different principal components. The principal component explaining the most variation (40.2%), the PC1, separated CD and NC samples. The patient sample CD_5 appeared to group with non-IBD controls, supported by the sample distance matrix where this patient clustered among the control group. PC2, explaining 17.2% of the variation, showed differences in expression within the sample groups. NC_4 and NC_6 appeared to have the largest difference in gene expression based on PC2 and the sample distance matrix.

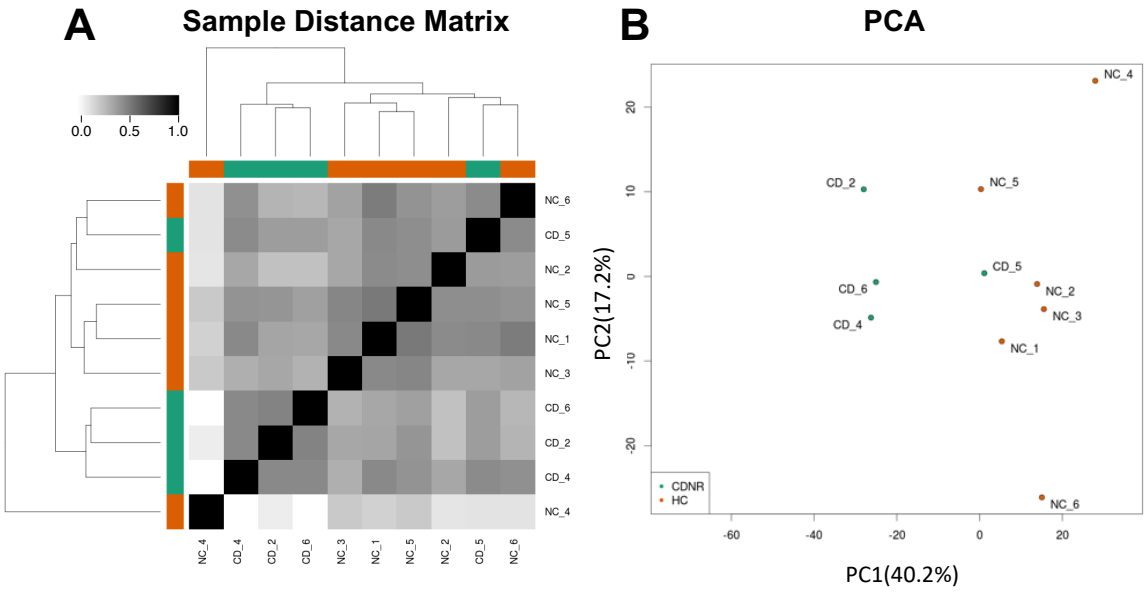


Figure 5.2: ***Sample distance matrix and Principal Component Analysis (PCA).***

RNA derived from colonic tissue of 6 non-IBD controls and 4 CD patients was analysed using bulk RNA sequencing. John Cole performed all analyses on R studio (R version 4.0.2 and below) using the Searchlight pipeline (Cole *et al.*, 2021). A) Sample distance matrix displaying similarity between samples based on whole gene expression. CD patients are annotated in green, NC samples are annotated in orange. B) PCA showing variation between CD patients (CDNR, blue) and controls (NC, pink) based on principal component (PC)1 and PC2. PC1 explains 40.2% and PC2 17.2%. of variation between the sample groups. All analysis was performed on R studio (R version 4.0.2 and below).

5.2.2 Differential gene expression and gene set enrichment analyses

Next, a volcano plot was generated to highlight the differences between non-IBD controls and CD patients, showing significant genes (adj. p<0.05) with a fold change <-1 or >1 in pink and genes with adjusted p-value >0.05 but fold change <-1 or >1 in blue (*Figure 5.3*). CD and non-IBD controls had 979 DEGs, 478 were upregulated, and 501 were downregulated in the patient group. The top 10 DEGs were summarised in *Figure 5.4A*. They included osteopontin (*SPP1*), a protein previously shown to be elevated in the mucosa and serum of CD patients (Sato *et al.*, 2005), and genes encoding the immunoglobulin components, immunoglobulin heavy constant gamma 3 (*IGHG3*) and 1 (*IGHG1*), which are indicators of an ongoing antibody response. Unsupervised hierarchical clustering of DEG expression was displayed as a heatmap in *Figure 4B* and showed a clear separation between the two sample groups.

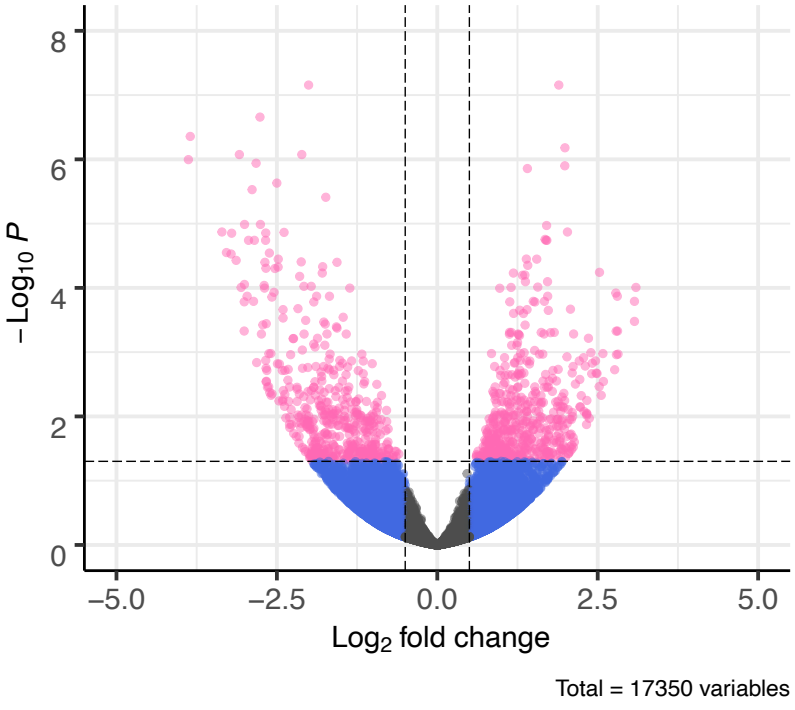


Figure 5.3: ***Volcano plot comparing differential gene expression between CD patients and non-IBD controls.***

The gene expression in colonic tissue derived from 4 CD patients and 6 non-IBD controls was compared. Differential gene expression was performed using the *DESeq* package. Fold change was displayed in $\log_2$ and p-value in $-\log_{10}$ transformation. Black dots: p-value>0.05, $-0.5<\log_2<0.5$, Blue dots: p-value>0.05, $-0.5<\log_2>0.5$, Pink dots: p-value$\leq$0.05,$-0.5>\log_2<0.5$. All analysis was performed on R studio (R version 4.0.2 and below). All analysis was performed on R studio (R version 4.0.2 and below).

A

Gene	adj.P.Val	logFC
IGHG3	7.57E-12	-4.5286624
IGHG1	3.19E-10	-4.4452565
SPP1	5.34E-10	-2.6049774
THBD	6.98E-08	-2.0058238
OAF	6.98E-08	1.89398649
LILRB4	2.20E-07	-2.763065
FCGR1A	4.40E-07	-3.8499871
TRIB3	6.61E-07	1.98947478
PLA2G2D	8.42E-07	-3.0845555
MS4A7	8.42E-07	-2.1115414

B

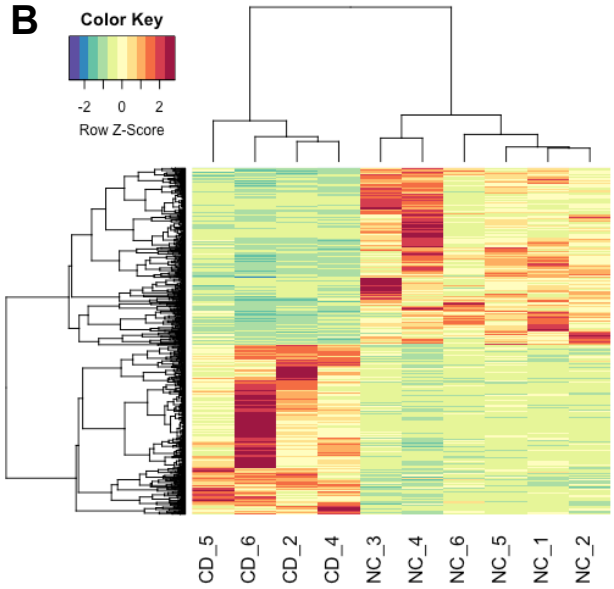


Figure 5.4: ***Differential gene expression between diseased and normal colonic tissue.***

The gene expression of colonic tissue derived from 4 CD patients and 6 non-IBD controls was compared. DEGs were selected based on p <0.05. Differential gene expression was performed using *DESeq* package. A) Top 10 DEGs between CD and controls. B) Heatmap visualising all DEGs with hierarchical clustering of samples and genes. All analysis was performed on R studio (R version 4.0.2 and below). *IGHG3:* immunoglobulin heavy gamma chain 3, *IGHG1:* immunoglobulin heavy gamma chain 1, *SPP1:* osteopontin, *THBD:* thrombomodulin, *OAF:* out at first homolog, *LILRB4:* leukocyte immunoglobulin like receptor B4, *FCGR1A:* FC gamma receptor 1A, *TRIB3:* tribbles pseudokinase 3, *PLA2G2D:* phospholipase A2 group 2D, *MS4A7:* Membrane Spanning 4-Domains A7.

A biological theme comparison analysis was performed next to test for enriched gene sets in the upregulated and downregulated genes (*Figure 5.5*). Amongst the upregulated pathways were the inflammatory conditions *Rheumatoid arthritis*, *Systemic lupus erythematosus* and asthma, as well as IBD, confirming that the patients sampled showed changes related to chronic inflammation. Additionally, a wide range of inflammation-related pathways and processes, such as TLR and cytokine signaling, were up-regulated, which are involved in the CD pathogenesis (Hausmann *et al.*, 2002; Guan and Zhang, 2017). Additionally, the gene set *Intestinal immune network for IgA production* was enriched among the upregulated genes, an immunological process crucial for host defence against the intestinal microbiota and pathogens and known to be aberrant in CD (Gutzeit, Magri and Cerutti, 2014). Overall, these analyses suggest that CD patients have intrinsic differences from non-IBD controls based on their state of inflammation.

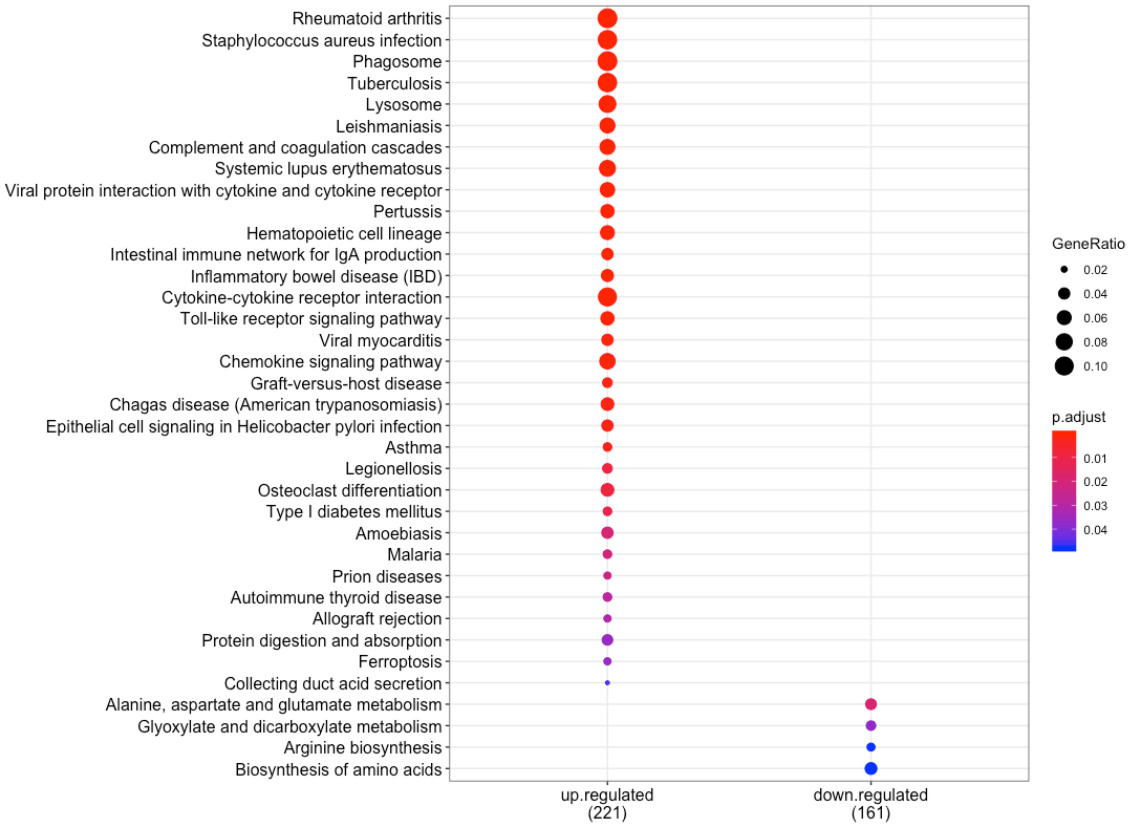


Figure 5.5: ***Biological theme comparison.***

Gene set enrichment analysis was performed based on differential gene expression analysis between colonic tissue of CD patients and non-IBD controls. The dot size refers to the GeneRatio, i.e., the proportion of upregulated genes in each biological theme. The colour is related to the adjusted p-value (p.adjust). Analysis was performed in R Studio (R version 4.0.2 and below) using the *clusterProfiler* package and *compareCluster* function with the KEGG database as reference.

5.3 Stratification of CD patients based on immune system-associated genes

To determine differences in the immune response amongst CD patients, the DEGs linked to the GO *immune system process* were selected. As previously described, this GO comprises all genes related to the development and function of the immune system (EMBL-EBI, 2018). By using this method, all differences in gene expression unrelated to changes in the immune system of patients and controls can be excluded. 224 of the DEGs were included under the GO *immune system process*. Subsequently, I focussed on upregulated genes to determine activated immunological pathways that could serve as potential treatment targets in the future (n =187).

5.3.1 RNA sequencing: Immune system process genes

The heatmap in *Figure 5.6* shows the gene expression of upregulated *immune system process* genes. Hierarchical clustering clustered non-IBD controls and CD patient samples separately and highlighted four gene clusters. One cluster was seemingly enriched in patient CD_6 and not in any other patient, suggesting that it may relate to an inflammatory response only found in a patient subset. The cluster is comprised of 116 genes which are summarised in *Table 5.1*.

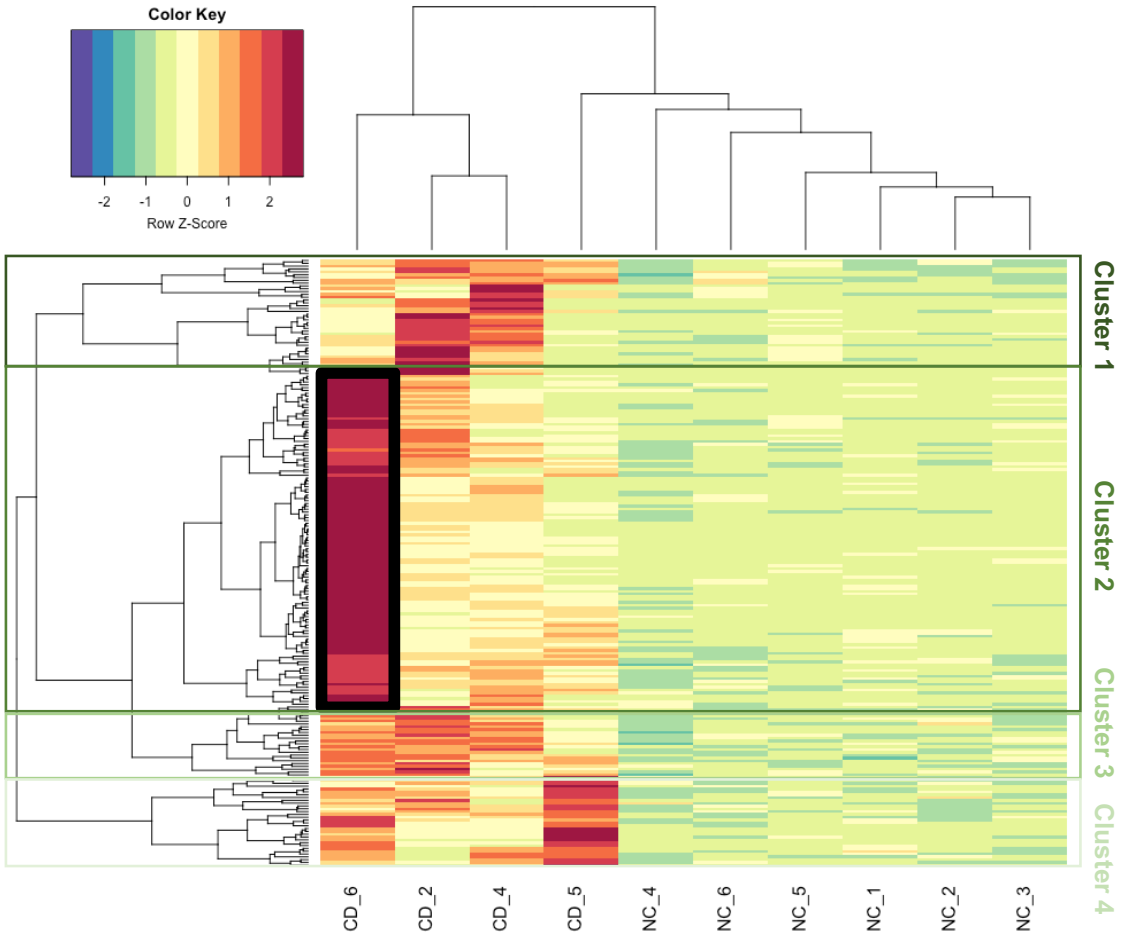


Figure 5.6: ***Immune system process genes upregulated in CD.***

Upregulated DEGs in intestinal tissue of CD patients (n =4) compared to controls (n =6) that were associated with the GO "Immune System Process" were selected. Heatmap visualising output with hierarchical clustering of samples and expression. All analysis was performed on R studio (R version 4.0.2 and below).

Immune cluster 2 signature genes					
ADA	*CD300C*	*DCSTAMP*	*HLA-DPA1*	*LYZ*	*SIGLEC5*
AIF1	*CD44*	*DUSP10*	*HLA-DPB1*	*MMP12*	*SIGLEC7*
AMPD3	*CD48*	*ENPP2*	*HLA-DQA1*	*MMP9*	*SLAMF7*
AP1S2	*CHI3L1*	*FCAR*	*HLA-DRA*	*NBN*	*SLC11A1*
APOE	*CLEC2B*	*FCER1G*	*HMOX1*	*NLRP3*	*SLC2A5*
ARL11	*CLEC4A*	*FCGR1A*	*IGSF6*	*ORM1*	*SLC7A7*
ATP6AP1	*CLEC7A*	*FCGR2A*	*IL1B*	*P2RX7*	*SNX10*
C1QA	*CMKLR1*	*FGR*	*IL1RN*	*PDGFRB*	*SPI1*
C1QB	*COL1A1*	*FLT1*	*IL6*	*PIK3AP1*	*TLR2*
C1QC	*COL1A2*	*FPR1*	*INHBA*	*PILRA*	*TLR8*
C1S	*COL3A1*	*FPR2*	*IRAK3*	*PLA2G2D*	*TNFAIP6*
C2	*CREG1*	*FTL*	*ITGAX*	*PLA2G7*	*TNFSF13B*
C3	*CRISPLD2*	*G6PD*	*ITGB2*	*PLEK*	*TNFSF15*
C3AR1	*CTSB*	*GM2A*	*KYNU*	*PLEKHO2*	*TREM1*
C5AR1	*CTSC*	*GNS*	*LAMP1*	*PSAP*	*TYROBP*
CCL20	*CTSH*	*GPNMB*	*LAMP2*	*PTPRJ*	*VIM*
CCL3	*CTSK*	*GPR84*	*LGALS1*	*S100B*	
CCL3L1	*CXCL8*	*HAVCR2*	*LILRB4*	*SEMA4A*	
CCR1	*CXCR1*	*HCAR2*	*LST1*	*SERPINA1*	
CD1B	*CYP27B1*	*HIST1H2BF*	*LYN*	*SIGLEC14*	

Table 5.1: ***Gene list of the immune cluster enriched in patient CD_6.***

5.3.2 Microarray: Immune system process genes

To understand whether this cluster could also be identified in a different set of patients, the same analysis was applied to a microarray dataset of RNA derived from colonic tissue of 19 CD patients and 6 non-IBD controls (*Table 5.3*). This publicly available data was deposited on NCBI under the accession number *GSE16879* by Arijs *et al*. (2010). 65 of the 116 immune cluster genes were included in the microarray probe, displayed in the heatmap of *Figure 5.7*. Interestingly, hierarchical clustering showed that the immune cluster signature not only separated CD samples from non-IBD controls but also clustered patients into two distinct groups, classified as immune cluster up (CD Up, n =*7*) and normal (CD Normal, n =*12*). Initial visual examination indicated that the patients CD_16 and CD_19 mimic the cluster 2 gene expression patterns of CD Up patients even though

they clustered separately. I, therefore, looked at the mean expression of all signature genes. The means of CD Normal patients ranged between 58.7 and 96.6, and those of CD Up patients were between 277.9 and 443. CD_16 and CD_19 had a mean signature gene expression of 237.8 and 241.7, indicating that their means were more comparable to the CD Up than CD Normal samples and were thus re-classified accordingly. Immune clusters 1, 3 and 4 identified in *Figure 5.6* were also studied in the validation dataset but did not distinguish between patients and were therefore not analysed further.

	CD (n= 19)	**NC (n= 6)**
Median Age (Range)	31.8 (23.7-46.5)	*N/A*
M:F ratio	1.375	*N/A*
Median CRP levels in mg/dl (Range)	10.2 (4.3-35)	*N/A*
Treatment **Monotherapies (%)** *Steroids* *Thiopurines* *5-ASA* **Combination Therapies** *Immunosuppressants + Steroids*	 4 (21.2%) 14 (73.7%) 8 (42.1%) 2 (10.5%)	*N/A*
Location (%) L1: *Ileal* L2: *Colonic* L3: *Ileocolonic*	 0 14 (73.7%) 5 (26.3%)	

Table 5.2: ***Summary of cohort information of GSE16879 microarray dataset.***

19 CD patient samples and 6 non-IBD volunteers (NC) were recruited by Arijs *et al*. (2010). CRP levels in mg/dl, current treatments and disease location were included for CD patients. *N/A:* Not Applicable. 5-ASA: aminosalicylates. CRP: C-reactive protein.

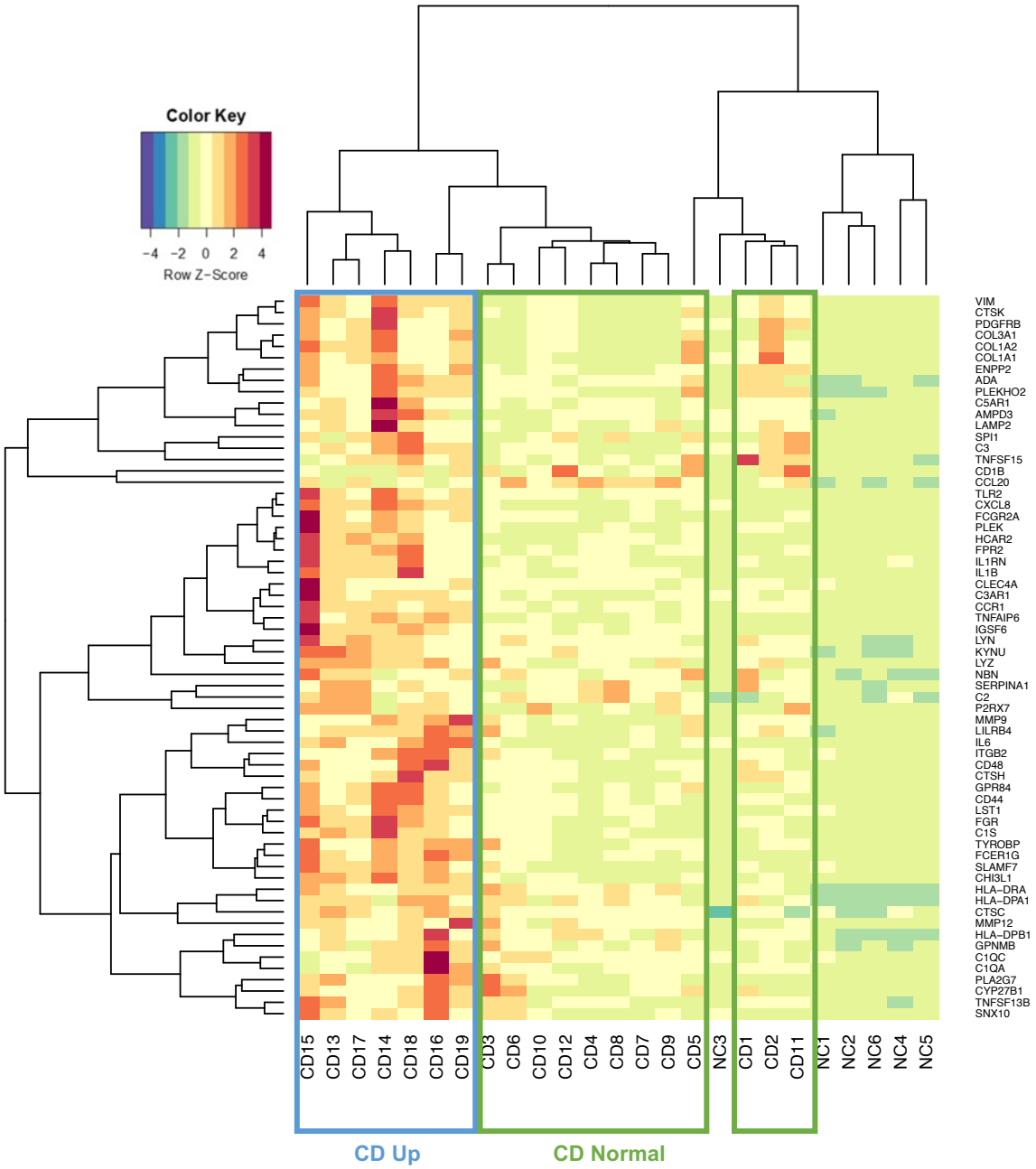


Figure 5.7: ***Gene expression of immune cluster genes in GSE16879.***

Upregulated DEGs in intestinal tissue of CD patients (n =19) compared to controls (n =6) associated with the GO *immune system process* were selected. Heatmap visualises the output with hierarchical clustering of samples and expression. CD patients were categorised based on the expression level of immune cluster genes into CD Normal (*n* =7) and upregulated levels (CD Up, *n* =12). Data used are deposited on NCBI under GSE94648 by Arijs *et al*. (2010). The visualisation was performed in R studio (R version 4.0.2 and below). Differential gene expression analysis was performed using the GEOR2 tool from the NCBI.

5.4 Stratification of CD patients based on their immune cell composition and immune pathways

After showing that the immune gene cluster 2 could be found not only in the bulk RNAseq data but also in the GSE16879 microarray dataset, other differences in immune cell composition and activated immune pathways between CD Up and CD Normal patients were studied.

5.4.1 Gene set enrichment and pathway analysis

First, GSE and pathway analyses were performed. These are valuable tools to understand which immunological processes and pathways are enriched in specific sample groups. The following analyses were based on all genes included in *Table 5.1* and fold changes between controls and CD patients in the bulk RNAseq data. MeSH and WikiPathway databases were used as references for the analyses.

Figure 5.8A displays the MeSH linked with the immune cluster. Subject headings enriched were associated with neutrophils (*neutrophil infiltration and activation*), granulation response (*respiratory burst and cell degranulation*) and other immune processes, such as *complement activation*, osteoclast activity (*bone resorption*) and activation of *adaptive immunity*. *Figure 5.8B* showed the enriched pathways, which had an association with toll-like receptor activity (*TLR4 signaling and tolerance, toll-like receptor signaling pathway* and *regulation of toll-like receptor signaling pathway*), complement (*complement and coagulation cascades, human complement system* and *complement activation*), anti-viral immune responses (*Hepatitis C and hepatocellular carcinoma, Ebola virus pathway on host*), pro-inflammatory cytokines (*photodynamic therapy-induced NF-kB survival signaling, IL1 and megakaryocytes in obesity, cytokines and inflammatory response*), and macrophage repair and pro-inflammatory responses (*spinal cord injury*, *lung fibrosis* and *microglia pathogen phagocytosis pathway*). These findings indicated that neutrophils and macrophages were potentially driving the inflammation in the immune cluster, potentially *via* complement and TLR activation and IL1 and NF-kB signaling.


A
Respiratory Burst
Neutrophil Infiltration
Neutrophil Activation
Graft Survival
Genes, MHC Class II
Dose–Response Relationship, Immunologic
Complement Activation
Cell Degranulation
Bone Resorption
Adaptive Immunity
ADA AIF1 APOE C1QA C1QB C1QC C1S C2 C3 C3AR1 C5AR1 CCL20 CCL3 CCR1 CD44 CD48 CH25L1 CLEC2B CLEC7A COL1A1 CTSB CTSC CTSK CXCL8 CXCR1 DCSTAMP ENPP2 FCAR FCER1G FCGR1A FCGR2A FCGR2C FGR FPR1 FPR2 G6PD HAVCR2 HLA-DPA1 HLA-DPB1 HLA-DQA1 HLA-DRA HMOX1 IL1B IL1RN IL6 INHBA IRAK3 ITGAX ITGB2 LAMP1 LAMP2 LGALS1 LYN LYZ MMP9 NLRP3 P2RX7 PILRA PLEK S100B SERPINA1 SIGLEC14 SIGLEC5 SIGLEC7 SLC11A1 SNX10 SPI1 TLR2 TLR8 TNFAIP6 TNFSF13B TNFSF15 TREM1 TYROBP VIM
B
TYROBP Causal Network
Toll–like Receptor Signaling Pathway
TLR4 Signaling and Tolerance
Spinal Cord Injury
Senescence and Autophagy in Cancer
Regulation of toll–like receptor signaling pathway
Photodynamic therapy–induced NF–kB survival signaling
Peptide GPCRs
Oxidative Damage
Overview of nanoparticle effects
Microglia Pathogen Phagocytosis Pathway
Lung fibrosis
Inflammatory Response Pathway
IL1 and megakaryocytes in obesity
Human Complement System
Hepatitis C and Hepatocellular Carcinoma
Ebola Virus Pathway on Host
Cytokines and Inflammatory Response
Complement and Coagulation Cascades
Complement Activation
Allograft Rejection
AIF1 C1QA C1QB C1QC C1S C2 C3 C3AR1 C5AR1 CCL3 CCR1 CD44 COL1A1 COL1A2 COL3A1 CREG1 CTSB CXCL8 CXCR1 FCER1G FCGR1A FCGR2A FCGR2C FPR1 FPR2 HAVCR2 HLA-DPA1 HLA-DPB1 HLA-DQA1 HLA-DRA HMOX1 IGSF6 IL1B IL6 INHBA IRAK3 ITGAX ITGB2 LYN MMP12 MMP9 NLRP3 PLA2G7 PLEK SERPINA1 SIGLEC7 SLC7A7 TLR2 TLR8 TNFSF13B TNFSF15 TREM1 TYROBP VIM

180

Figure 5.8: ***Gene set enrichment and pathway analysis of immune cluster 2 genes.***

112 signature genes were tested for significant associations with medical subheadings (MeSH) and pathways. A) MeSH analysis was performed with the *enrichMESH()* function from the meshes package accessing the pub2pubmed database curated by the NCBI. B) WikiPathway analysis was performed with the *enricher()* function using the wpid2gene database (Version 20180810) as reference. Data used are deposited on NCBI under GSE94648 by Arijs *et al.* (2010). All analysis was performed in R studio (R version 4.0.2 and below) and required the clusterProfiler package.

5.4.2 Digital cytometry: CIBERSORTx deconvolution tool

After identifying neutrophils and macrophages as potential cell drivers of the immune response related to the cluster 2 gene signature, CIBERSORTx was applied to estimate immune cell composition, as previously explained in *section 4.5.3*. The signature matrix, called LM22, used as a reference, included 22 leukocyte subsets derived from human PBMCs and was quantified with the microarray analysis (Newman *et al.*, 2019b). *Figure 5.9A* shows the fractions of the 22 leukocyte subsets derived from peripheral blood in all RNAseq dataset samples. The following 16 immune cell types were identified: monocytes, M1 macrophages, M2 macrophages, M0 macrophages, neutrophils, naïve and memory B cells, plasma cells, $CD8^+$ T cells, naïve $CD4^+$ T cells, memory and activated memory $CD4^+$ T cells, regulatory T cells (Tregs), NK cells resting, DCs activated, activated mast cells and eosinophils. Non-IBD controls and CD patients did not show consistent differences in immune cell composition. When looking closer at CD patients, it became evident that patient CD_6, the patient enriched for the immune cluster, had a high proportion of the pro-inflammatory M1 macrophages and activated mast cells (*Figure 5.9B*). No statistical analysis could be performed due to a lack of samples in CD subsets.

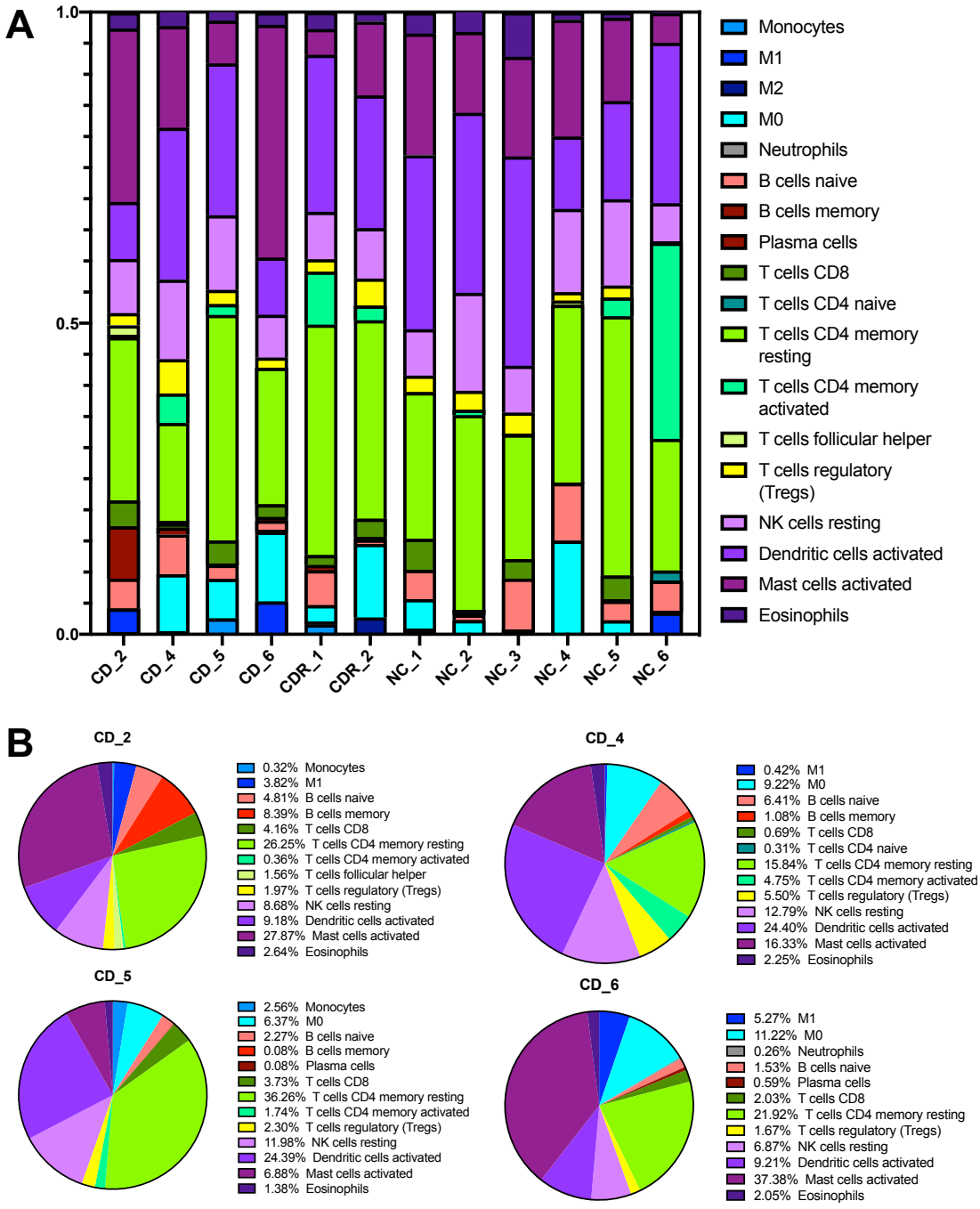


Figure 5.9: ***Immune cell composition in colonic tissue of active CD and non-IBD control.***

RNA derived from colonic tissue 6 non-IBD controls and CD patients was analysed using bulk RNA sequencing. Immune cell composition was estimated using the digital cytometry CIBERSORTx platform. A) Immune cell composition displayed for CD and control (NC) samples as fractions of 1 visualised as stacked bar plots. B) Percentage of immune cells in CD samples visualised as pie charts.

Because of the small sample size in the initial patient cohort, I repeated the analysis on the GSE16879 microarray data. However, all apart from 3 samples had a correlation coefficient below 0.5, suggesting that the CIBERSORTx analysis did not accurately estimate immune cell fractions in 88% of samples. Therefore, I decided not to include the results of the analysis.

5.4.3 Intestinal macrophage signatures

The WikiPathway analysis displayed a significant association with the pathways *spinal cord injury*, *lung fibrosis* and *microglia pathogen phagocytosis pathway* (*Figure 5.8*), potentially referring to increased macrophage activity in the immune cluster 2 signature. Since the CIBERSORTx analysis was not successful for the GSE16879 dataset, I next used two intestinal macrophage signatures to understand whether the CD subsets had differences in gene expression related to this immune cell type. Peters *et al*. (2017) identified a set of genes driving IBD pathogenesis in the intestine, titled "Macrophage Driver Genes" (*RAC2, ITGB2, SASH3, LILRB1, GMFG, SLA, NCF1, P2RX5, GPSM3, CORO1A, CTLA4, LILRA6, AIF1, RGS19, NCKAP1L, FMNL1, ITGAL, ARHGDIB, DOK3, LILRB4, SLC15A3, SNX20, TBC1D10C, NFAM1, CD4, ARHGAP30, MS4A4A, PPP1R16B, CD163, SELPLG, NCF1C, RASSF5, LCP2, APOBEC3G, PIK3CD, SLAMF1, PTPN7, SLAMF8, CD37, IPCEF1, LST1, TNFAIP8L2, SELL, ADA, HCK, RHOH, ITGAX, CYTIP, CCR5, PCED1B, CD40, ARHGAP25, SP140, PLEKHO2, FGR, CYTH4, JAK3, BCL2A1, ARHGAP9, PARVG, CD53, APBB1IP, LAMP3, FCER1G, RASGRP2, KCNAB2, GBP5, NCF2, APOL3, LEF1, ODF3B, ADAM19, ANXA6, DOCK11, GLYR1, TRBC1, MNDA, DEF6, WARS, FYB, CD28, PTPRC, PLEK, EVI2A, WAS, SP110, SAMSN1, SPOCK2, SEMA4D, RASSF2, PRKCH, HLA-DRA, IL26, DENND1C, HHEX, LCP1, IDO1, KYNU, RASGRP1, CTSK, CSF2RB, LY96, CLEC2B, YME1L1, ELK3, C3, LYN, TRIM22, MAFB, C1S, ANKRD44, FYN, TNFAIP3, TNFAIP8, ALOX5AP, FPR1, ICAM1, GBP4, TCF4, GPR65, ATOX1, C17ORF62, LGALS1, CHST2, CD3D, MSN, RFX5, LPCAT1, CHI3L1, LRRK2, SLFN11, VIM, LINC00152*) (*Figure 5.12A*). scRNAseq of CD-derived intestinal tissue was published by Martin *et al*. (2019) and elucidated a macrophage gene set, titled "scRNAseq Macrophage Genes" (*CD163L1, LILRB5, FOLR2, FUCA1, GPNMB, APOC1, LGMN, STAB1, SEPP1, CSF1R, MS4A4A, SLC40A1, SLCO2B1, MAFB, MRC1, DAB2, CD209, DNASE1L3, RGS2, JAML, VSIG4, CPVL, CLEC10A, RNASE6, CEPBD, C1QA, C1QB, C1QC, CD14, CD68, FCGR3A, PLD3, CCL4L2, CXCL2, CXCL3, CD83, CXCL8, SOD2, TXN, CCL3, CCL3L3, CCL4, NFKBIA, CD44, PKM, S100A6, IL1B, PLAUR, IER3, C15orf48, G0S2, S100A9, S100A8, CCL20, IDO1, INHBA, IL1RN, TNF, TNFAIP6, CXCL10, CXCL9, MIR155HG, NINJ1, IL6, GBP1, CAPG, AQP, MMP9, LDHA, KYNU, IL23A, S100A12, WTAP, PTGS2, TRAF1, PLEK, GK, IL1A, SERPINA1, NFKB1, PIM3, CFLAR, IL4I1, BIRC3, LILRA1, SNX10, CD40, ICAM1, TSPO, PDE4DIP, TNFAIP2, GBP5, DUSP2, CDKN1A, IRF1*) (*Figure 5.12B*)..

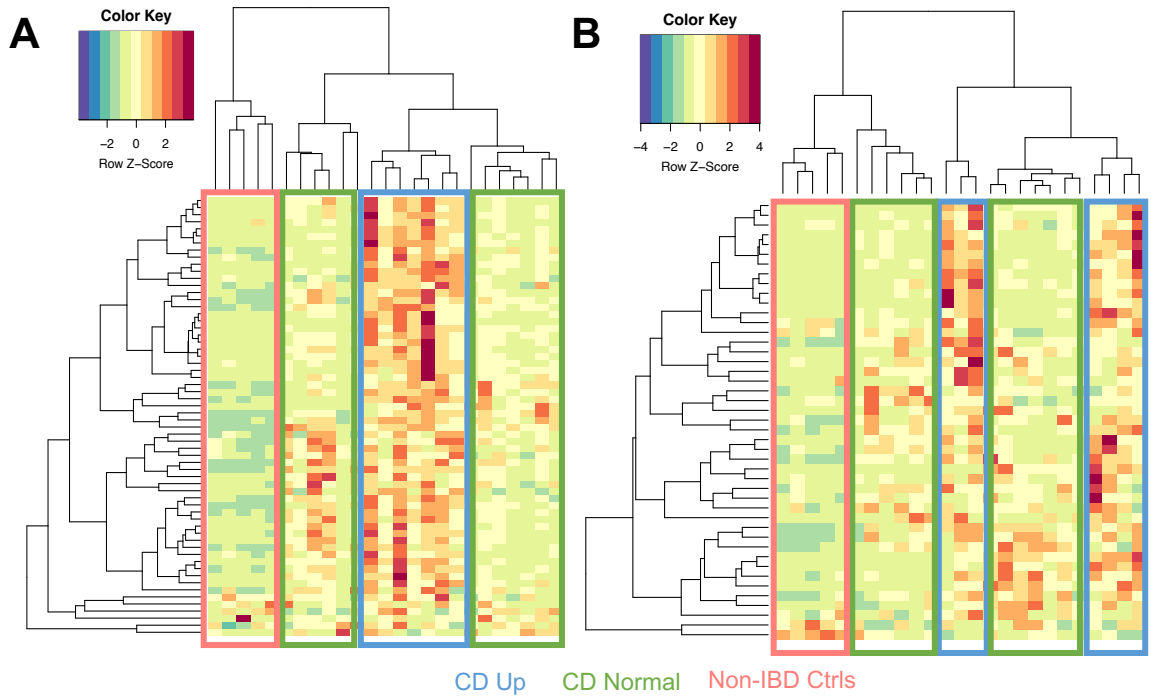


Figure 5.10: ***Macrophage gene expression in CD and non-IBD controls in the GSE16879 dataset.***

Macrophage-associated gene expression in colonic tissue was compared between CD patients (n =19) and controls (n =6). CD patients were categorised based on the expression level of immune cluster genes into CD Normal (*n* =12) and upregulated levels (CD Up, *n* =7). A) Heatmap showing macrophage driver genes as defined by Peters et al. (2017). B) Heatmap showing macrophage-associated scRNAseq cluster as defined by Martin *et al.* (2019). Data used is deposited on NCBI under GSE16879 by Arijs *et al.* (2009). All statistical analysis was performed in R studio (R version 4.0.2 and below).

To test whether the CD Up and CD Normal patients have a distinct expression of intestinal macrophage genes, the genes from the signatures were selected in the GSE16879 dataset. Hierarchical clustering of expression values distinguished the two CD subsets into the CD Up and CD Normal groups (*Figure 5.10*). For both signatures, CD Up patients upregulated most macrophage genes, whereas CD Normal patients only increased expression of a subset. To confirm these findings, multiple comparison analysis was performed, comparing all the intestinal macrophage signature genes between the CD subsets. 86 genes of the 126 Macrophage Driver Genes were up-regulated in the CD Up subgroup compared to the Normal CD patient subset (*data not shown*). Additionally, 34 of the 86 genes in the scRNAseq Macrophage Gene signature were increased in patients enriched for the immune cluster 2 (*data not shown*). To understand the relationship between the macrophage signatures and the immune cluster 2, a Venn diagram was made to visualise the overlap of genes between the different gene lists. 31.89% of immune cluster genes (37 out of 116) overlapped with the macrophage

signatures (*Figure 5.11*), indicating that not only neutrophils (shown by the WikiPathway analysis in *Figure 5.8*) and mast (shown by the CIBERSORTx analysis in *Figure 5.9*), but also intestinal macrophages were involved in driving the immune response in the CD Up patient subset.

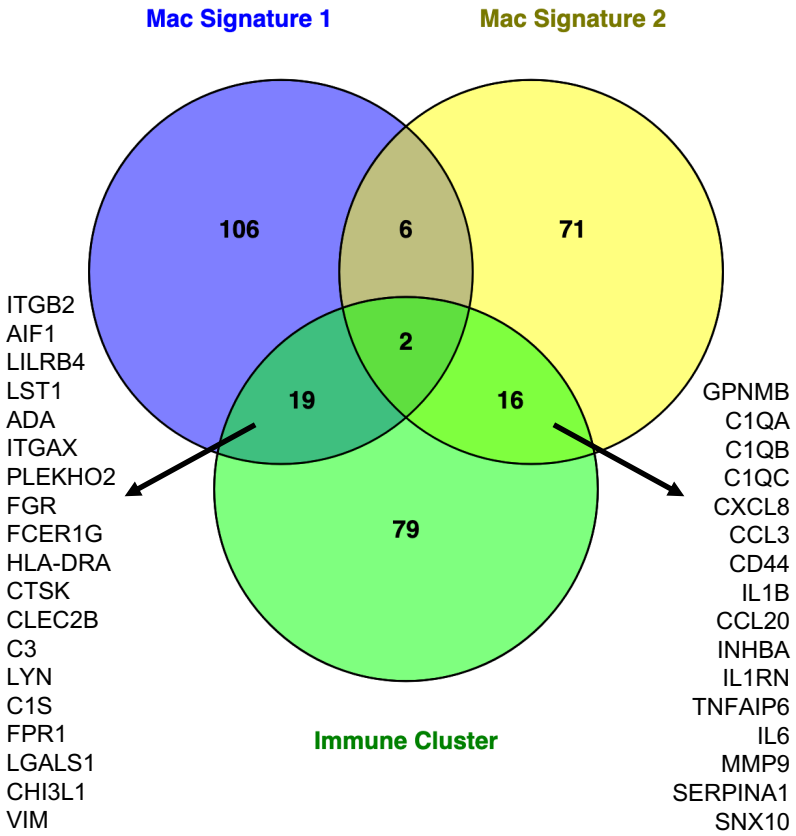


Figure 5.11: ***Venn diagram showing the overlap between macrophage signatures and the immune cluster signature.***

Mac Signature 1 refers to the "Macrophage Driver Genes" (Peters et al., 2017) and Mac Signature 2 to the "scRNAseq Macrophage Genes" (Martins *et al.*, 2019). Immune cluster refers to the immune cluster 2 gene signature identified in *section 5.3.1*. Diagram was generated using the Venny 2.1.0 tool (Oliveros, 2007).

5.5 Association between anti-TNF response and immune cluster up-regulated patients.

After investigating which immune cells and pathways were potentially involved in the immune response characterising immune cluster 2, it was evaluated whether this immunopathotype could be related to anti-TNF non-response in CD. Patients of the *GSE16879* dataset were sampled at baseline of anti-TNF treatment, and

their response to the therapy was followed up 4 to 6 weeks post-treatment initiation. The response was evaluated based on endoscopic examination and a histological inflammation score previously published by D'Haens *et al*. (1998). A histology score below 3 and complete mucosal healing assessed during endoscopy were treated as a response to anti-TNF. Based on this scoring system, 7 CD patients were classified as non-responders and 12 as responders to anti-TNF therapy. Interestingly, when overlapping the treatment response, all patients in the CD Up group were non-responders, and all CD Normal patients were responders.

To further elucidate the association between the immune cluster 2 identified in *section 5.3.2* and anti-TNF non-response, differential gene expression analysis was performed between responders and non-responders. 344 DEGs were identified in the analysis, 35 of which were part of the immune cluster. Differential gene expression is not necessarily reflective of predictive power for treatment response. Therefore, ROC analysis was performed next, which calculates the AUC reflective for the measure of separability between the two groups, i.e., the closer the AUC to 1, the better the variable is at discriminating between anti-TNF responders and non-responders. When testing the 35 immune cluster 2 signature genes, which were also DEGs between responders and non-responders, 9 genes had an AUC of 1.000 for predicting non-response to anti-TNF in the *GSE16879* datasets. These were *CHI3L1*, *TLR2*, TNF alpha induced protein 6 (*TNFAIP6*), *IL6*, *IL1B*, complement C1s (*C1S*), FGR proto-oncogene (*FGR*), FC gamma receptor IIa (*FCGR2A)*, and immunoglobulin superfamily member 6 (*IGSF6*). When correlating the gene expression values of these 9 genes, they all showed significant positive correlations ($r^2 > 0.56$ and $p < 0.05$, *Figure 5.12)*.

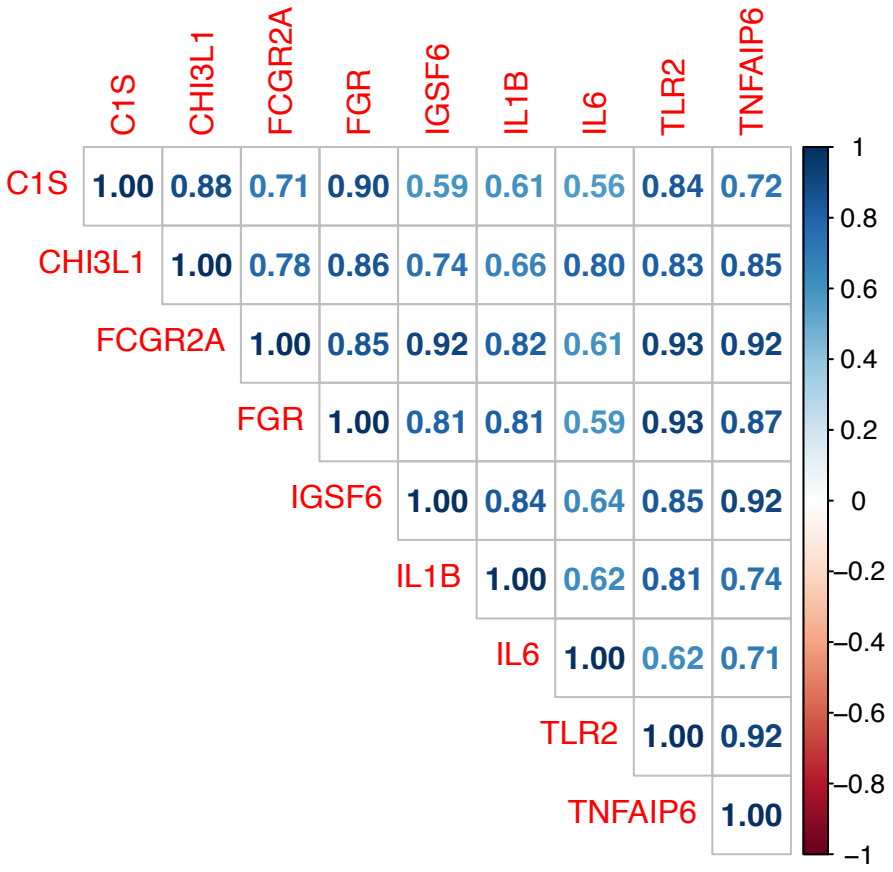


Figure 5.12: ***Correlation plot showing anti-TNF prediction signature.***

9 signature genes identified as predictors for non-response to anti-TNF are shown and their R^2-value displayed in each box. All genes had a significant correlation ($p < 0.05$). Correlation was performed using the *rcorr()* function of the Hmisc package. The correlation plot was generated using the corrplot package. Data used is deposited on NCBI under GSE94648 by Arijs *et al.* (2009). All analysis was performed in R studio (R version 4.0.2 and below).

I subsequently performed a random forest analysis using these 9 genes as variable input to test whether they have the same predictive power altogether (*Figure 5.13*). Random forest analysis is a regression method that uses machine learning to identify the predictive power of a signature (in this case the 9 target genes) for dependent variables (such as current treatment, inflammatory markers, and disease phenotype). Many decision trees are used to generate the random forest algorithm that discriminates between dependent variables. Subsequently, the ability of this algorithm to accurately predict each variable is tested. The predictive power of these variables was calculated by correlating the observed variable value, with the prediction error for each data point, which is referred to as out-of-bag error (OOB). Random forest analysis using categorical data as dependent variable, generates an area-under-curve (AUC) value output, visualised as receiver operating characteristic (ROC), as well as two measures demonstrating

the importance of each gene within the model, called mean decrease accuracy and mean decrease Gini. The higher the mean decrease accuracy, the more the model's accuracy is impacted if the gene is not part of the model signature. The higher the mean decrease Gini, the higher the weight of the gene during decision-making at each decision tree node.

The analysis categorised the samples into three groups: CD non-responders, CD responders and non-IBD controls. The OOB for the random forest model was 11.78%, which is very low for human data. The variables with the highest mean decrease accuracy and mean decrease Gini were *IL1B* and *CHI3L1*, identifying them as the genes with the most impact in the model. The AUC values for the prediction of all three categories (CD non-responders, CD responders and non-IBD) were 1.000, reflecting the results of the ROC analysis of the individual genes. These findings indicated that the expression level of genes in the immune cluster could distinguish non-responders from responders to anti-TNF within colonic CD lesions without fail. It should be noted that obtaining an AUC of 1 is unusual in human data. Therefore, these analyses should be repeated in an independent anti-TNF naïve cohort at the baseline of treatment initiation to confirm these findings.

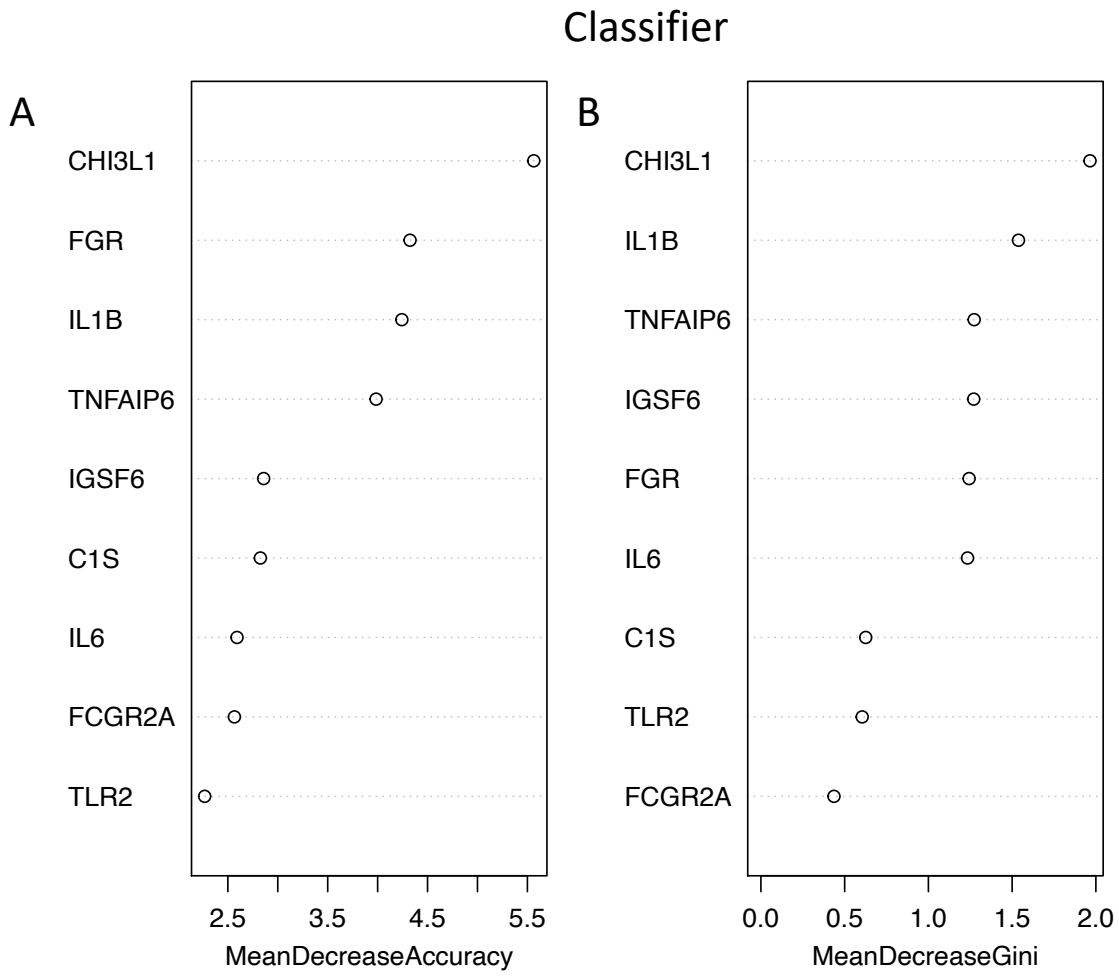


Figure 5.13: ***Random Forest output of prediction signature for anti-TNF treatment response in GSE16879 dataset.***

A) Random Forest classifiers mean decrease accuracy. B) Random Forest classifier mean decrease Gini. All analysis was performed in R studio (R version 4.0.2 and below) and required the *randomForest* package.

5.6 Quantification of CHI3L1 and *IL1B* expression in colonic tissue sections of CD patients and non-IBD controls using histology assessments

After identifying the genes with the highest predictive power for anti-TNF non-response within immune cluster 2, I wanted to test whether these targets can be quantified in colonic tissue in an additional cohort. I used IHC and *in situ* hybridisation to quantify protein and gene expression in 21 CD patients with endoscopically confirmed active disease and 15 non-IBD controls (*Table 5.2*). Colonic pinch biopsies were collected during routine endoscopies and fixed in

formalin for up to 72 hours (*Table 5.3*). The samples were embedded in paraffin before IHC and *in situ* hybridisation (ISH) staining. To exclude non-specific binding as well as background staining from the analysis, negative controls were included in the form of isotype antibody staining and staining without primary antibody (referred to as "no primary"). IHC and ISH staining was quantified using QuPath (Bankhead *et al.*, 2017), a digital quantitative pathology software. It can be used to recognise cells in tissue sections based on their shape and colour and can identify them as either positive or negative for specific staining characteristics.

	CD (n= 21)	**NC (n= 15)**
Median Age (Range)	31.8 (16-74)	42 (19-76)
M:F ratio	1.1	0.85

Table 5.3: ***Summary of cohort 1 information.***

21 CD patient samples and 15 non-IBD volunteers were recruited at the GRI endoscopy clinics. Colonic biopsies were collected in formalin and were subsequently fixed for up to 72hrs before further processing.

5.6.1 CHI3L1 protein expression in intestinal tissue of CD patients

In *section 5.5*, *CHI3L1* was identified as one of the genes with the highest predictive power to separate CD patients into responders and non-responders to anti-TNF. To understand whether CD patients have distinct patterns of CHI3L1 expression, it was measured in an additional patient cohort on a protein level; an anti-CHI3L1 primary antibody was used for IHC (*Figure 5.14*). CHI3L1 was found to be expressed in the lamina propria as well as the epithelial cells surrounding the crypts. Unlike non-IBD controls, CD patients frequently displayed deposits of CHI3L1 (*Figure 5.14A*). Only one non-IBD control exhibited the same staining pattern. Even though no significant difference was detected between the mean frequencies of CHI3L1-positive cells between normal and CD, patients could be differentiated into two groups based on their intestinal CHI3L1 protein expression (*Figure 5.14B*). Patients with low CHI3L1 expression exhibited less than 20% positive cells.

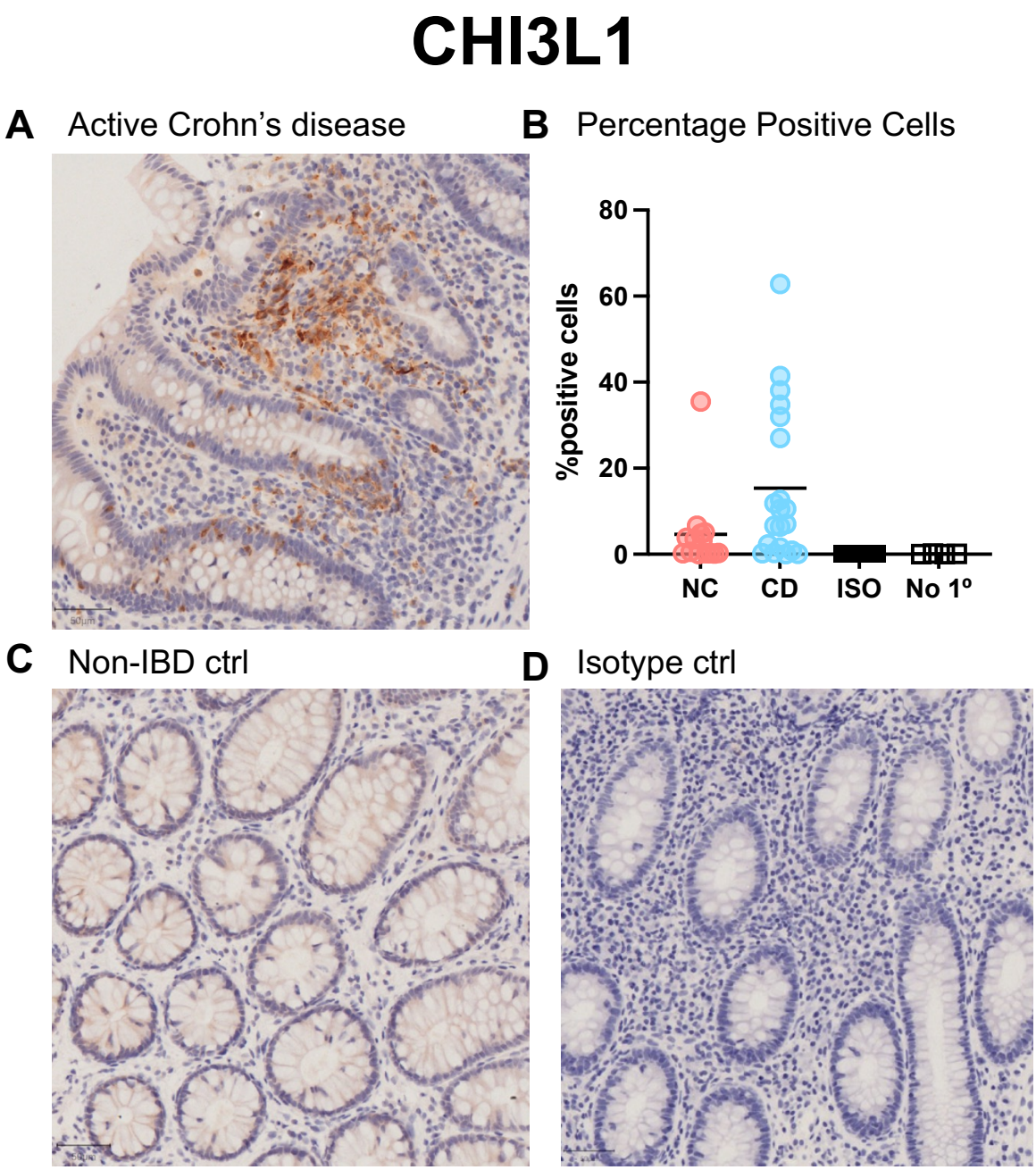


Figure 5.14: ***The quantification of CHI3L1 expression in intestinal pinch biopsies of CD patients and non-IBD controls by IHC.***

Intestinal pinch biopsies of 21 CD patients and 15 non-IBD controls were collected, formalin fixed, and paraffin-embedded. A) Representative CHI3L1 staining in diseased intestinal tissue. B) Quantification of percentage positive cells of all detected cells showing non-IBD controls (NC, red), CD patients (CD), isotype (ISO) and no primary antibody (No 1°). C) Representative CHI3L1 staining in normal intestinal tissue. D) Negative control using isotype control instead of primary CHI3L1 antibody. Data were generated in four independent experiments with age and sex-matched NC and CD samples. Data were analysed using QuPath (version 0.3.2 and below). Lines across the dot plot represent the mean. A Kruskal-Wallis followed-by Dunn's post-test as performed in B). Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below).

5.6.2 IL1B spatial RNA expression in intestinal tissue of CD patients

Detecting IL-1ß with IHC is problematic because it has two forms: before and after cleavage in the inflammasome. When using an anti-IL-1ß antibody, the main detection is focused on the cleaved IL-1ß protein released into the tissue. The previous analysis quantified the gene expression, i.e., the mRNA of IL-1ß currently or about to be transcribed within the cells, and not the soluble protein. To quantify *IL1B* transcription and compare its spatial distribution to CHI3L1, the ISH method was used, which allowed to hybridise and amplify *IL1B* mRNA in FFPE-fixed intestinal tissue. I attempted to co-stain the *IL1B* ISH with an anti-cluster of differentiation 68 (CD68) antibody using IHC to identify intestinal macrophages (*further discussed in section 1.8*). However, the protease cleavage step required to ensure access to RNA for the target probe cleaved the antibody:CD68 epitope binding and only ISH staining was detectable. Because of the strong signal in the positive control, the ISH *IL1B* data was still analysed and used. *Figure 5.15A* shows the expression signal of the *IL1B* probe, positive control, a probe against human *PPIB*, and negative control (a probe detecting the bacterial *dapb* gene). The expression was quantified using automated QuPath analysis and compared in *Figure 5.15B*. No significant difference was detected between non-IBD controls and CD patients (*Figure 5.15C*).

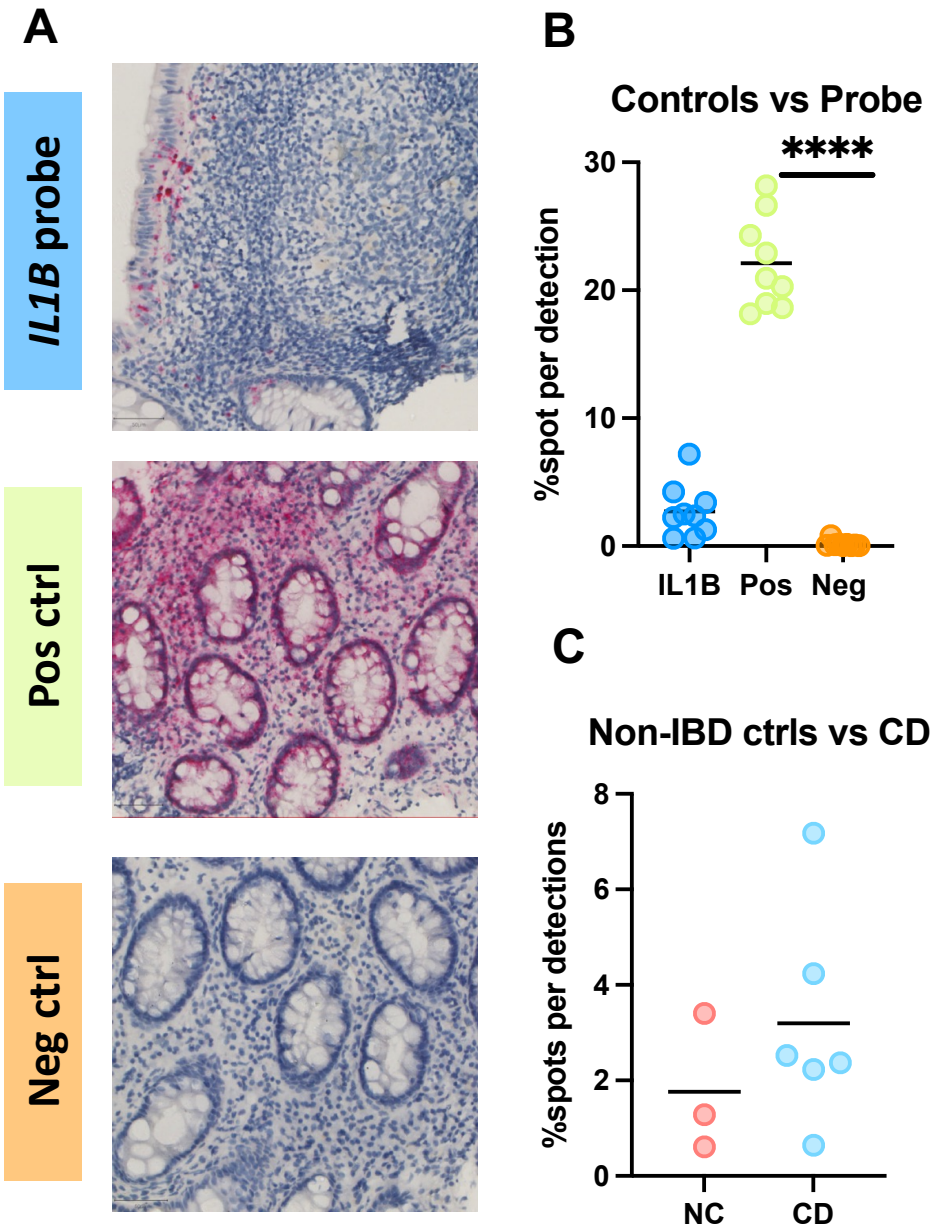


Figure 5.15: ***The quantification of IL1B mRNA expression in intestinal pinch biopsies of CD patients and non-IBD controls by ISH.***

Intestinal pinch biopsies of 6 CD patients and 3 non-IBD controls were collected, formalin fixed, and paraffin embedded. A) Representative staining of target *IL1B* gene, positive control *PPIB* gene (pos ctrl) and negative control *dapB* gene (neg ctrl) probes. B) Comparison of percentage mRNA spots per cell detections within intestinal tissue comparing target and control probes. D) Comparison of *IL1B* mRNA expression between non-IBD controls (NC) and CD patients (CD). Data was analysed using QuPath (version 0.3.2 and below). Lines represent the mean. One-way ANOVA with Holm-Sidaks's post-test was performed in C). Unpaired T-Test was performed in C). Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below). ****p<0.0001. *No 1º*: No Primary Antibody control. *Iso*: Isotype control.

5.7 Quantification of *CHI3L1* and *IL1B* gene expression in colonic tissue of CD patients and non-IBD controls

qPCR is another clinically available method that can be used for potential biomarker screening. An additional patient cohort, referred to as cohort 2, was recruited with matched intestinal biopsies collected in RNA*later*™ Stabilization Solution (Cat. No. AM7020, Thermo Fisher) (*Table 5.3*). To investigate the potential link between inflammation and the *IL1B*/*CHI3L1* signature, patients with active and inactive disease, determined by endoscopic examination, were included.

	CD (n= 18)	**NC (n= 14)**
Median Age (Range)	50.67 (30-77)	55.1 (19-73)
F:M ratio	0.8	0.75

Table 5.3: ***Summary of cohort 2 information.***

18 CD patients and 14 non-IBD volunteers were recruited at the GRI endoscopy clinic. Colonic biopsies were collected in RNA*later*™ Stabilization Solution (Cat. No. AM7020, Thermo Fisher) and stored at -80°C until RNA extraction.

qPCR analysis of this cohort found that *CHI3L1* ($p = 0.0118$) and *IL1B* ($p = 0.0181$) exhibited an increase in CD patients compared to non-IBD controls (*Figure 5.16A*). When stratified into CD active and remission, *CHI3L1* expression also distinguished active from inactive disease ($p = 0.0189$) (*Figure 5.16B*). Like previous IHC and transcriptomic data, *CHI3L1* and *IL1B* expression could be used to stratify CD patients into two groups.

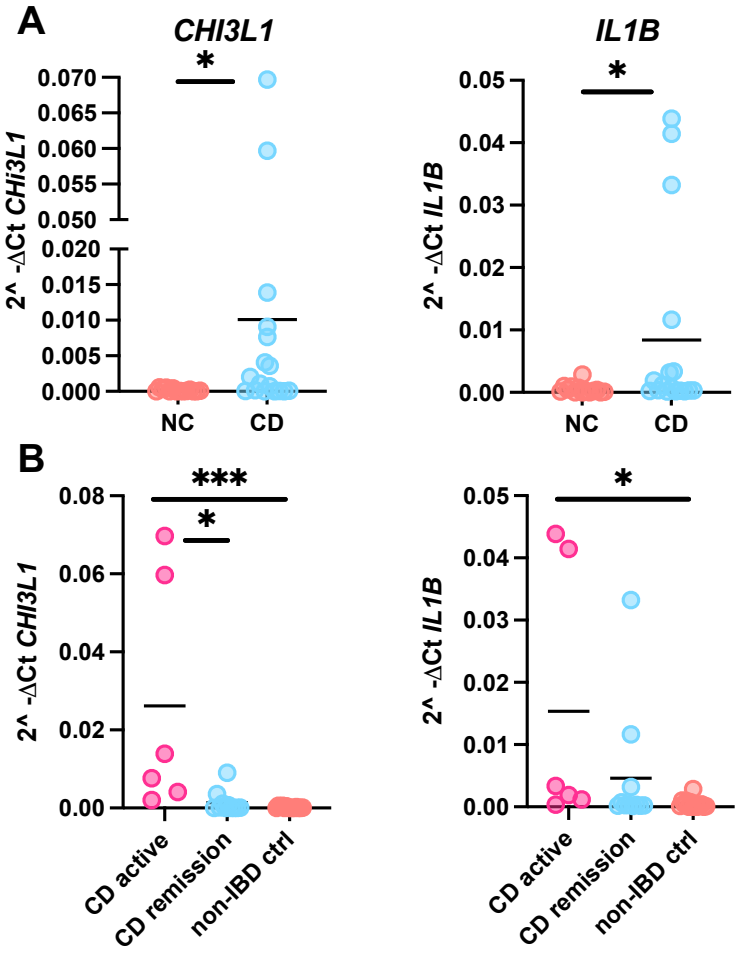


Figure 5.16: ***The quantification of CHI3L1 and IL1B mRNA expression in intestinal pinch biopsies of CD patients and non-IBD controls by qPCR.***

Intestinal pinch biopsies of 17 CD patients and 14 non-IBD controls were collected and stored in RNA fixative immediately before RNA isolation. Data were analysed using the 2^-ΔCt method using the human large ribosomal protein gene (*RPLPO*) as a housekeeping gene for normalisation purposes. A) Comparison of *CHI3L1* (left panel) and *IL1B* gene expression between non-IBD controls (NC) and CD patients. B) Comparison of *CHI3L1* (left panel) and *IL1B* gene expression between non-IBD controls (NC) and CD active and remission patients. Data were generated in four independent experiments with age and sex-matched NC and CD samples. Lines represent the mean. Mann-Whitney was performed in A). Kruskal-Wallis was performed in B). Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below). *p<0.05, ***p<0.001.

5.8 CD68 protein expression in intestinal tissue of active CD patients and non-IBD controls

After quantifying CHI3L1 protein and mRNA and *IL1B* mRNA, I next aimed to understand which cells in particular are responsible for their expression. The analyses in *section 5.4* suggested that neutrophils, mast cells, and macrophages are potentially linked to the immune cell cluster 2. CD68 is a lysosomal-associated membrane protein commonly used as a histological marker to identify macrophages (Ferenbach and Hughes, 2008). It is located within the lysosome and has been found to be associated with pro-inflammatory and tissue repair functions of macrophages. Even though the majority of literature relates to it as a macrophage marker, it has been shown also to be upregulated in neutrophils during inflammation in patients with IBD (Amanzada *et al.*, 2013). *Figure 5.17* shows the expression of CD68 in intestinal tissue of patients in the IHC cohort *(*cohort 1, *summarised in Table 5.2)* with active CD compared to non-IBD controls. No significant difference was detected between patients and controls; however, intestinal tissue affected with CD showed a wider spread of CD68 expression *Figure 5.17B*.

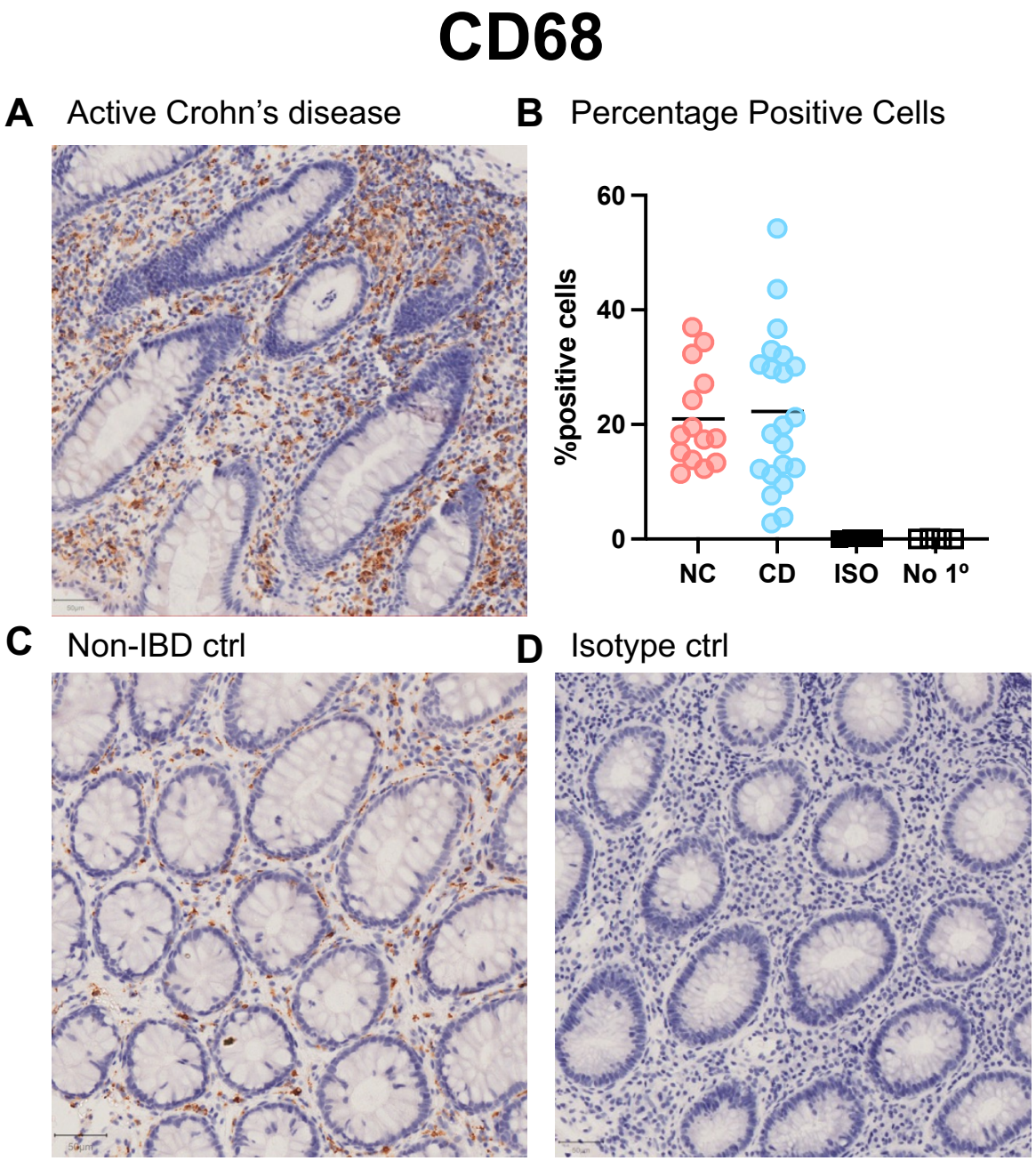


Figure 5.17: ***The quantification of CD68 expression in intestinal pinch biopsies of CD patients and non-IBD controls by IHC.***

Intestinal pinch biopsies of 21 CD patients and 15 non-IBD controls were collected, formalin fixed, and paraffin-embedded. A) Representative CD68 staining in diseased intestinal tissue. B) Quantification of percentage positive CD68 cells of all detected cells showing non-IBD controls (NC, red), CD patients (CD), isotype (ISO) and no primary antibody (No 1°). C) Representative CD68 staining in normal intestinal tissue. D) Negative control using isotype control instead of primary CD68 antibody. Data were generated in four independent experiments with age and sex-matched NC and CD samples. Data were analysed using QuPath (version 0.3.2 and below). Lines represent the mean. A Kruskal-Wallis followed-by Dunn's post-test as performed in B). Normal distribution was tested using Anderson-Darling, D-Agostino-Pearson, Shapiro-Wilk and Kolmogorov Smirnov tests. All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below).

5.9 Correlation of target protein and gene expression

The signature identified in *section 5.3* showed significant correlations between *TLR2, CHI3L1* and *IL1B*. To test whether this was the case in the additional patient cohorts, linear regression analysis was performed between the individual targets (*Figure 5.18*). No significant correlation was found between CHI3L1 and CD68 protein expression in cohort 1 (*Figure 5.18A*). When comparing the expression of CHI3L1 protein and *IL1B* mRNA expression, a positive correlation was detected (R^2 =0.7818, p =0.0161) (*Figure 5.18B*). Similarly, when correlating *CHI3L1* and *IL1B* expression quantified by qPCR in cohort 2 the two gene targets significantly correlated (R^2 =0.7844, p <0.0001) (*Figure 5.18C*). Due to small sample size one patient was identified with the CHI3L1/*IL1B* signature in cohort 1 and two in the second cohort.

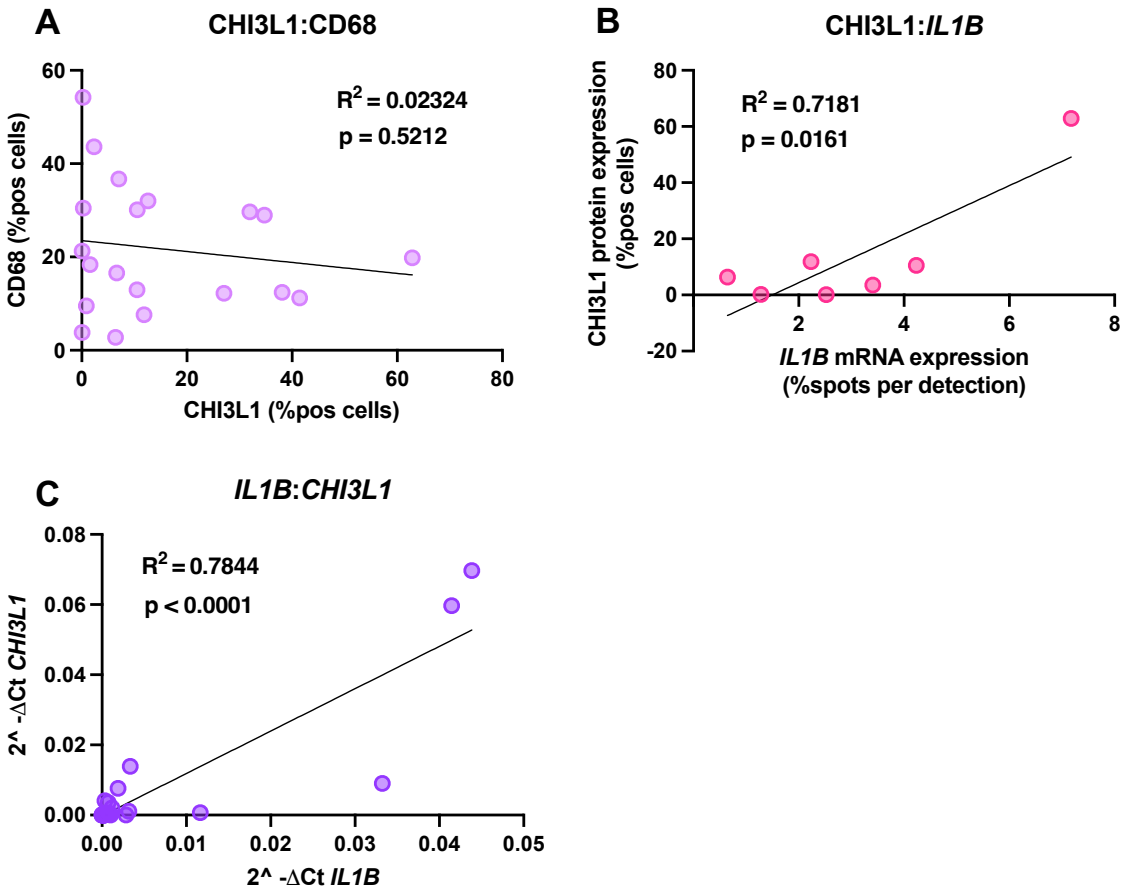


Figure 5.18: ***Correlation of target expression in intestinal pinch biopsies of CD patients and non-IBD controls.***

A) Linear regression between CHI3L1 and CD68 protein expression quantified by IHC in *sections 5.6.1* and *5.8.* in cohort 1 B) Linear regression between CHI3L1 protein and *IL1B* mRNA expression quantified by IHC (*section 5.6.1*) and ISH (*section 5.6.2*), respectively, in cohort 1. C) Linear regression between *IL1B* and *CHI3L1* mRNA expression quantified by qPCR (*section 5.7*) in cohort 2. Positive cells for IHC and ISH experiments were measured using QuPath (version 0.3.2 and below). qPCR data were analysed using the 2^-ΔCt method. All statistical analysis was performed in GraphPad Prism (Version 9.3.1 and below).

5.10 Co-localisation

Co-expression may indicate associations between genes and proteins; however, co-localisation can provide more information about which cell types and structures express the targets. Figure 5.19A/B shows representative IHC data from two CD patients displaying CD68 and CHI3L1 expression. Areas on the tissue section exhibiting positive staining for anti-CHI3L1 antibody also appeared to exhibit high $CD68^{+}$ cell infiltration. The tissue structure displayed histological characteristics of a granuloma response, including an expanded lamina propria, leukocyte infiltration and general distortion of tissue architecture.

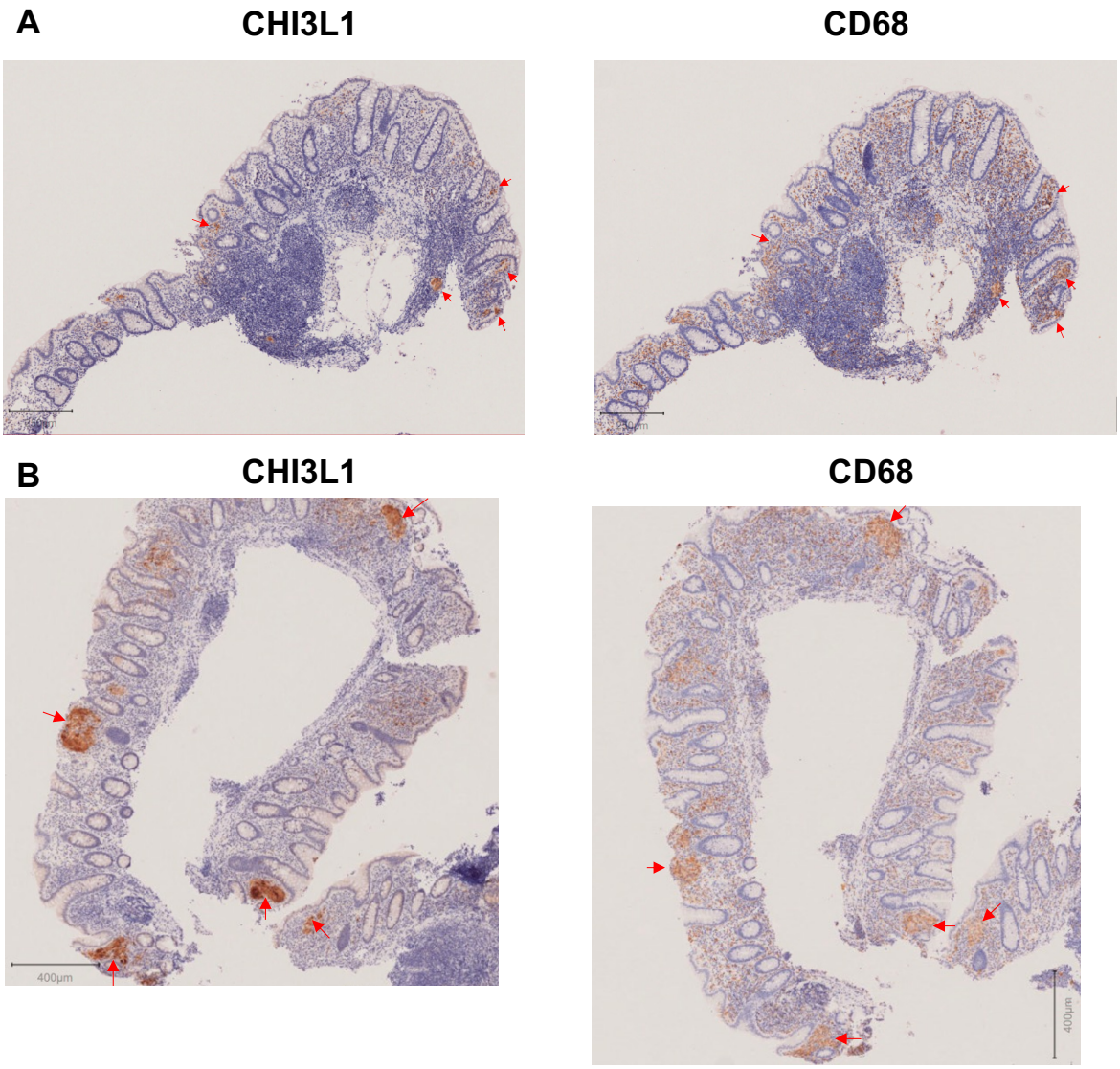


Figure 5.19: ***Co-localisation of CHI3L1 and CD68 in diseased intestinal tissue from cohort 1.***

Intestinal pinch biopsies of 21 CD patients and 15 non-IBD controls were collected, formalin fixed, paraffin-embedded, and stained with anti-CD68 and anti-CHI3L1 primary antibody for IHC. A) CHI3L1 protein expression in intestinal tissue of two representative CD patients. B) CD68 protein expression in intestinal tissue of two representative CD patients. Red arrows indicate CHI3L1 infiltrated areas.

5.11 Discussion

This chapter aimed to stratify patients based on their intestinal immune signatures. *Section 5.2* compared bulk RNA sequencing data from colonic pinch biopsies between active CD and non-IBD controls. Differential gene expression analysis of the two sample groups showed that 978 genes were significantly up or downregulated. Biological theme comparison analyses tested for enriched gene sets in up and down-regulated DEGs. Most gene sets were enriched for immune cluster 2 upregulated pathways, including IBD and other chronic inflammatory

conditions such as RA, SLE, and asthma (*Figure 5.5*). Overall, based on the differential gene expression and biological theme comparison analysis, intestinal tissue affected with CD unsurprisingly exhibited abnormal and chronic inflammatory characteristics.

After disease-associated differences in mucosal colonic gene expression were identified, CD patients were stratified based on their immune system process genes *(Figure 5.6/7*). This highlighted a cluster of 116 genes, which had selectively enriched gene expression in a subset of patients, referred to as immune cluster 2. One patient exhibited enrichment for immune cluster 2 in the bulk RNAseq data, and so did 7 patients in the microarray validation dataset GSE16879, referred to as CD Up patients. GSE analysis suggested enrichment in TLR signalling as well as complement and inflammasome activity in CD Up patients (*Figure 5.8*). After taking a closer look at the immune mechanisms underlying immune cluster 2 response, I moved on to understanding which immune cells were mainly involved. GSE and pathway analyses showed a link to a neutrophil/granulocyte response and macrophage-induced tissue repair. Digital cytometry, used for deconvolution of bulk transcriptomics to estimate immune cell composition, suggested an increase in the proportion of M1 macrophages and activated mast cells in colonic samples of the patient with the upregulated immune cluster 2 when compared to unaffected patients in the bulk RNAseq dataset (*Figure 5.9*). When looking further into intestinal macrophages, the CD Up patient subset in the GSE16879 data had increased intestinal macrophage-associated signatures compared to the rest of the CD cohort (*Figure 5.10*).

In terms of the TLR activity, TLRs 2 and 8 were part of this distinct immune cluster. TLR2 forms heterodimers with either TLR1 or TLR6 for signaling (Takeda and Akira, 2005). This receptor complex recognises gram-positive bacteria and induces NF-kB and MAPK-dependent signalling. *TLR2* variants have been linked to IBD and other mycobacterial infections (e.g., tuberculosis and mycobacterial lung disease) (Yim *et al.*, 2006, 2008; Bank *et al.*, 2014). TLR8 can signal on its own and responds to viral stimuli, triggering the MYD88 cascade (Takeda and Akira, 2005). UC disease severity correlates with the expression of *TLR8*, potentially linking it to colon-specific inflammation (Sánchez-Muñoz *et al.*, 2011). The complement system plays a vital role in maintaining barrier function in healthy and diseased intestines (Sina, Kemper and Derer, 2018). It has been shown that mice lacking as few as one gene

from the complement cascade results in stronger intestinal inflammation in a chemical-induced mouse model of chronic colitis (Johswich *et al.*, 2009; Elvington, Schepp-Berglind and Tomlinson, 2015). The gene encoding for complement component 4 (*C4*) has been identified as a CD susceptibility gene by GWAS (Cleynen *et al.*, 2016). Expression of complement components *C3* and *C4* were predominantly found in intestinal epithelial cells (IECs) located in the crypts as well as macrophages in the submucosal layer of CD patients using IHC and ISH (Laufer *et al.*, 2001), again suggesting macrophage involvement. Inflammasome activity, which induces the cleavage of non-functional pro-IL-1β and pro-IL18, has been shown to positively correlate with CD disease severity (Reinecker *et al.*, 1993; Swanson, Deng and Ting, 2019), underlining its importance in the driving of pathogenesis. The downstream effect of all these inflammatory mediators is damage to the intestinal mucosa, requiring constant repair, with mucosal healing being the final stage of recovery from a CD flare.

When looking at the clinical response of patients to anti-TNF treatment, non-responders to anti-TNF blocking agents exhibited an increase in the immune cluster 2 gene signature expression. Anti-TNF therapy refraction poses a big issue in disease management of CD since other biologic treatments are often less effective post previous TNF blocking (Sands *et al.*, 2017; Verstockt *et al.*, 2020). The genes most effective in predicting non-response were *IL1B, IL6, IGSF6, FGR, CHI3L1, FCGR2A, TNFAIP6, TLR2* and *C1S* (*Figure 5.13)*. It should be noted that 6 out of the 9 genes (apart from *FCGR2A*, *IGSF6* and *TLR2*) were part of the intestinal macrophage signatures used in *Figure 5.10,* suggesting that they are linked to macrophage activity in IBD. Furthermore, most of these genes have been associated with inflammatory diseases, including IBD.

FCGR2A is known to induce phagocytosis, and a gene variant has been linked to IBD (Weersma *et al.*, 2010). Serum TNFAIP6 protein expression has been suggested to function as a biomarker for disease activity in both CD and UC (Yu *et al.*, 2016), and its mucosal gene expression been previously linked to anti-TNF non-response by Arijs *et al.* (2010). It should be noted that this publication generated the GSE16879 dataset used in this study to identify links to anti-TNF non-response and this association has not been confirmed in a study with an additional patient cohort. CHI3L1 is a non-enzymatic hydrolase that can be secreted by macrophages, neutrophils, and epithelial cells in response to specific immune

challenges (Zhao *et al.*, 2020), linking it to both neutrophil and macrophage activity suspected in immune cluster 2. It has been suggested as an autoantigenic target in IBD (Deutschmann *et al.*, 2019), found to induce macrophage recruitment in colorectal cancer (Kawada *et al.*, 2011) and to also drive epithelial cell proliferation and tissue repair (He *et al.*, 2013). These findings support the hypotheses that neutrophil and macrophage activity/infiltration can be found in CD Up patients. Furthermore, macrophage-induced tissue repair/fibrosis might also take place.

IL-1β encoding gene as well as protein expression has previously been identified as a predictor of non-response after treatment initiation of anti-TNF blocker because its expression remains upregulated in non-responders during treatment (Leal *et al.*, 2015). The same study showed that other inflammatory cytokines, such as IL-6 (also part of the immune cluster 2 signature), were altered in response to the therapeutic, and their gene expression was downregulated, suggesting that IL-1β-dependent pathways drive disease during ant-TNF treatment. IL-1β requires cleavage of its nonfunctional pro-IL-1β form *via* an inflammasome response. The gene encoding the inflammasome mediator NLPR3 was also upregulated in the immune cluster signature, supporting the hypothesis that an increase in IL-1β signaling might be responsible for driving non-response in patients with anti-TNF refractory patients. In addition to these findings, Friedrich *et al.* (2021) found that IL-1β-dependent interaction between stroma and neutrophils is another hallmark of anti-TNF non-response. The levels of plasma IL-1β levels have also been found to be upregulated in anti-TNF non-responders at baseline, and the capacity of their blood-derived monocytes to produce IL-1β was higher compared to responders (Gaiani *et al.*, 2020). Bösl *et al.* (2018) found that co-stimulation of TLR2 and TLR8 of human monocytes induced an inflammatory response causing enhanced secretion of the neutrophil chemoattractants IL-8 and CXCL1 as well as suppressed TNFα, IL-1α, IL-1β, and IL12p70 production. In the same study they showed *in vitro* that blood-derived neutrophils exhibited a higher migration potential when exposed to supernatant of co-stimulated cultures compared to single TLR stimuli. The signature identified in this chapter showed that *IL1B* was not suppressed in response to the TLR2 and TLR8 pathway because the gene expression values positively correlated, contradicting the results from Bösl *et al.* (2018). However, their study looked at cells derived from healthy volunteers, so

it is a possibility that my findings relate to an aberrant TLR functionality in CD Up patients. This aberrant signaling might be in response to TLR2:TLR8 co-stimulation in intestinal macrophages inducing IL-1ß production, which in turn drives neutrophil infiltration.

In *Figure 5.14* CHI3L1 positive staining could be seen around and in crypts of both non-IBD controls and CD patients. CHI3L1 has been found to be expressed by epithelial cells in human intestines (Mizoguchi, 2006), therefore the IHC signal found in and around the crypts potentially stems from the epithelial layer. Apart from CHI3L1 staining around the crypts and sparse positive signals in the LP, the protein was highly accumulated in granuloma-like structures. The same areas were also affected by influx of $CD68^+$ cells. Granuloma have been linked to a selection of clinical phenotypes in CD, including severe disease and inflammation penetrating the mucosal layer (Molnár *et al.*, 2005). They are formed by the accumulation of epithelial cells and macrophages in response to a continuous stimulus whilst they are struggling to clear the infection (Pagán and Ramakrishnan, 2018). TGF-ß1 and the IL-4/IL-13 axis are thought to drive granuloma formation (Pagán and Ramakrishnan, 2018). Interestingly, CHI3L1 has been shown to promote tumour growth and metastasis *via* TGF-ß1 stimulation in mice (He *et al.*, 2013; Geng *et al.*, 2018; Qiu *et al.*, 2018). Additionally, TLR2 has previously been shown trigger granuloma formation in response to tuberculosis (Reiling *et al.*, 2002). Huppertz *et al.* (2020) showed CD68 deposition in the lung granuloma of sarcoidosis. Colonic granulomas in chronic granulomatous disease (CGD), a condition with a genetic defect in phagocytosis, have shown infiltration of macrophages. Interestingly, when quantifying $CD68^+$ cells in affected intestinal tissue of CDG, CD and non-diseased controls, CDG patients showed a significant decrease (Liu *et al.*, 2009). In sarcoidosis Huppertz *et al.* (2020) also identified co-localisation of the granuloma with IL-1ß and caspase-1, a crucial component of the NLPR3 inflammasome. Using ISH, I did not detect the same accumulation of *IL1B* in the granuloma structures as described by Huppertz *et al.* (2020) in sarcoidosis. However, I used the assay together with anti-CD68 antibody staining, which could have potentially interfered with the hybridisation of the *IL1B* probe. The experiment should be repeated without the IHC co-stain to increase the quality of the data.

Overall, this study identified a potential subset of CD patients with enhanced granulomatous disease. Based on this observation, a granuloma response may explain why the cohort exhibited patient groups defined by low and high $CD68^+$ cell counts. Furthermore, the literature in the last paragraph suggests a potential involvement of CHI3L1 in driving TGF-ß1-dependent granuloma formation and TLR2 and IL-1ß-dependent signaling of $CD68^+$ cells, a marker for neutrophil and/or macrophages, in contributing to the initiation or progression of the response.

5.12 Conclusion

Bulk RNA sequencing elucidated differences in the gene expression between intestinal tissue from healthy individuals or people with CD. Stratifying the colonic transcriptomic profile of CD-affected lesions based on immune system-associated genes identified a subset of patients with a distinct inflammatory signature. Pathway and GSE analysis suggested a TLR-driven pathway, with complement and inflammasome activation and NF-kB signaling. Additionally, together with digital cytometry, these analyses identified neutrophil activity as well as healing responses potentially due to a macrophage tissue repair response. When looking at clinical characteristics, patients with an increase in the distinct colonic inflammatory signature were all refractory to anti-TNF treatment. Using ROC and random forest models to test the predictive power of individual genes in the signature and all genes combined, I found that both analyses predicted anti-TNF non-response with 100% accuracy. Subsequently, it was shown that CHI3L1-encoding gene and protein as well as *IL1B* expression has the potential to stratify patients into two distinct groups. CHI3L1 protein and *IL1B* mRNA expression were found to correlate in cohort 1. Additionally, upregulation of *CHI3L1* mRNA in colonic tissue of cohort 2 distinguished active disease from remission. In terms of co-localisation, the CD68 expression of the macrophage/neutrophil marker was co-localised in tissue areas with high levels of CHI3L1 protein deposition.

Even though a patient subset with increased CHI3L1 and *IL1B* levels was identified in two additional cohorts, the response of these patients to anti-TNF treatment is unknown. Ongoing work in the Milling lab is currently investigating whether histology and qPCR assessments can be used to quantify these two targets and

whether their expression levels are predictive for anti-TNF treatment response in an additional cohort.

Chapter 6: General Discussion

The main aims of my project were to stratify CD patients based on their PB and intestinal immunopathotypes, and link distinct phenotypes to clinical parameters related to the state of intestinal inflammation or treatment response. I hypothesised that CD patients exhibit different immune signatures in their blood and intestine and that these link to specific disease phenotypes and treatment responses. In *chapter 3,* I found that although patients were classified as relapsing/disease active based on HBI they did not show consistent changes in their peripheral blood (*Figure 3.9*). However, a subset of patients with a distinct PB phenotype was identified, characterised by reduced levels of T cell subpopulations (*Figure 3.16, summarised in Figure 6.1*). Analyses of matched blood and biopsy samples showed increased cytotoxic, anti-bacterial and anti-viral immune responses in patients with low T cells (*Figure 4.16*). Additionally, overall $CD8^+$ T cells or gene expression of genes associated with this population were consistently found to be reduced in PB (*Figure 4.1, 4.7,4.8, 4.11*) together with α4ß7-encoding gene expression (*Figure 4.21*/*4.22*).

The potential of PB and intestinal $CD8^+$ T cells to function as biomarkers was studied by Boschetti *et al.* (2016). They investigated $CD8^+$ T cell levels in 73 CD patients that have recently undergone ileal resection surgery. They linked an increase in mucosal T cells to the recurrence of disease within a year of surgery. However, they also identified an increase of $CD8^+$ cells in the circulation of the patients undergoing relapse, contradicting my findings in *chapter 4.* The monoclonal antibody etrolizumab binds the integrin subunit ß7. It inhibits immune infiltration into the intestines by blocking α4ß7 and promotes egress to draining lymph nodes by blocking αEß7, which has been shown in a murine model of colitis (Dai *et al.*, 2021). A randomised placebo-controlled phase 3 trial of etrolizumab in CD patients with moderate to severe disease showed a decrease in intestinal $CD8^+$ T cells 14 weeks post-treatment initiation (Sandborn *et al.*, 2020; Dai *et al.*, 2021). Patients with the LT blood phenotype, which includes an increase in intestinal $CD8^+$ T cells could potentially be a target group for this therapeutic. Another treatment already in clinical use is vedolizumab which blocks α4ß7 only. A large study (*n* =1,115) testing vedolizumab as maintenance therapy for CD found initial response in 60% of patients (assessed at week 4) and long-term response in

36-39% of patients, depending on dosage and frequency of administration (Sandborn *et al.*, 2013b). Low T cell levels in PB might be an indicator of their migration to the intestines, thus the LT patient phenotype could be a biomarker for response to vedolizumab.

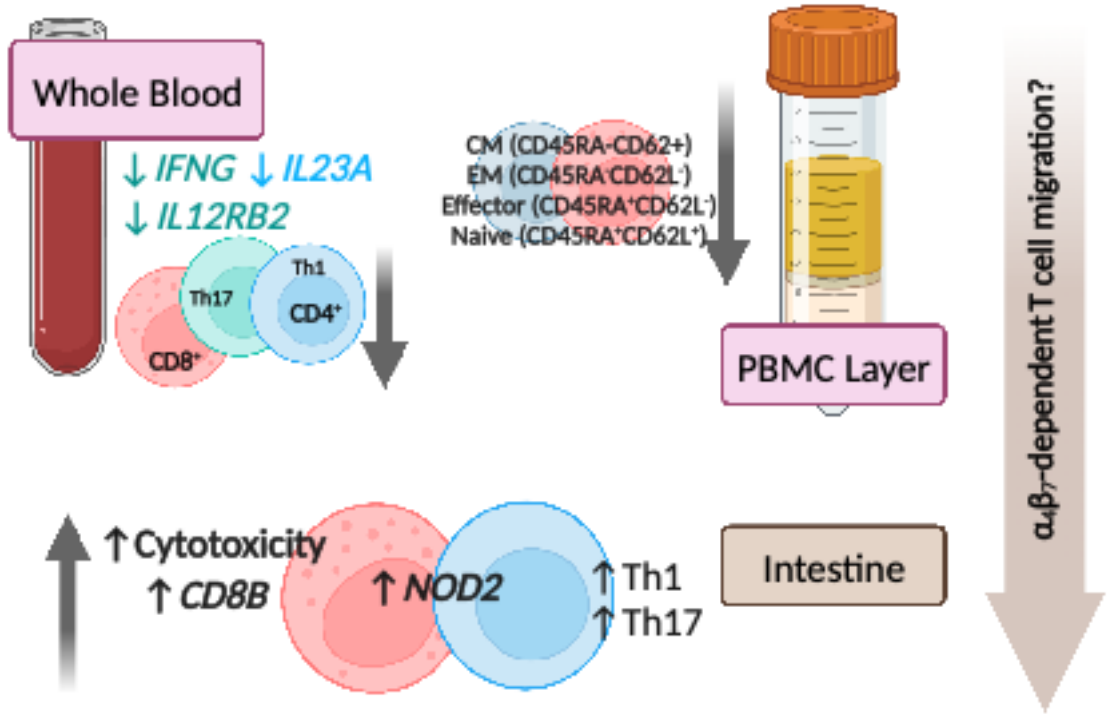


Figure 6.1: ***Overview of LT phenotype identified in*** **chapters 3** ***and*** **4.**

Patients with the LT phenotype exhibited lower levels of T cell subpopulations (both $CD4^+$ and $CD8^+$ in their peripheral blood. Additionally, Th1 and $CD8^+$ T cells were reduced based on digital cytometry analysis in adult CD patients, as was the gene expression of *IFNG*, *IL12RB2* (Th1 markers) and *IL23A* (Th17 markers). They also had reduced levels of the genes encoding for α4β7 integrin suggesting that T cells positive for this gut-homing marker may have migrated to the intestines. In their intestine, the LT patients had an increase in Th1/Th17 immune responses, usual for CD pathology, $CD8^+$ T cells and high levels of cytotoxicity. They also showed more intestinal *NOD2* expression, a genetic risk factor for CD. This figure was created using Biorender.com.

Further research to test these hypotheses a study would require the collection of matched blood and intestinal tissue of inflamed patients with endoscopically confirmed active disease at baseline and follow-up after vedolizumab/placebo treatment. I would not suggest studying etrolizumab because it is currently not approved for CD treatment. PBMCs and LPMCs as well as RNA from whole blood and intestinal tissue, should be isolated and analysed based on their T cell phenotype. They can be used for extensive T cell phenotyping using flow cytometry, including extracellular and intracellular staining to determine different Th, $CD8^+$ and Treg subpopulations. Gene expression analysis either uses clinically available methods such as qPCR or more expensive but more informative

analyses with scRNAseq to identify changes in T cell profiles and potentially other immune cells related to the LT PB phenotype. I predict that patients responding to vedolizumab during follow-up exhibit the LT patient phenotype and have distinct intestinal gene expression at baseline compared to the non-responders.

My PhD project also identified an intestinal gene expression signature potentially associated with non-response to anti-TNF treatment (*Figure 5.13*). Hallmarks of this signature were CHI3L1^{+} granulomas and increased *IL1B* expression as well as neutrophil and macrophage activity, supporting previous findings from Friedrich *et al.* (2021) who identified an increase in IL1-ß-driven neutrophil/stroma activation in anti-TNF non-responders. Granuloma formation has been previously linked to severe CD (Molnár *et al.*, 2005). Anakinra, an IL-1 receptor antagonist, is currently being tested as a first-line therapeutic in 214 UC patients as part of the IASO trial (Thomas *et al.*, 2019). The results of this trial are projected to be published later in 2022. Pre-clinical data generated in murine models have already shown promising results. In Winnie-TNF knockout mice, modelling anti-TNF refractory UC, and Inositol polyphosphate 5′-phosphatase–deficient (SHIP) knockout mice, mimicking ileal CD (McLarren *et al.*, 2011), anakinra significantly reduced inflammation (Ngoh *et al.*, 2016; Liso *et al.*, 2022). Winnie-TNF knockout mice spontaneously develop TNF-independent colitis due to a missense mutation in the gene encoding mucin 2 and deletion of *TNF*. SHIP$^{-/-}$ mice also exhibit spontaneous inflammation, which is limited to their ileum and has a similar phenotype to CD, and develop granuloma-like structures in their intestines, similar to those in the non-responder signature identified in *chapter 5* (McLarren *et al.*, 2011). These pre-clinical data, together with my findings, suggest that anakinra might be a suitable alternative for anti-TNF treatment for patients with this immunopathotype.

Testing the predictive power of signature target gene expression and the presence of CHI3L1^{+} granuloma in intestinal tissue of responders and non-responders at baseline in an additional cohort would be the next step in validating this signature. To further understand the molecular mechanism underlying the non-response, the targets should also be quantified in different immune cells. Multiplex immuno-fluorescent staining is a valuable tool for studying the co-localisation of IL1B, CHI3L1 and TLR2 with neutrophils, macrophages, mast cells and epithelial cells.

Since macrophages and neutrophils share many markers, including CD68 (Amanzada *et al.*, 2013), co-staining with macrophage and neutrophil-specific markers (such as TREM1 (Caër *et al.*, 2021) and CD66b (Therrien *et al.*, 2019b)) with the signature proteins will help understand which cell types express which target. SHIP$^{-/-}$ mice exhibit a similar phenotype to the non-responders. Further characterisation and comparison of the inflammatory signatures in mice and humans may also be helpful to understand their potential as a representative mouse model for the immunopathoype. An enteroid-myeloid organoid model developed by Staab *et al.*, 2020 could be used to study further the interaction between cell types in a human disease model. It allows for colon-derived epithelial organoids to be co-cultured with macrophages. Using this model in a subset of CD patients at anti-TNF treatment baseline can aid in elucidating the interaction between epithelial cells and macrophages in anti-TNF non-responders.

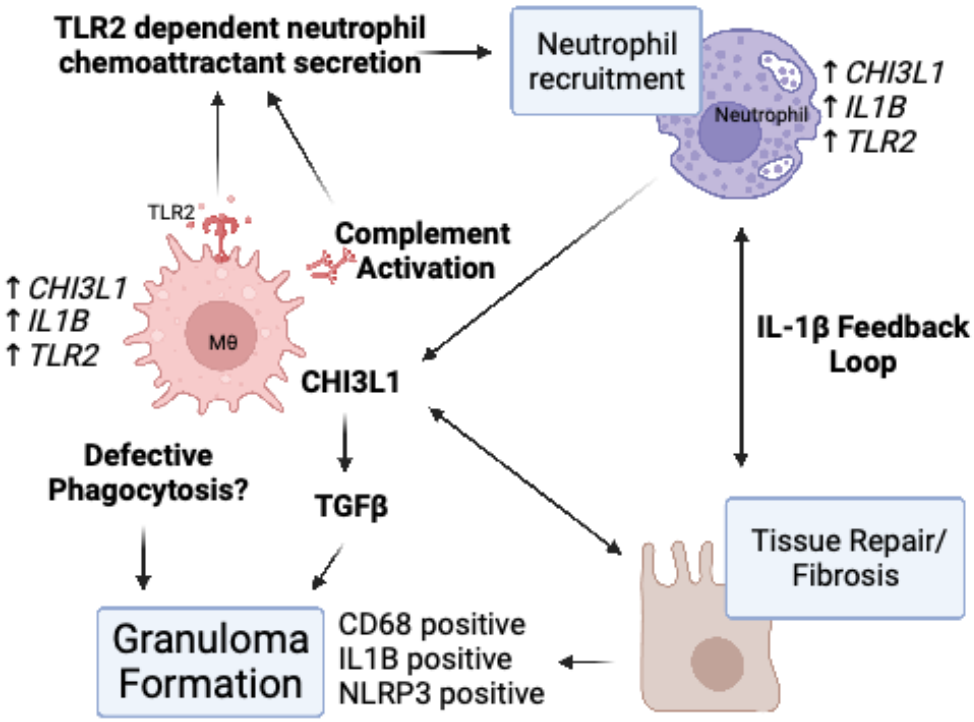


Figure 6.2: ***Overview of hypothesised mechanism in intestinal signature in non-responders to anti-TNF treatment identified in*** **chapter 5.**

CD patients with non-response to anti-TNF exhibited a complex intestinal immunopathotype. This hypothesis was generated on the genes with the highest predictive power for non-response to anti-TNF. TLR2 hyporeactive macrophages (Mθ) release chemoattractants, such as IL-8 and CXCL1, recruiting neutrophils to the intestine. Both macrophages and neutrophils release IL1-β and CHI3L1 to induce tissue repair and granuloma formation. The latter is further promoted by macrophages' defective phagocytosis of translocated bacteria in the mucosa and aberrant fibrosis. Epithelial cells and neutrophils stimulate each other via an IL1-β feedback loop. This figure was created using Blorender.com.

Chapter 7: Conclusion

My project has identified two distinct CD immunopathotypes by stratifying patients based on their PB T cell and intestinal immune signature. The PB immunopathotype was characterised by reduced levels of T cells, which was linked to an increase in $CD8^+$ T cells and cytotoxicity in paediatric patients and potentially more severe CD. IL1-ß, TLR activity, complement activity, and granuloma formation characterised the distinct intestinal phenotype. Interestingly, it appeared enriched in patients with non-response to anti-TNF before treatment initiation. Biomarkers for both disease severity and treatment responses to therapeutics are currently lacking in the management of CD. Further research into the mechanisms behind the immunopathotypes discussed in this thesis may therefore aid clinical management of CD and reduce the need for surgical interventions.

References

Aandahl, E.M., Sandberg, J.K., Beckerman, K.P., Taskén, K., Moretto, W.J. and Nixon, D.F. (2003) 'CD7 Is a Differentiation Marker That Identifies Multiple CD8 T Cell Effector Subsets', The Journal of Immunology, 170(5):2349–2355.

Abbas, A.R., Baldwin, D., Ma, Y., Ouyang, W., Gurney, A., Martin, F., Fong, S., van Lookeren Campagne, M., Godowski, P., Williams, P.M., Chan, A.C. and Clark, H.F. (2005) 'Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data', Genes & Immunity, 6(4):319–331.

Acosta-Rodriguez, E. V., Napolitani, G., Lanzavecchia, A. and Sallusto, F. (2007) 'Interleukins 1ß and 6 but not transforming growth factor-ß are essential for the differentiation of interleukin 17–producing human T helper cells', Nature Immunology, 8(9):942–949.

Acuto, O. and Michel, F. (2003) 'CD28-mediated co-stimulation: a quantitative support for TCR signalling', Nature Reviews Immunology, 3(12):939–951.

Aggarwal, S., Ghilardi, N., Xie, M.H., De Sauvage, F.J. and Gurney, A.L. (2003) 'Interleukin-23 promotes a distinct CD4 T cell activation state characterized by the production of interleukin-17', The Journal of biological chemistry, 278(3):1910–1914.

Ahern, P.P., Izcue, A., Maloy, K.J. and Powrie, F. (2008) 'The interleukin-23 axis in intestinal inflammation', Immunological Reviews, 226(1):147–159.

Albina, J., Mills, C., Henry Jr, W. and Caldwell, M.D. (1990) 'Temporal expression of different pathways of 1-arginine metabolism in healing wounds. | The Journal of Immunology', Journal of Immunology, 144(10):3877–3880.

Amanzada, A., Ahmed Malik, I., Blaschke, M., Khan, S., Rahman, H., Ramadori, G. and Moriconi, F. (2013) 'Identification of CD68+ neutrophil granulocytes in in vitro model of acute inflammation and inflammatory bowel disease', International Journal of Clinical and Experimental Pathology, 6(4):561.

Ananthakrishnan, A.N., Khalili, H., Higuchi, L.M., Bao, Y., Korzenik, J.R., Giovannucci, E.L., Richter, J.M., Fuchs, C.S. and Chan, A.T. (2012) 'Higher predicted vitamin D status is associated with reduced risk of Crohn's disease', Gastroenterology, 142(3):482–489.

Ananthakrishnan, A.N., McGinley, E.L. and Binion, D.G. (2008) 'Excess hospitalisation burden associated with Clostridium difficile in patients with inflammatory bowel disease', Gut, 57(2):205–210.

Anderson, G., Jenkinson, E.J., Moore, N.C. and Owen, J.J.T. (1993) 'MHC class II-positive epithelium and mesenchyme cells are both required for T-cell development in the thymus', Nature 1993 362:6415, 362(6415):70–73.

Aniwan, S., Tremaine, W.J., Raffals, L.E., Kane, S. V. and Loftus, E. V. (2018) 'Antibiotic Use and New-Onset Inflammatory Bowel Disease in Olmsted County, Minnesota: A Population-Based Case-Control Study', Journal of Crohn's & colitis, 12(2):137–144.

Araki, A., Kanai, T., Ishikura, T., Makita, S., Uraushihara, K., Iiyama, R., Totsuka, T., Takeda, K., Akira, S. and Watanabe, M. (2005) 'MyD88-deficient mice develop severe intestinal inflammation in dextran sodium sulfate colitis', Journal of gastroenterology, 40(1):16–23.

Arijs, I., Quintens, R., Lommel, L. Van, Van Steen, K., De Hertogh, G., Lemaire, K., Schraenen, A., Perrier, C., Van Assche, G., Vermeire, S., Geboes, K., Schuit, F. and Rutgeerts, P. (2010) 'Predictive value of epithelial gene expression profiles for response to infliximab in Crohn's disease', Inflammatory bowel diseases, 16(12):2090–2098.

Artis, D., Humphreys, N.E., Potten, C.S., Wagner, N., Müller, W., Mcdermott, J.R., Grencis, R.K. and Else, K.J. (2000) 'Beta7 integrin-deficient mice: delayed leukocyte recruitment and attenuated protective immunity in the small intestine during enteric helminth infection', European journal of immunology, 30:1656–1664.

Atreya, R., Neumann, H., Neufert, C., Waldner, M.J., Billmeier, U., Zopf, Y.,

Willma, M., App, C., Münster, T., Kessler, H., Maas, S., Gebhardt, B., Heimke-Brinck, R., Reuter, E., Dörje, F., Rau, T.T., Uter, W., Wang, T.D., Kiesslich, R., Vieth, M., Hannappel, E. and Neurath, M.F. (2014) 'In vivo imaging using fluorescent antibodies to tumor necrosis factor predicts therapeutic response in Crohn's disease.', Nature medicine, 20(3):313–318.

Auer, I., Goetz, S., Ziemer, E., Malchow, H. and Ehms, H. (1979) 'Immune status in Crohn's disease', Gut, 20(4):261–268.

Bader, L., Gullaksen, S.E., Blaser, N., Brun, M., Bringeland, G.H., Sulen, A., Gjesdal, C.G., Vedeler, C. and Gavasso, S. (2019) 'Candidate markers for stratification and classification in rheumatoid arthritis', Frontiers in Immunology, 10:1488.

Bagnasco, D., Caminati, M., Ferrando, M., Aloè, T., Testino, E., Canonica, G.W. and Passalacqua, G. (2018) 'Anti-IL-5 and IL-5Ra: Efficacy and safety of new therapeutic strategies in severe uncontrolled asthma', BioMed Research International, 2018(5698212):1-8.

Bain, C.C., Bravo-Blas, A., Scott, C.L., Gomez Perdiguero, E., Geissmann, F., Henri, S., Malissen, B., Osborne, L.C., Artis, D. and Mowat, A.M.I. (2014) 'Constant replenishment from circulating monocytes maintains the macrophage pool in the intestine of adult mice', Nature Immunology, 15(10):929–937.

Bain, C.C. and Schridde, A. (2018) 'Origin, Differentiation, and Function of Intestinal Macrophages', Frontiers in Immunology, 9:2733.

Bain, C.C., Scott, C.L., Uronen-Hansson, H., Gudjonsson, S., Jansson, O., Grip, O., Guilliams, M., Malissen, B., Agace, W.W. and Mowat, A.M.I. (2013) 'Resident and pro-inflammatory macrophages in the colon represent alternative context-dependent fates of the same Ly6Chi monocyte precursors', Mucosal immunology, 6(3):498–510.

Banerjee, A., Gordon, S.M., Intlekofer, A.M., Paley, M.A., Mooney, E.C., Lindsten, T., Wherry, E.J. and Reiner, S.L. (2010) 'The Transcription Factor Eomesodermin Enables CD8+ T Cells to Compete for the Memory Cell Niche',

Journal of immunology, 185:4988–4992.

Bank, S., Skytt Andersen, P., Burisch, J., Pedersen, N., Roug, S., Galsgaard, J., Ydegaard Turino, S., Broder Brodersen, J., Rashid, S., Kaiser Rasmussen, B., Avlund, S., Bastholm Olesen, T., Jurgen Hoffmann, H., Kragh Thomsen, M., Thomsen, V., Frydenberg, M., Andersen Nexø, B., Sode, J., Vogel, U. and Andersen, V. (2014) 'Polymorphisms in the Inflammatory Pathway Genes TLR2, TLR4, TLR9, LY96, NFKBIA, NFKB1, TNFA, TNFRSF1A, IL6R, IL10, IL23R, PTPN22, and PPARG Are Associated with Susceptibility of Inflammatory Bowel Disease in a Danish Cohort', PLoS ONE, 9(6):e98815.

Bankhead, P., Loughrey, M.B., Fernández, J.A., Dombrowski, Y., McArt, D.G., Dunne, P.D., McQuaid, S., Gray, R.T., Murray, L.J., Coleman, H.G., James, J.A., Salto-Tellez, M. and Hamilton, P.W. (2017) 'QuPath: Open source software for digital pathology image analysis', Scientific Reports 2017, 7(1):1–7.

Barnes, P.J. (2006) 'How corticosteroids control inflammation: Quintiles Prize Lecture 2005', British Journal of Pharmacology, 148(3):245–254.

Barrett, J.C., Hansoul, S., Nicolae, D.L., Cho, J.H., Duerr, R.H., Rioux, J.D., Brant, S.R., Silverberg, M.S., Taylor, K.D., Michael, M., Bitton, A., Dassopoulos, T., Datta, L.W., Green, T., Griffiths, M., Kistner, E.O., Murtha, M.T., Regueiro, M.D., Jerome, I., Schumm, L.P., Steinhart, A.H., Targan, S.R., Ramnik, J., Prescott, J., Onnie, C.M., Fisher, S.A., Marchini, J. and Ghori, J. (2009) 'Genome-wide association defines more than thirty distinct susceptibility loci for Crohn's disease', Nature Genetics, 40(8):955–962.

Bauer, C., Duewell, P., Mayer, C., Lehr, H.A., Fitzgerald, K.A., Dauer, M., Tschopp, J., Endres, S., Latz, E. and Schnurr, M. (2010) 'Colitis induced in mice with dextran sulfate sodium (DSS) is mediated by the NLRP3 inflammasome', Gut, 59(9):1192–1199.

Bek, S., Nielsen, J. V, Bojesen, A.B., Franke, A., Bank, S., Vogel, U. and Andersen, V. (2016) 'Systematic review: genetic biomarkers associated with anti-TNF treatment response in inflammatory bowel diseases.', Alimentary pharmacology & therapeutics, 44(6):554–67.

Bennett, C.L., Christie, J., Ramsdell, F., Brunkow, M.E., Ferguson, P.J., Whitesell, L., Kelly, T.E., Saulsbury, F.T., Chance, P.F. and Ochs, H.D. (2001) 'The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3', Nature Genetics, 27(1):20–21.

Berg, A.M., Dam, A.N. and Farraye, F.A. (2013) 'Environmental Influences on the Onset and Clinical Course of Crohn's Disease—Part 2: Infections and Medication Use', Gastroenterology & Hepatology, 9(12):803-10.

Bernardo, D., Marin, A.C., Fernández-Tomé, S., Montalban-Arques, A., Carrasco, A., Tristán, E., Ortega-Moreno, L., Mora-Gutiérrez, I., Díaz-Guerra, A., Caminero-Fernández, R., Miranda, P., Casals, F., Caldas, M., Jiménez, M., Casabona, S., De La Morena, F., Esteve, M., Santander, C., Chaparro, M. and Gisbert, J.P. (2018) 'Human intestinal pro-inflammatory CD11c high CCR2 + CX3CR1 + macrophages, but not their tolerogenic CD11c - CCR2 - CX3CR1 - counterparts, are expanded in inflammatory bowel disease', Mucosal immunology, 11(4):1114–1126.

Bernstein, C.N., Blanchard, J.F., Rawsthorne, P. and Yu, N. (2001) 'The prevalence of extraintestinal diseases in inflammatory bowel disease: a population-based study', The American Journal of Gastroenterology, 96(4):1116–1122.

Best, W.R. (2006) 'Predicting the Crohn's disease activity index from the harvey-bradshaw index', Inflammatory Bowel Diseases, 12(4):304–310.

Best, W.R., Becktel, J.M., Singleton, J.W. and Kern, F. (1976) 'Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study.', Gastroenterology, 70(3):439–444.

de Bie, C., Kindermann, A. and Escher, J. (2013) 'Use of exclusive enteral nutrition in paediatric Crohn's disease in The Netherlands', Journal of Crohn's & colitis, 7(4):263–270.

Bielinski, S.J., Hall, J.L., Pankow, J.S., Boerwinkle, E., Matijevic-Aleksic, N., He, M., Chambless, L. and Folsom, A.R. (2011) 'Genetic variants in TLR2 and

TLR4 are associated with markers of monocyte activation: The atherosclerosis risk in communities MRI study', Human Genetics, 129(6):655–662.

Birchenough, G.M.H., Johansson, M.E. V, Gustafsson, J.K., Bergström, J.H. and Hansson, G.C. (2015) 'New developments in goblet cell mucus secretion and function', Mucosal Immunology, 8:712–719.

Bjarnason, I. (2017) 'The Use of Fecal Calprotectin in Inflammatory Bowel Disease', Gastroenterology & Hepatology, 13(1):53-56.

Blackburn, S.D., Shin, H., Haining, W.N., Zou, T., Workman, C.J., Polley, A., Betts, M.R., Freeman, G.J., Vignali, D.A.A. and Wherry, E.J. (2008) 'Coregulation of CD8+ T cell exhaustion by multiple inhibitory receptors during chronic viral infection', Nature Immunology, 10(1):29–37.

Bohner, P., Chevalier, M.F., Cesson, V., Rodrigues-Dias, S.C., Dartiguenave, F., Burruni, R., Tawadros, T., Valerio, M., Lucca, I., Nardelli-Haefliger, D., Jichlinski, P. and Derré, L. (2019) 'Double positive CD4+CD8+ T cells are enriched in urological cancers and favor T helper-2 polarization', Frontiers in Immunology, 10:622.

Bohr, S., Patel, S.J., Vasko, R., Shen, K., Golberg, A., Berthiaume, F. and Yarmush, M.L. (2015) 'The role of CHI3L1 (Chitinase-3-Like-1) in the pathogenesis of infections in burns in a mouse model', PLoS ONE, 10(11):e0140440

Booth, C., O'Shea, J.A. and Potten, C.S. (1999) 'Maintenance of Functional Stem Cells in Isolated and Cultured Adult Intestinal Epithelium', Experimental Cell Research, 249(2):359–366.

Boschetti, G., Nancey, S., Moussata, D., Cotte, E., Francois, Y., Flourié, B. and Kaiserlian, D. (2016) 'Enrichment of Circulating and Mucosal Cytotoxic CD8+ T Cells Is Associated with Postoperative Endoscopic Recurrence in Patients with Crohn's Disease', Journal of Crohn's & colitis, 10(3):338–345.

Bösl, K., Giambelluca, M., Haug, M., Bugge, M., Espevik, T., Kandasamy, R.K. and Bergstrøm, B. (2018) 'Coactivation of TLR2 and TLR8 in primary human

monocytes triggers a distinct inflammatory signaling response', Frontiers in Physiology, 9:618.

Bouazzi, M., Bak, N.F., Agnholt, J., Wewer, V., Malham, M. and Bendix, M. (2022) 'Paediatric Crohn's Disease Patients Have Increased Inflammatory Markers Compared to Adult Patients prior to Biological Treatment', GastroHep, 2022(9550908):1–11.

Buisson, A., Vazeille, E., Minet-Quinard, R., Goutte, M., Bouvier, D., Goutorbe, F., Pereira, B., Barnich, N. and Bommelaer, G. (2016) 'Faecal chitinase 3-like 1 is a reliable marker as accurate as faecal calprotectin in detecting endoscopic activity in adult patients with inflammatory bowel diseases', Alimentary Pharmacology & Therapeutics, 43(10):1069–1079.

Breese, E., Braegger, C.P., Corrigan, C.J., Walker-Smith, J.A. and MacDonald, T.T. (1993) 'Interleukin-2- and interferon-gamma-secreting T cells in normal and diseased human intestinal mucosa.', Immunology, 78(1):127.

Bressenot, A., Salleron, J., Bastien, C., Danese, S., Boulagnon-Rombi, C. and Peyrin-Biroulet, L. (2015) 'Comparing histological activity indexes in UC', Gut, 64(9):1412–1418.

Broberger, O. and Perlmann, P. (1959) 'AUTOANTIBODIES IN HUMAN ULCERATIVE COLITIS', Journal of Experimental Medicine, 110(5):657–674.

Brunkow, M.E., Jeffery, E.W., Hjerrild, K.A., Paeper, B., Clark, L.B., Yasayko, S.A., Wilkinson, J.E., Galas, D., Ziegler, S.F. and Ramsdell, F. (2001) 'Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse', Nature Genetics, 27(1):68–73.

Bujko, A., Atlasy, N., Landsverk, O.J.B., Richter, L., Yaqub, S., Horneland, R., Øyen, O., Aandahl, E.M., Aabakken, L., Stunnenberg, H.G., Bækkevold, E.S. and Jahnsen, F.L. (2018) 'Transcriptional and functional profiling defines human small intestinal macrophage subsets', The Journal of experimental medicine, 215(2):441–458.

Bunn, S.K., Bisset, W.M., Main, M.J.C. and Golden, B.E. (2001) 'Fecal calprotectin as a measure of disease activity in childhood inflammatory bowel disease', Journal of pediatric gastroenterology and nutrition, 32(2):171–177.

Caër, C., Gorreja, F., Forsskåhl, S.K., Brynjolfsson, S.F., Szeponik, L., Magnusson, M.K., Börjesson, L.G., Block, M., Bexe-Lindskog, E. and Wick, M.J. (2021) 'TREM-1+ Macrophages Define a Pathogenic Cell Subset in the Intestine of Crohn's Disease Patients', Journal of Crohn's & Colitis, 15(8):1346-1361.

De Calisto, J., Villablanca, E.J. and Mora, J.R. (2012) 'FcγRI (CD64): an identity card for intestinal macrophages', European journal of immunology, 42:3136-3140.

Cario, E., Gerken, G. and Podolsky, D.K. (2007) 'Toll-like receptor 2 controls mucosal inflammation by regulating epithelial barrier function', Gastroenterology, 132(4):1359–1374.

Cario, E. and Podolsky, D.K. (2000) 'Differential Alteration in Intestinal Epithelial Cell Expression of Toll-Like Receptor 3 (TLR3) and TLR4 in Inflammatory Bowel Disease', Infection and Immunity, 68(12):7010-7017.

Carr, M.W., Roth, S.J., Luther, E., Rose, S.S. and Springer, T.A. (1994) 'Monocyte chemoattractant protein 1 acts as a T-lymphocyte chemoattractant', Proceedings of the National Academy of Sciences of the United States of America, 91(9):3652–3656.

Chaudhry, A., Samstein, R.M., Treuting, P., Liang, Y., Pils, M.C., Heinrich, J.M., Jack, R.S., Wunderlich, F.T., Brüning, J.C., Müller, W. and Rudensky, A.Y. (2011) 'Interleukin-10 signaling in regulatory T cells is required for suppression of Th17 cell-mediated inflammation', Immunity, 34(4):566.

Cheifetz, A.S. (2013) 'Management of active Crohn disease', JAMA, 309(20): 2150–2158.

Chen, C.C., Pekow, J., Llado, V., Kanneganti, M., Lau, C.W., Mizoguchi, A., Mino-Kenudson, M., Bissonnette, M. and Mizoguchi, E. (2011) 'Chitinase 3-Like-1

Expression in Colonic Epithelial Cells as a Potentially Novel Marker for Colitis-Associated Neoplasia', The American Journal of Pathology, 179(3):1494-1503.

Chertov, O., Ueda, H., Xu, L.L., Tani, K., Murphy, W.J., Wang, J.M., Howard, O.M.Z., Sayers, T.J. and Oppenheim, J.J. (1997) 'Identification of human neutrophil-derived cathepsin G and azurocidin/CAP37 as chemoattractants for mononuclear cells and neutrophils', The Journal of experimental medicine, 186(5):739–747.

Choi, S.M., Park, H.J., Choi, E.A., Jung, K.C. and Lee, J. Il (2021) 'Cellular heterogeneity of circulating CD4+CD8+ double-positive T cells characterized by single-cell RNA sequencing', Scientific Reports, 11(1):1–12.

Chu, K.M., Watermeyer, G., Shelly, L., Janssen, J., May, T.D., Brink, K., Benefeld, G. and Li, X. (2013) 'Childhood helminth exposure is protective against inflammatory bowel disease: a case control study in South Africa', Inflammatory bowel diseases, 19(3):614–620.

Cleynen, I., Konings, P., Robberecht, C., Laukens, D., Amininejad, L., Theatre, E., MacHiels, K., Arijs, I., Rutgeerts, P., Louis, E., Franchimont, D., De Vos, M., Van Steen, K., Georges, M., Moreau, Y., Vermeesch, J. and Vermeire, S. (2016) 'Genome-Wide Copy Number Variation Scan Identifies Complement Component C4 as Novel Susceptibility Gene for Crohn's Disease', Inflammatory bowel diseases, 22(3):505–515.

Coffman, S. (1998) 'Cytokines Induce the Development Review of Functionally Heterogeneous T Helper Cell Subsets', Immunity, 8:275–283.

Cole, J.J., Faydaci, B.A., McGuinness, D., Shaw, R., Maciewicz, R.A., Robertson, N.A. and Goodyear, C.S. (2021) 'Searchlight: automated bulk RNA-seq exploration and visualisation using dynamically generated R scripts', BMC Bioinformatics, 22:411.

Cosnes, J., Beaugerie, L., Carbonnel, F. and Gendre, J.P. (2001) 'Smoking cessation and the course of Crohn's disease: An intervention study', Gastroenterology, 120(5):1093–1099.

Crohn's & Colitis UK (2022): New research shows over 1 in 123 people in UK living with Crohn's or Colitis. Available at: https://crohnsandcolitis.org.uk/news-stories/news-items/new-research-shows-over-1-in-123-people-in-uk-living-with-crohn-s-or-colitis (Accessed: 16 July 2022).

Crohn, B.B., Ginzburg, L. and Oppenheimer, G.D. (1932) 'REGIONAL ILEITIS: A PATHOLOGIC AND CLINICAL ENTITY', Journal of the American Medical Association, 99(16):1323–1329.

Cummings, R.J., Barbet, G., Bongers, G., Hartmann, B.M., Gettler, K., Muniz, L., Furtado, G.C., Cho, J., Lira, S.A. and Blander, J.M. (2016) 'Different tissue phagocytes sample apoptotic cells to direct distinct homeostasis programs', Nature, 539(7630):565–569.

D'Haens, G.R., Geboes, K., Peeters, M., Baert, F., Penninckx, F. and Rutgeerts, P. (1998) 'Early lesions of recurrent Crohn's disease caused by infusion of intestinal contents in excluded ileum', Gastroenterology, 114(2):262–267.

Dadi, H.K., Simon, A.J. and Roifman, C.M. (2003) 'Effect of CD3delta deficiency on maturation of alpha/beta and gamma/delta T-cell lineages in severe combined immunodeficiency', The New England journal of medicine, 349(19):1821–1828.

Dai, B., Hackney, J.A., Ichikawa, R., Nguyen, A., Elstrott, J., Orozco, L.D., Sun, K.H., Modrusan, Z., Gogineni, A., Scherl, A., Gubatan, J., Habtezion, A., Deswal, M., Somsouk, M., Faubion, W.A., Chai, A., Sharafali, Z., Hassanali, A., Oh, Y.S., Tole, S., McBride, J., Keir, M.E. and Yi, T. (2021) 'Dual targeting of lymphocyte homing and retention through α4ß7 and αEß7 inhibition in inflammatory bowel disease', Cell Reports Medicine, 2(8):100381.

Daperno, M., D'Haens, G., Van Assche, G., Baert, F., Bulois, P., Maunoury, V., Sostegni, R., Rocca, R., Pera, A., Gevers, A., Mary, J.Y., Colombel, J.F. and Rutgeerts, P. (2004) 'Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD', Gastrointestinal Endoscopy, 60(4):505–512.

Dejaco, C., Lichtenberger, C., Miehsler, W., Oberhuber, G., Herbst, F., Vogelsang, H., Gangl, A. and Reinisch, W. (2003) 'An open-label pilot study of granulocyte colony-stimulating factor for the treatment of severe endoscopic postoperative recurrence in Crohn's disease', Digestion, 68(2–3):63–70.

Denning, T.L., Wang, Y.C., Patel, S.R., Williams, I.R. and Pulendran, B. (2007) 'Lamina propria macrophages and dendritic cells differentially induce regulatory and interleukin 17–producing T cell responses', Nature Immunology, 8(10):1086–1094.

Deshmane, S.L., Kremlev, S., Amini, S. and Sawaya, B.E. (2009) 'Monocyte Chemoattractant Protein-1 (MCP-1): An Overview', Journal of Interferon & Cytokine Research, 29(6):313-326.

Deusch, K., Lüling, F., Reich, K., Classen, M., Wagner, H. and Pfeffer, K. (1991) 'A major fraction of human intraepithelial lymphocytes simultaneously expresses the gamma/delta T cell receptor, the CD8 accessory molecule and preferentially uses the V delta 1 gene segment', European journal of immunology, 21(4):1053–1059.

Deutschmann, C., Sowa, M., Murugaiyan, J., Roesler, U., Röber, N., Conrad, K., Laass, M.W., Bogdanos, D., Sipeki, N., Papp, M., Rödiger, S., Roggenbuck, D. and Schierack, P. (2019) 'Identification of Chitinase-3-Like Protein 1 as a Novel Neutrophil Antigenic Target in Crohn's Disease', Journal of Crohn's & Colitis, 13(7):894-904.

Djurić, Z., Šaranac, L., Budić, I., Pavlović, V. and Djordjević, J. (2018) 'Therapeutic role of methotrexate in pediatric Crohn's disease', Bosnian Journal of Basic Medical Sciences, 18(3):211-216.

Donskoy, E. and Goldschneider, I. (1992) 'Thymocytopoiesis is maintained by blood-borne precursors throughout postnatal life. A study in parabiotic mice', Journal of Immunology, 148(6):1604–12.

Doyle, A.G., Herbein, G., Montaner, L.J., Minty, A.J., Caput, D., Ferrara, P. and Gordon, S. (1994) 'Interleukin-13 alters the activation state of murine

macrophages in vitro: comparison with interleukin-4 and interferon-gamma', European journal of immunology, 24(6):1441–1445.

Duerr, R.H., Taylor, K.D., Brant, S.R., Rioux, J.D., Silverberg, M.S., Daly, M.J., Steinhart, A.H., Abraham, C., Regueiro, M., Griffiths, A., Dassopoulos, T., Bitton, A., Yang, H., Targan, S., Datta, L.W., Kistner, E.O., Schumm, L.P., Lee, A.T., Gregersen, P.K., Barmada, M.M., Rotter, J.I., Nicolae, D.L. and Cho, J.H. (2006) 'A genome-wide association study identifies IL23R as an inflammatory bowel disease gene', Science, 314(5804):1461–1463.

Duman, M., Gencpinar, P., Biçmen, M., Arslan, N., Özden, Ö., Üzüm, Ö., Çelik, D., Sayiner, A.A. and Gülay, Z. (2015) 'Fecal calprotectin: can be used to distinguish between bacterial and viral gastroenteritis in children?', The American journal of emergency medicine, 33(10):1436–1439.

Duricova, D., Burisch, J., Jess, T., Gower-Rousseau, C. and Lakatos, P.L. (2014) 'Age-related differences in presentation and course of inflammatory bowel disease: An update on the population-based literature', Journal of Crohn's and Colitis, 8(11):1351–1361.

Eljaafari, A., Yuruker, O., Ferrand, C., Farre, A., Addey, C., Tartelin, M.-L., Thomas, X., Tiberghien, P., Simpson, E., Rigal, D. and Scott, D. (2013) 'Isolation of human CD4/CD8 double-positive, graft-versus-host disease-protective, minor histocompatibility antigen-specific regulatory T cells and of a novel HLA-DR7-restricted HY-specific CD4 clone', Journal of immunology (Baltimore, Md. : 1950), 190(1):184–194.

Elvington, M., Schepp-Berglind, J. and Tomlinson, S. (2015) 'Regulation of the alternative pathway of complement modulates injury and immunity in a chronic model of dextran sulphate sodium-induced colitis', Clinical and experimental immunology, 179(3):500–508.

EMBL-EBI (2018) QuickGO::Term GO:0002376 immune system process. Available at: https://www.ebi.ac.uk/QuickGO/term/GO:0002376 (Accessed: 12 July 2022).

Esensten, J.H., Helou, Y.A., Chopra, G., Weiss, A. and Bluestone, J.A. (2016) 'CD28 costimulation: from mechanism to therapy', Immunity, 44(5):973-988.

Fairbanks, L.D., Rückemann, K., Qiu, Y., Hawrylowicz, C.M., Richards, D.F., Swaminathan, R., Kirschbaum, B. and Simmonds, H.A. (1999) 'Methotrexate inhibits the first committed step of purine biosynthesis in mitogen-stimulated human T-lymphocytes: a metabolic basis for efficacy in rheumatoid arthritis?', Biochemistry Journal, 342(Pt1):143–152.

Feakins, R.M. (2014) 'Ulcerative colitis or crohn's disease? Pitfalls and problems', *Histopathology*, 64(3):317–335.

Ferenbach, D. and Hughes, J. (2008) 'Macrophages and dendritic cells: what is the difference?', Kidney International, 74(1):5–7.

Ferenczi, K., Burack, L., Pope, M., Krueger, J.G. and Austin, L.M. (2000) 'CD69, HLA-DR and the IL-2R Identify Persistently Activated T Cells in Psoriasis Vulgaris Lesional Skin: Blood and Skin Comparisons by Flow Cytometry', Journal of Autoimmunity, 14(1):63–78.

Ferrante, M., Henckaerts, L., Joossens, M., Pierik, M., Joossens, S., Dotan, N., Norman, G.L., Altstock, R.T., Van Steen, K., Rutgeerts, P., Van Assche, G. and Vermeire, S. (2007) 'New serological markers in inflammatory bowel disease are associated with complicated disease behaviour.', Gut, 56(10):1394–403.

Feuerstein, J.D. and Cheifetz, A.S. (2017) 'Crohn Disease: Epidemiology, Diagnosis, and Management', Mayo Clinic Proceedings, 92(7):1088–1103.

Fontenot, J.D., Gavin, M.A. and Rudensky, A.Y. (2003) 'Foxp3 programs the development and function of CD4+CD25+ regulatory T cells', Nature Immunology, 4(4):330–336.

Franchimont, D., Vermeire, S., El Housni, H., Pierik, M., Van Steen, K., Gustot, T., Quertinmont, E., Abramowicz, M., Van Gossum, A., Devière, J. and Rutgeerts, P. (2004) 'Deficient host-bacteria interactions in inflammatory bowel disease? The toll-like receptor (TLR)-4 Asp299gly polymorphism is associated

with Crohn's disease and ulcerative colitis', Gut, 53(7):987–992.

Frau, A., Ijaz, U.Z., Slater, R., Jonkers, D., Penders, J., Campbell, B.J., Kenny, J.G., Hall, N., Lenzi, L., Burkitt, M.D., Pierik, M., Darby, A.C. and Probert, C.S.J. (2021) 'Inter-kingdom relationships in Crohn's disease explored using a multi-omics approach', Gut Microbes, 13(1).

Freeman, H.J. (2004) 'Comparison of longstanding pediatric-onset and adult-onset Crohn's disease', Journal of pediatric gastroenterology and nutrition, 39(2):183–186.

Freeman, K., Ryan, R., Parsons, N., Taylor-Phillips, S., Willis, B.H. and Clarke, A. (2021) 'The incidence and prevalence of inflammatory bowel disease in UK primary care: a retrospective cohort study of the IQVIA Medical Research Database', BMC Gastroenterology, 21(1):1–7.

Friedrich, M., Pohin, M., Jackson, M.A., Korsunsky, I., Bullers, S.J., Rue-Albrecht, K., Christoforidou, Z., Sathananthan, D., Thomas, T., Ravindran, R., Tandon, R., Peres, R.S., Sharpe, H., Wei, K., Watts, G.F.M., Mann, E.H., Geremia, A., Attar, M., Barone, F., Brenner, M., Buckley, C.D., Coles, M., Frei, A.P., Lassen, K.G., Powrie, F.M., McCuaig, S., Thomas, L., Collantes, E., Uhlig, H.H., Sansom, S.N., Easton, A., Raychaudhuri, S., Travis, S.P. and Powrie, F.M. (2021) 'IL-1-driven stromal–neutrophil interactions define a subset of patients with inflammatory bowel disease that does not respond to therapies', Nature Medicine, 27(11):1970–1981.

Frolkis, A.D., Dykeman, J., Negrón, M.E., Debruyn, J., Jette, N., Fiest, K.M., Frolkis, T., Barkema, H.W., Rioux, K.P., Panaccione, R., Ghosh, S., Wiebe, S. and Kaplan, G.G. (2013) 'Risk of Surgery for Inflammatory Bowel Diseases Has Decreased Over Time: A Systematic Review and Meta-analysis of Population-Based Studies', Gastroenterology, 145(5):996–1006.

Fu, Y., Zhan, X., Wang, Y., Jiang, X., Liu, M., Yang, Y., Huang, Y., Du, X., Zhong, X.-P., Li, L., Ma, L. and Hu, S. (2019) 'NLRC3 expression in dendritic cells attenuates CD4+ T cell response and autoimmunity', The EMBO Journal, 38(16):e101397.

Fujino, S., Andoh, A., Bamba, S., Ogawa, A., Hata, K., Araki, Y. and Bamba, T. (2003) 'Increased expression of interleukin 17 in inflammatory bowel disease', Gut, 52:65–70.

Fukata, M., Michelsen, K.S., Eri, R., Thomas, L.S., Hu, B., Lukasek, K., Nast, C.C., Lechago, J., Xu, R., Naiki, Y., Soliman, A., Arditi, M. and Abreu, M.T. (2005) 'Toll-like receptor-4 is required for intestinal response to epithelial injury and limiting bacterial translocation in a murine model of acute colitis', American Journal of Physiology - Gastrointestinal and Liver Physiology, 288(5 51-5):1055–1065.

Funderburg, N.T., Stubblefield Park, S.R., Sung, H.C., Hardy, G., Clagett, B., Ignatz-Hoover, J., Harding, C. V., Fu, P., Katz, J.A., Lederman, M.M. and Levine, A.D. (2013) 'Circulating CD4+ and CD8+ T cells are activated in inflammatory bowel disease and are associated with plasma markers of inflammation', Immunology, 140(1):87–97.

Fuss, I.J., Neurath, M., Boirivant, M., Klein, J.S., de la Motte, C., Strong, S.A., Fiocchi, C. and Strober, W. (1996) 'Disparate CD4+ lamina propria (LP) lymphokine secretion profiles in inflammatory bowel disease. Crohn's disease LP cells manifest increased secretion of IFN-gamma, whereas ulcerative colitis LP cells manifest increased secretion of IL-5.', The Journal of Immunology, 157(3):1261-1270.

Gaiani, F., Rotoli, B.M., Ferrari, F., Barilli, A., Visigalli, R., Carra, M.C., de'Angelis, G.L., de'Angelis, N. and Dall'Asta, V. (2020) 'Monocytes from infliximab-resistant patients with Crohn's disease exhibit a disordered cytokine profile', Scientific Reports, 10(1):1–10.

Galkina, E., Tanousis, K., Preece, G., Tolaini, M., Kioussis, D., Florey, O., Haskard, D.O., Tedder, T.F. and Ager, A. (2003) 'L-Selectin Shedding Does Not Regulate Constitutive T Cell Trafficking but Controls the Migration Pathways of Antigen-activated T Lymphocytes', The Journal of Experimental Medicine, 198(9):1323-1335.

García-Sánchez, V., Iglesias-Flores, E., González, R., Gisbert, J.P., Gallardo-

Valverde, J.M., González-Galilea, Á., Naranjo-Rodríguez, A., de Dios-Vega, J.F., Muntané, J. and Gómez-Camacho, F. (2010) 'Does fecal calprotectin predict relapse in patients with Crohn's disease and ulcerative colitis?', Journal of Crohn's & colitis, 4(2):144–152.

Gately, M.K., Warrier, R.R., Honasoge, S., Carvajal, D.M., Faherty, D.A., Connaughton, S.E., Anderson, T.D., Sarmiento, U., Hubbard, B.R. and Murphy, M. (1994) 'Administration of recombinant IL-12 to normal mice enhances cytolytic lymphocyte activity and induces production of IFN-γ in vivo', International Immunology, 6(1): 157–167.

Gauchotte-Lindsay, C., Aspray, T.J., Knapp, M. and Ijaz, U.Z. (2019) 'Systems biology approach to elucidation of contaminant biodegradation in complex samples – integration of high-resolution analytical and molecular tools', Faraday Discussions, 218(0):481–504.

Gaya, D.R., Lyon, T.D.B., Duncan, A., Neilly, J.B., Han, S., Howell, J., Liddell, C., Stanley, A.J., Morris, A.J. and Mackenzie, J.F. (2005) 'Faecal calprotectin in the assessment of Crohn's disease activity', QJM: An International Journal of Medicine, 98(6):435–441.

Gebert, A., Rothkötter, H.J. and Pabst, R. (1996) 'M Cells in Peyer's Patches of the Intestine', International Review of Cytology, 167:91–159.

Geisler, C., Larsen, J.K. and Plesner, T. (1988) 'Identification of αß and γδ T Cell Receptor-Positive Cells', Scandinavian Journal of Immunology, 28(6):741–745.

Geng, B., Pan, J., Zhao, T., Ji, J., Zhang, C., Che, Y., Yang, J., Shi, H., Li, J., Zhou, H., Mu, X., Xu, C., Wang, C., Xu, Y., Liu, Z., Wen, H. and You, Q. (2018) 'Chitinase 3-like 1-CD44 interaction promotes metastasis and epithelial-to-mesenchymal transition through ß-catenin/Erk/Akt signaling in gastric cancer', Journal of Experimental and Clinical Cancer Research, 37(1):1–20.

Gerloff, K., Pereira, D.I.A., Faria, N., Boots, A.W., Kolling, J., Förster, I., Albrecht, C., Powell, J.J. and Schins, R.P.F. (2013) 'Influence of simulated

gastro-intestinal conditions on particle-induced cytotoxicity and interleukin-8 regulation in differentiated and undifferentiated Caco-2 cells', Nanotoxicology, 7(4):53.

Girardin, S.E., Boneca, I.G., Viala, J., Chamaillard, M., Labigne, A., Thomas, G., Philpott, D.J. and Sansonetti, P.J. (2003) 'Nod2 Is a General Sensor of Peptidoglycan through Muramyl Dipeptide (MDP) Detection', Journal of Biological Chemistry, 278(11):8869–8872.

Gisbert, J. and Panés, J. (2009) 'Loss of Response and Requirement of Infliximab Dose Intensification in Crohn's Disease', American Journal of Gastroenterology, 104(3):760–767.

Glocker, E.-O., Kotlarz, D., Boztug, K., Gertz, E.M., Schäffer, A.A., Noyan, F., Perro, M., Diestelhorst, J., Allroth, A., Murugan, D., Hätscher, N., Pfeifer, D., Sykora, K.-W., Sauer, M., Kreipe, H., Lacher, M., Nustede, R., Woellner, C., Baumann, U., Salzer, U., Koletzko, S., Shah, N., Segal, A.W., Sauerbrey, A., Buderus, S., Snapper, S.B., Grimbacher, B. and Klein, C. (2009) 'Inflammatory bowel disease and mutations affecting the interleukin-10 receptor', The New England journal of medicine, 361(21): 2033–2045.

Graham, D.B. and Xavier, R.J. (2013) 'From genetics of inflammatory bowel disease towards mechanistic insights', Trends in immunology, 34(8):371–378.

Gratchev, A., Kzhyshkowska, J., Köthe, K., Muller-Molinet, I., Kannookadan, S., Utikal, J. and Goerdt, S. (2006) 'Mφ1 and Mφ2 can be re-polarized by Th2 or Th1 cytokines, respectively, and respond to exogenous danger signals', Immunobiology, 211(6–8):473–486.

Guan, Q. and Zhang, J. (2017) 'Recent Advances: The Imbalance of Cytokines in the Pathogenesis of Inflammatory Bowel Disease', Mediators of Inflammation, 2017:4810258.

Guidi, L., Marzo, M., Andrisani, G., Felice, C., Pugliese, D., Mocci, G., Nardone, O., De Vitis, I., Papa, A., Rapaccini, G., Forni, F. and Armuzzi, A. (2014) 'Faecal calprotectin assay after induction with anti-Tumour Necrosis Factor α agents in

inflammatory bowel disease: Prediction of clinical response and mucosal healing at one year', Digestive and Liver Disease, 46(11):974–979.

Guti-acerrez, A., Francés, R., Amorós, A., Zapater, P., Garmendia, M., NDongo, M., Caño, R., Jover, R., Such, J. and Pérez-Mateo, M. (2009) 'Cytokine Association with Bacterial DNA in Serum of Patients with Inflammatory Bowel Disease', Inflammatory Bowel Diseases, 15(4):508–514.

Gutzeit, C., Magri, G. and Cerutti, A. (2014) 'Intestinal IgA production and its role in host-microbe interaction', Immunological reviews, 260(1):76-85.

Ha, C. and Kornbluth, A. (2014) 'Vedolizumab as a Treatment for Crohn's Disease and Ulcerative Colitis.', Gastroenterology & hepatology, 10(12):793–800.

Hamann, D., Baars, P.A., Rep, M.H.G., Hooibrink, B., Kerkhof-Garde, S.R., Klein, M.R. and Van Lier, R.A.W. (1997) 'Phenotypic and Functional Separation of Memory and Effector Human CD8+ T Cells', Journal of Experimental Medicine, 186(9):1407–1418.

Hampe, J., Cuthbert, A., Croucher, P.J.P., Mirza, M.M., Mascheretti, S., Fisher, S., Frenzel, H., King, K., Hasselmeyer, A., MacPherson, A.J.S., Bridger, S., Van Deventer, S., Forbes, A., Nikolaus, S., Lennard-Jones, J.E., Foelsch, U.R., Krawczak, M., Lewis, C., Schreiber, S. and Mathew, C.G. (2001) 'Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations', The Lancet, 357(9272):1925–1928.

Han, B.K., Wysham, K.D., Cain, K.C., Tyden, H., Bengtsson, A.A. and Lood, C. (2020) 'Neutrophil and lymphocyte counts are associated with different immunopathological mechanisms in systemic lupus erythematosus', Lupus Science & Medicine, 7:e000382.

Hanauer, S.B., Feagan, B.G., Lichtenstein, G.R., Mayer, L.F., Schreiber, S., Colombel, J.F., Rachmilewitz, D., Wolf, D.C., Olson, A., Bao, W. and Rutgeerts, P. (2002) 'Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial', The Lancet, 359(9317):1541–1549.

Harvey, R.F. and Bradshaw, J.M. (1980) ‘A SIMPLE INDEX OF CROHN’S-DISEASE ACTIVITY’, The Lancet, 315(8167):514.

Hashimoto, D., Chow, A., Noizat, C., Teo, P., Beasley, M.B., Leboeuf, M., Becker, C.D., See, P., Price, J., Lucas, D., Greter, M., Mortha, A., Boyer, S.W., Forsberg, E.C., Tanaka, M., van Rooijen, N., García-Sastre, A., Stanley, E.R., Ginhoux, F., Frenette, P.S. and Merad, M. (2013) ‘Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes’, Immunity, 38(4):792–804.

Hausmann, M., Kiessling, S., Mestermann, S., Webb, G., Spöttl, T., Andus, T., Schölmerich, J., Herfarth, H., Ray, K., Falk, W. and Rogler, G. (2002) ‘Toll-like receptors 2 and 4 are up-regulated during intestinal inflammation’, Gastroenterology, 122(7):1987–2000.

He, C.H., Lee, C.G., DelaCruz, C.S., Lee, C.M., Zhou, Y., Ahangari, F., Ma, B., Herzog, E.L., Rosenberg, S.A., Li, Y., Nour, A.M., Parikh, C.R., Schmidt, I., Modis, Y., Cantley, L. and Elias, J.A. (2013) ‘Chitinase 3-like 1 regulates cellular and tissue responses via IL-13 receptor α2’, Cell Reports, 4(4):830–841.

He, X.C., Zhang, J., Tong, W.G., Tawfik, O., Ross, J., Scoville, D.H., Tian, Q., Zeng, X., He, X., Wiedemann, L.M., Mishina, Y. and Li, L. (2004) ‘BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt–β-catenin signaling’, Nature Genetics, 36(10):1117–1121.

Hedl, M., Li, J., Cho, J.H. and Abraham, C. (2007) ‘Chronic stimulation of Nod2 mediates tolerance to bacterial products’, Proceedings of the National Academy of Sciences of the United States of America, 104(49):19440–19445.

Heimann, T.M., Bolnick, K. and Aufses, A.H. (1986) ‘Prognostic significance of severe preoperative lymphopenia in patients with Crohn’s disease’, Annals of surgery, 203(2):132–135.

Hendriks, J., Gravestein, L.A., Tesselaar, K., Van Lier, R.A.W., Schumacher, T.N.M. and Borst, J. (2000) ‘CD27 is required for generation and long-term maintenance of T cell immunity’, Nature Immunology, 1(5):433–440.

Heresbach, D., Alexandre, J.L., Branger, B., Bretagne, J.F., Cruchant, E., Dabadie, A., Dartois-Hoguin, M., Girardot, P.M., Jouanolle, H., Kerneis, J., Le Verger, J.C., Louvain, V., Politis, J., Richecoeur, M., Robaszkiewiczs, M. and Seyrig, J.A. (2005) 'Frequency and significance of granulomas in a cohort of incident cases of Crohn's disease', *Gut*, 54(2):215.

Hewitt, E.W. (2003) 'The MHC class I antigen presentation pathway: strategies for viral immune evasion', Immunology, 110(2):163-169.

Hisamatsu, T., Suzuki, M., Reinecker, H.C., Nadeau, W.J., McCormick, B.A. and Podolsky, D.K. (2003) 'CARD15/NOD2 functions as an antibacterial factor in human intestinal epithelial cells', Gastroenterology, 124(4):993–1000.

Hoffmann, P., Eder, R., Boeld, T.J., Doser, K., Piseshka, B., Andreesen, R. and Edinger, M. (2006) 'Only the CD45RA+ subpopulation of CD4+CD25high T cells gives rise to homogeneous regulatory T-cell lines upon in vitro expansion', Blood, 108(13):4260–4267.

Homer, C.R., Richmond, A.L., Rebert, N.A., Achkar, J. and McDonald, C. (2010) 'ATG16L1 and NOD2 Interact in an Autophagy-Dependent Antibacterial Pathway Implicated in Crohn's Disease Pathogenesis', Gastroenterology, 139(5):1630-1641.e2.

Hommes, D.W. (2006) 'Step-Up Versus Top-Down Therapy in the Treatment of Crohn's Disease', Gastroenterology&Hepatology, 2(8):546–547.

Hornecker, B. and Veauthier, J. (2018) 'Crohn's Disease: Diagnosis and Management', American Family Physician, 98(11):661–669.

Horowitz, J.E., Warner, N., Staples, J., Crowley, E., Gosalia, N., Murchie, R., Van Hout, C., Fiedler, K., Welch, G., King, A.K., Reid, J.G., Overton, J.D., Baras, A., Shuldiner, A.R., Griffiths, A., Gottesman, O., Muise, A.M. and Gonzaga-Jauregui, C. (2021) 'Mutation spectrum of NOD2 reveals recessive inheritance as a main driver of Early Onset Crohn's Disease', Scientific Reports, 11(1):5595.

Hu, M.D. and Edelblum, K.L. (2017) ‘Sentinels at the frontline: the role of intraepithelial lymphocytes in inflammatory bowel disease’, Current pharmacology reports, 3(6):321.

Hu, S., Du, X., Huang, Y., Fu, Y., Yang, Y., Zhan, X., He, W., Wen, Q., Zhou, X., Zhou, C., Zhong, X.P., Yang, J., Xiong, W., Wang, R., Gao, Y. and Ma, L. (2018) ‘NLRC3 negatively regulates CD4+ T cells and impacts protective immunity during Mycobacterium tuberculosis infection’, PLOS Pathogens, 14(8):e1007266.

Hu, Z., Gu, W., Wei, Y., Liu, G., Wu, S. and Liu, T. (2019) ‘NKT Cells in Mice Originate from Cytoplasmic CD3-Positive, CD4−CD8− Double-Negative Thymocytes that Express CD44 and IL-7Rα’, Scientific Reports, 9(1):1–12.

Huan, Y., Kong, Q., Mou, H. and Yi, H. (2020) ‘Antimicrobial Peptides: Classification, Design, Application and Research Progress in Multiple Fields’, Frontiers in Microbiology, 11:582779.

Huang, H., Hao, S., Li, F., Ye, Z., Yang, J. and Xiang, J. (2007) ‘CD4+ Th1 cells promote CD8+ Tc1 cell survival, memory response, tumor localization and therapy by targeted delivery of interleukin 2 via acquired pMHC I complexes’, Immunology, 120(2):148.

Huang, W., Na, L., Fidel, P.L. and Schwarzenberger, P. (2004) ‘Requirement of interleukin-17A for systemic anti-Candida albicans host defense in mice’, The Journal of infectious diseases, 190(3):624–631.

Huang, Z., Zak, J., Pratumchai, I., Shaabani, N., Vartabedian, V.F., Nguyen, N., Wu, T., Xiao, C. and Teijaro, J.R. (2019) ‘IL-27 promotes the expansion of self-renewing CD8+ T cells in persistent viral infection’, Journal of Experimental Medicine, 216(8):1791–1808.

Huard, B., Gaulard, P., Faure, F., Hercend, T. and Triebel, F. (1994) ‘Cellular expression and tissue distribution of the human LAG-3-encoded protein, an MHC class II ligand’, Immunogenetics, 39(3):213–217.

Hueber, W., Charles Barish, C., Research Associates, W., Burakoff, R., Fedorak,

R., Karczewski, M., Karczewski, Jacek, Kaser, A., Mayer, L., Wehkamp, Jan, Gastroenterologie, R.-B.-K., Yao, Michael, Feagan, B., R Higgins, P.D., Karczewski, J, Reinisch, W., Sands, B., Travis, S., Wehkamp, J and Yao, M (2012) 'Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial', Gut, 61(12):1693–1700.

Huppertz, C., Jäger, B., Wieczorek, G., Engelhard, P., Oliver, S.J., Bauernfeind, F.G., Littlewood-Evans, A., Welte, T., Hornung, V. and Prasse, A. (2020) 'The NLRP3 inflammasome pathway is activated in sarcoidosis and involved in granuloma formation', European Respiratory Journal, 55:1900119.

Hurlimann, J., Thorbecke, G.J. and Hochwald, G.M. (1966) 'The liver as the site of C-reactive protein formation', The Journal of experimental medicine, 123(2):365–378.

Hyun, K.K., De La Luz Sierra, M., Williams, C.K., Gulino, A.V. and Tosato, G. (2006) 'G-CSF down-regulation of CXCR4 expression identified as a mechanism for mobilization of myeloid cells', Blood, 108(3):812.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U. and Speed, T.P. (2003) 'Exploration, normalization, and summaries of high density oligonucleotide array probe level data', Biostatistics (Oxford, England), 4(2):249–264.

Itoh, M., Takahashi, T., Sakaguchi, N., Kuniyasu, Y., Shimizu, J., Otsuka, F. and Sakaguchi, S. (1999) 'Thymus and Autoimmunity: Production of CD25+CD4+ Naturally Anergic and Suppressive T Cells as a Key Function of the Thymus in Maintaining Immunologic Self-Tolerance | The Journal of Immunology', The Journal of Immunology, 162(9):5317–5326.

Ivanova, E.A. and Orekhov, A.N. (2015) 'T Helper Lymphocyte Subsets and Plasticity in Autoimmunity and Cancer: An Overview', BioMed Research International, 2015:327470.

Jackson, A.L., Matsumoto, H., Janszen, M., Maino, V., Blidy, A. and Shye, S.

(1990) ‘Restricted expression of p55 interleukin 2 receptor (CD25) on normal T cells’, Clinical Immunology and Immunopathology, 54(1):126–133.

Janeway, C.A. (1992) ‘The T cell receptor as a multicomponent signalling machine: CD4/CD8 coreceptors and CD45 in T cell activation’, Annual review of immunology, 10:645–674.

Jang, H.J., Kang, B. and Choe, B.H. (2019) ‘The difference in extraintestinal manifestations of inflammatory bowel disease for children and adults’, Translational Pediatrics, 8(1):4-15.

Johnson, W.E., Li, C. and Rabinovic, A. (2007) ‘Adjusting batch effects in microarray expression data using empirical Bayes methods’, Biostatistics (Oxford, England), 8(1):118–127.

Johswich, K., Martin, M., Bleich, A., Kracht, M., Dittrich-Breiholz, O., Gessner, J.E., Suerbaum, S., Wende, E., Rheinheimer, C. and Klos, A. (2009) ‘Role of the C5a receptor (C5aR) in acute and chronic dextran sulfate-induced models of inflammatory bowel disease’, Inflammatory bowel diseases, 15(12):1812–1823.

Jonuleit, H., Schmitt, E., Stassen, M., Tuettenberg, A., Knop, J. and Enk, A.H. (2001) ‘Identification and Functional Characterization of Human Cd4+Cd25+ T Cells with Regulatory Properties Isolated from Peripheral Blood’, Journal of Experimental Medicine, 193(11):1285–1294.

Joosse, M.E., Menckeberg, C.L., De Ruiter, L.F., Rolien, H.(, Raatgeep,) C, Van Berkel, L.A., Simons-Oosterhuis, Y., Tindemans, I., Femke, A.(, Muskens,) M, Hendriks, R.W., Hoogenboezem, R.M., Cupedo, T., De Ridder, L., Escher, J.C., Veenbergen, S. and Samsom, J.N. (2018) ‘Frequencies of circulating regulatory TIGIT + CD38 + effector T cells correlate with the course of inflammatory bowel disease’, Mucosal Immunology, 12:154-163.

Jostins, L., Ripke, S., Weersma, R.K., Duerr, R.H., McGovern, D.P., Hui, K.Y., Lee, J.C., Schumm, L.P., et al. (2012) ‘Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease’, Nature, 491(7422):119–124.

Kamada, N., Hisamatsu, T., Okamoto, S., Chinen, H., Kobayashi, T., Sato, T., Sakuraba, A., Kitazume, M.T., Sugita, A., Koganei, K., Akagawa, K.S. and Hibi, T. (2008) 'Unique CD14+ intestinal macrophages contribute to the pathogenesis of Crohn disease via IL-23/IFN-γ axis', The Journal of Clinical Investigation, 118(6):2269–2280.

Kaplan, G.G. and Ng, S.C. (2017) 'Understanding and Preventing the Global Increase of Inflammatory Bowel Disease', Gastroenterology, 152(2):313-321.e2.

Kasahara, S., Ando, K., Saito, K., Sekikawa, K., Ito, H., Ishikawa, T., Ohnishi, H., Seishima, M., Kakumu, S. and Moriwaki, H. (2003) 'Lack of Tumor Necrosis Factor Alpha Induces Impaired Proliferation of Hepatitis B Virus-Specific Cytotoxic T Lymphocytes', Journal of Virology, 77(4):2469–2476.

Kashiwagi, I., Morita, R., Schichita, T., Komai, K., Saeki, K., Matsumoto, M., Takeda, K., Nomura, M., Hayashi, A., Kanai, T. and Yoshimura, A. (2015) 'Smad2 and Smad3 Inversely Regulate TGF-ß Autoinduction in Clostridium butyricum-Activated Dendritic Cells', Immunity, 43(1):65–79.

Kawada, M., Chen, C.C., Arihiro, A., Nagatani, K., Watanabe, T. and Mizoguchi, E. (2008) 'Chitinase 3-like-1 enhances bacterial adhesion to colonic epithelial cells through the interaction with bacterial chitin-binding protein', Laboratory Investigation 2008 88:8, 88(8):883–895.

Kawada, M., Seno, H., Kanda, K., Nakanishi, Y., Akitake, R., Komekado, H., Kawada, K., Sakai, Y., Mizoguchi, E. and Chiba, T. (2011) 'Chitinase 3-like 1 promotes macrophage recruitment and angiogenesis in colorectal cancer', Oncogene, 31(26):3111–3123.

Keating, C., Bolton-Warberg, M., Hinchcliffe, J., Davies, R., Whelan, S., L Wan, A.H., Fitzgerald, R.D., Davies, S.J., Ijaz, U.Z. and Smith, C.J. (2021) 'Temporal changes in the gut microbiota in farmed Atlantic cod (Gadus morhua) outweigh the response to diet supplementation with macroalgae', Animal Microbiome, 3:7.

Khoshnevisan, R., Vakili, F., Klein, C., Kotlarz, D., Nasirian, M., Sherkat, R. and

Rezaei, A. (2019) ‘An analysis and survey of interleukin-10 receptor mutation in inflammatory bowel disease (IBD) in the first Iranian IBD cohort’, Journal of Laboratory Medicine, 43(4):185–189.

Kleiveland, C. (2015) ‘Peripheral Blood Mononuclear Cells’, in The Impact of Food Bioactives on Health: In Vitro and Ex Vivo Models. Springer:161–167.

Knall, C., Worthen, G.S. and Johnson, G.L. (1997) ‘Interleukin 8-stimulated phosphatidylinositol-3-kinase activity regulates the migration of human neutrophils independent of extracellular signal-regulated kinase and p38 mitogen-activated protein kinases’, Proceedings of the National Academy of Sciences of the United States of America, 94(7):052–3057.

Kolaczkowska, E. and Kubes, P. (2013) ‘Neutrophil recruitment and function in health and inflammation’, Nature Reviews Immunology, 13(3):159–175.

Korzenik, J.R. and Dieckgraefe, B.K. (2005) ‘An open-labelled study of granulocyte colony-stimulating factor in the treatment of active Crohn’s disease’, Alimentary pharmacology & therapeutics, 21(4):391–400.

Kosoy, R., Kim-Schulze, S., Rahman, A., Friedman, J.R., Huang, R., Peters, L.A., Amir, E. ad, Perrigoue, J., Stojmirovic, A., Song, W. min, Ke, H., Ungaro, R., Mehandru, S., Cho, J., Dubinsky, M., Curran, M., Brodmerkel, C., Schadt, E.E., Sands, B.E., Colombel, J.F., Kasarskis, A., Argmann, C.A. and Suárez-Fariñas, M. (2021) ‘Deep Analysis of the Peripheral Immune System in IBD Reveals New Insight in Disease Subtyping and Response to Monotherapy or Combination Therapy’, CMGH, 12(2):599–632.

Kotanides, H. and Reich, N.C. (1996) ‘Interleukin-4-induced STAT6 recognizes and activates a target site in the promoter of the interleukin-4 receptor gene’, The Journal of biological chemistry, 271(41):25555–25561.

Krause, P., Morris, V., Greenbaum, J.A., Park, Y., Bjoerheden, U., Mikulski, Z., Muffley, T., Shui, J.W., Kim, G., Cheroutre, H., Liu, Y.C., Peters, B., Kronenberg, M. and Murai, M. (2015) ‘IL-10-producing intestinal macrophages prevent excessive antibacterial innate immunity by limiting IL-23 synthesis’,

Nature communications, 6:7055.

Kristensen, V.A., Cvancarova, M., Høivik, M.L., Moum, B. and Vatn, M.H. (2020) 'Serological antibodies and surgery in a population-based inception cohort of Crohn's disease patients – the IBSEN study', Scandinavian journal of gastroenterology, 55(4):436–441.

Kronman, M.P., Zaoutis, T.E., Haynes, K., Feng, R. and Coffin, S.E. (2012) 'Antibiotic Exposure and IBD Development Among Children: A Population-Based Cohort Study', Pediatrics, 130(4):e794-e803.

Kucuksezer, U.C., Aktas Cetin, E., Esen, F., Tahrali, I., Akdeniz, N., Gelmez, M.Y. and Deniz, G. (2021) 'The Role of Natural Killer Cells in Autoimmune Diseases', Frontiers in Immunology, 12:622306.

Kuhbandner, K., Hammer, A., Haase, S., Terbrack, E., Hoffmann, A., Schippers, A., Wagner, N., Hussain, R.Z., Miller-Little, W.A., Koh, A.Y., Stoolman, J.S., Segal, B.M., Linker, R.A. and Stüve, O. (2019) 'MaDCAM-1-mediated intestinal lymphocyte homing is critical for the development of active experimental autoimmune encephalomyelitis', Frontiers in Immunology, 10(APR), p. 903.

Kumar, B. V., Connors, T.J. and Farber, D.L. (2018) 'Human T cell development, localization, and function throughout life', Immunity, 48(2):202-213.

Kurachi, M., Barnitz, R.A., Yosef, N., Odorizzi, P.M., Diiorio, M.A., Lemieux, M.E., Yates, K., Godec, J., Klatt, M.G., Regev, A., Wherry, E.J. and Haining, W.N. (2014) 'The transcription factor BATF operates as an essential differentiation checkpoint in early effector CD8+ T cells', Nature Immunology 2014 15:4, 15(4):373–383.

Kurowski, J.A., Milinovich, A., Ji, X., Bauman, J., Sugano, D., Kattan, M.W. and Achkar, J.P. (2021) 'Differences in Biologic Utilization and Surgery Rates in Pediatric and Adult Crohn's Disease: Results From a Large Electronic Medical Record-derived Cohort', Inflammatory Bowel Diseases, 27(7):1035–1044.

Kusnadi, A., Ramírez-Suástegui, C., Fajardo, V., Chee, S.J., Meckiff, B.J.,

Simon, H., Pelosi, E., Seumois, G., Ay, F., Vijayanand, P. and Ottensmeier, C.H. (2021) 'Severely ill COVID-19 patients display impaired exhaustion features in SARS-CoV-2-reactive CD8+ T cells', Science Immunology, 6(55):eabe4782.

Kvedaraite, E., Lourda, M., Ideström, M., Chen, P., Olsson-Akefeldt, S., Forkel, M., Gavhed, D., Lindforss, U., Mjösberg, J., Henter, J.I. and Svensson, M. (2016) 'Tissue-infiltrating neutrophils represent the main source of IL-23 in the colon of patients with IBD', Gut, 65(10):1632–1641.

Lamb, C.A., Kennedy, N.A., Raine, T., Hendy, P.A., Smith, P.J., Limdi, J.K., Hayee, B., Lomer, M.C.E., Parkes, G.C., Selinger, C., Barrett, K.J., Davies, R.J., Bennett, C., Gittens, S., Dunlop, M.G., Faiz, O., Fraser, A., Garrick, V., Johnston, P.D., Parkes, M., Sanderson, J., Terry, H., Gaya, D.R., Iqbal, T.H., Taylor, S.A., Smith, M., Brookes, M., Hansen, R. and Hawthorne, A.B. (2019) 'British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults', Gut, 68:s1–s106.

Lamps, L.W., Madhusudhan, K.T., Havens, J.M., Greenson, J.K., Bronner, M.P., Chiles, M.C., Dean, P.J. and Scott, M.A. (2003) 'Pathogenic Yersinia DNA is detected in bowel and mesenteric lymph nodes from patients with Crohn's disease', The American journal of surgical pathology, 27(2):220–227.

Langley, B.O., Guedry, S.E., Goldenberg, J.Z., Hanes, D.A., Beardsley, J.A. and Ryan, J.J. (2021) 'Inflammatory bowel disease and neutrophil–lymphocyte ratio: A systematic scoping review', Journal of Clinical Medicine, 10(18):4219.

Lanier, T.J., Schall, R.D., Berahovich, N.L., Lai, Z. and Wei, L.L. (2022) 'Chemokine Receptor Expression Evidence for NK Cell Subsets Based on'. Journal of Immunology, 177(11):7833-7840.

Laufer, J., Oren, R., Goldberg, I., Horwitz, A., Kopolovic, J., Chowers, Y. and Passwell, J.H. (2001) 'Cellular localization of complement C3 and C4 transcripts in intestinal specimens from patients with Crohn's disease', Clinical and Experimental Immunology, 120(1):30–37.

Leal, R.F., Planell, N., Kajekar, R., Lozano, J.J., Ordás, I., Dotti, I., Esteller, M.,

Masamunt, M.C., Parmar, H., Ricart, E., Panés, J. and Salas, A. (2015a) 'Identification of inflammatory mediators in patients with Crohn's disease unresponsive to anti-TNFα therapy', Gut, 64(2):233–242.

Lee, C.M., He, C.H., Nour, A.M., Zhou, Y., Ma, B., Park, J.W., Kim, K.H., Cruz, C. Dela, Sharma, L., Nasr, M.L., Modis, Y., Lee, C.G. and Elias, J.A. (2016) 'IL-13Rα2 uses TMEM219 in chitinase 3-like-1-induced signalling and effector responses', Nature Communications 2016 7:1, 7(1):1–13.

Lei, H., Kuchenbecker, L., Streitz, M., Sawitzki, B., Vogt, K., Landwehr-Kenzel, S., Millward, J., Juelke, K., Babel, N., Neumann, A., Reinke, P. and Volk, H.D. (2015) 'Human CD45RA− FoxP3hi Memory-Type Regulatory T Cells Show Distinct TCR Repertoires With Conventional T Cells and Play an Important Role in Controlling Early Immune Activation', American Journal of Transplantation, 15(10):2625–2635.

Leist, T.P., Kohler, M., Eppler, M. and Zinkernagel, R.M. (1989) 'Effects of treatment with IL-2 receptor specific monoclonal antibody in mice. Inhibition of cytotoxic T cell responses but not of T help.', The Journal of Immunology, 143(2):628-632.

Leon, F., Roldan, E., Sanchez, L., Camarero, C., Bootello, A. and Roy, G. (2003) 'Human small-intestinal epithelium contains functional natural killer lymphocytes', Gastroenterology, 125(2):345–356.

Ley, R.E., Peterson, D.A. and Gordon, J.I. (2006) 'Ecological and evolutionary forces shaping microbial diversity in the human intestine', Cell, 124(4):837–848.

Li, J., He, Y., Hao, J., Ni, L. and Dong, C. (2018) 'High Levels of Eomes Promote Exhaustion of Anti-tumor CD8+ T Cells', Frontiers in Immunology, 9:2981.

Li, X., Kang, B., Eom, Y., Zhong, J., Lee, H.K., Kim, H.M. and Song, J.S. (2022) 'Comparison of cytotoxicity effects induced by four different types of nanoparticles in human corneal and conjunctival epithelial cells', Scientific Reports, 12(1):1–10.

Liang, S.C., Tan, X.Y., Luxenberg, D.P., Karim, R., Dunussi-Joannopoulos, K., Collins, M. and Fouser, L.A. (2006) 'Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides', Journal of Experimental Medicine, 203(10):2271–2279.

Libermann, T.A. and Baltimore, D. (1990) 'Activation of interleukin-6 gene expression through the NF-kappa B transcription factor.', Molecular and Cellular Biology, 10(5):2327-2334.

Lichtenstein, G.R., Targan, S.R., Dubinsky, M.C., Rotter, J.I., Barken, D.M., Princen, F., Carroll, S., Brown, M., Stachelski, J., Chuang, E., Landers, C.J., Stempak, J.M., Singh, S. and Silverberg, M.S. (2011) 'Combination of genetic and quantitative serological immune markers are associated with complicated Crohn's disease behavior', Inflammatory bowel diseases, 17(12), pp. 2488–2496.

Lima, M., Leander, M., Santos, M., Santos, A.H., Lau, C., Queirós, M.L., Gonçalves, M., Fonseca, S., Moura, J., Teixeira, M.D.A. and Orfao, A. (2015) 'Chemokine Receptor Expression on Normal Blood CD56+ NK-Cells Elucidates Cell Partners That Comigrate during the Innate and Adaptive Immune Responses and Identifies a Transitional NK-Cell Population', Journal of Immunology Research, 2015:839684.

Linares, R., Francés, R., Gutiérrez, A. and Juanola, O. (2021) 'Bacterial Translocation as Inflammatory Driver in Crohn's Disease', Frontiers in Cell and Developmental Biology, 9:703310.

Liso, M., Verna, G., Cavalcanti, E., De Santis, S., Armentano, R., Tafaro, A., Lippolis, A., Campiglia, P., Gasbarrini, A., Mastronardi, M., Pizarro, T.T., Cominelli, F., Lopetuso, L.R. and Chieppa, M. (2022) 'Interleukin 1β Blockade Reduces Intestinal Inflammation in a Murine Model of Tumor Necrosis Factor–Independent Ulcerative Colitis', Cellular and Molecular Gastroenterology and Hepatology, 14(1):151-171.

Liu, F.D.M., Kenngott, E.E., Schröter, M.F., Kühl, A., Jennrich, S., Watzlawick, R., Hoffmann, U., Wolff, T., Norley, S., Scheffold, A., Stumhofer, J.S., Saris, C.J.M., Schwab, J.M., Hunter, C.A., Debes, G.F. and Hamann, A. (2014) 'Timed

Action of IL-27 Protects from Immunopathology while Preserving Defense in Influenza', PLOS Pathogens, 10(5):e1004110.

Liu, J.Z., Van Sommeren, S., Huang, H., Ng, S.C., Alberts, R., Takahashi, A., Ripke, S., Lee, J.C., Jostins, L., Shah, T., Abedian, S., Cheon, J.H., Cho, J., Daryani, N.E., Franke, L., Fuyuno, Y., Hart, A., Juyal, R.C., Juyal, G., Kim, W.H., Morris, A.P., Poustchi, H., Newman, W.G., Midha, V., Orchard, T.R., Vahedi, H., Sood, A., Sung, J.J.Y., Malekzadeh, R., Westra, H.J., Yamazaki, K., Yang, S.K., Barrett, J.C., Franke, A., Alizadeh, B.Z., Parkes, M., Thelma, B.K., Daly, M.J., Kubo, M., Anderson, C.A. and Weersma, R.K. (2015) 'Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations', Nature genetics, 47(9):979–986.

Liu, L., Dong, Y., Ye, M., Jin, S., Yang, J., Joosse, M.E., Sun, Y., Zhang, J., Lazarev, M., Brant, S.R., Safar, B., Marohn, M., Mezey, E. and Li, X. (2017) 'The Pathogenic Role of NLRP3 Inflammasome Activation in Inflammatory Bowel Diseases of Both Mice and Humans', Journal of Crohn's and Colitis, 11(6):737–750.

Liu, S., Ren, J., Han, G., Wang, G., Gu, G., Xia, Q. and Li, J. (2012) 'Mean platelet volume: A controversial marker of disease activity in Crohn's disease', European Journal of Medical Research, 17(1):1–7.

Liu, S., Russo, P.A., Baldassano, R.N. and Sullivan, K.E. (2009) 'CD68 Expression Is Markedly Different in Crohn's Disease and the Colitis Associated with Chronic Granulomatous Disease', Inflammatory Bowel Diseases, 15(8):1213–1217.

Liu, T., Zhang, L., Joo, D. and Sun, S.C. (2017) 'NF-κB signaling in inflammation', Signal Transduction and Targeted Therapy, 2(1):1–9.

Liu, Z., Liu, J.Q., Talebian, F., Wu, L.C., Li, S. and Bai, X.F. (2013) 'IL-27 enhances the survival of tumor antigen-specific CD8+ T cells and programs them into IL-10-producing, memory precursor-like effector cells', European journal of immunology, 43(2):468–479.

Louis, E., Collard, A., Oger, A.F., Degroote, E., Aboul Nasr El Yafi, F. and

Belaiche, J. (2001) ‘Behaviour of Crohn’s disease according to the Vienna classification: changing pattern over the course of the disease’, Gut, 49(6):777–782.

Ludwiczek, O., Vannier, E., Borggraefe, I., Kaser, A., Siegmund, B., Dinarello, C.A. and Tilg, H. (2004) ‘Imbalance between interleukin-1 agonists and antagonists: relationship to severity of inflammatory bowel disease’, Clinical and Experimental Immunology, 138(2):323-329.

Ma, D., Huang, H. and Huang, Z. (2010) ‘STAT1 signaling is required for optimal Th1 cell differentiation in mice’, Chinese Science Bulletin, 55(11):1032–1040.

Maaser, C., Sturm, A., Vavricka, S.R., Kucharzik, T., Fiorino, G., Annese, V., Calabrese, E., Baumgart, D.C., Bettenworth, D., Borralho Nunes, P., Burisch, J., Castiglione, F., Eliakim, R., Ellul, P., González-Lama, Y., Gordon, H., Halligan, S., Katsanos, K., Kopylov, U., Kotze, P.G., Krustiņš, E., Laghi, A., Limdi, J.K., Rieder, F., Rimola, J., Taylor, S.A., Tolan, D., Van Rheenen, P., Verstockt, B. and Stoker, J. (2019) ‘ECCO-ESGAR Guideline for Diagnostic Assessment in IBD Part 1: Initial diagnosis, monitoring of known IBD, detection of complications’, Journal of Crohn’s and Colitis, 13(2):144-164K.

MacDonald, K.P.A., Palmer, J.S., Cronau, S., Seppanen, E., Olver, S., Raffelt, N.C., Kuns, R., Pettit, A.R., Clouston, A., Wainwright, B., Branstetter, D., Smith, J., Paxton, R.J., Cerretti, D.P., Bonham, L., Hill, G.R. and Hume, D.A. (2010) ‘An antibody against the colony-stimulating factor 1 receptor depletes the resident subset of monocytes and tissue- and tumor-associated macrophages but does not inhibit inflammation’, Blood, 116(19):3955–3963.

Mack, D.R., Beedle, S., Warren, J., Davis, J. and Gross, T. (2002) ‘Peripheral Blood Intracellular Cytokine Analysis in Children Newly Diagnosed with Inflammatory Bowel Disease’, Pediatric Research, 51(3):328-332.

Maheshwari, A., Kelly, D.R., Nicola, T., Ambalavanan, N., Jain, S.K., Murphyullrich, J., Athar, M., Shimamura, M., Bhandari, V., Aprahamian, C., Dimmitt, R.A., Serra, R. and Ohls, R.K. (2011) ‘TGF-β2 suppresses macrophage cytokine production and mucosal inflammatory responses in the developing

intestine', Gastroenterology, 140(1):242–253.

Mantovani, A., Sica, A., Sozzani, S., Allavena, P., Vecchi, A. and Locati, M. (2004) 'The chemokine system in diverse forms of macrophage activation and polarization', Trends in Immunology, 25(12):677–686.

Marks, D.J.B., Harbord, M.W.N., MacAllister, R., Rahman, F.Z., Young, J., Al-Lazikani, B., Lees, W., Novelli, M., Bloom, S. and Segal, A.W. (2006) 'Defective acute inflammation in Crohn's disease: a clinical investigation', The Lancet, 367(9511):668–678.

Martin, J.C., Chang, C., Boschetti, G., Ungaro, R., Giri, M., Grout, J.A., Gettler, K., Chuang, L. shiang, Nayar, S., Greenstein, A.J., Dubinsky, M., Walker, L., Leader, A., Fine, J.S., Whitehurst, C.E., Mbow, M.L., Kugathasan, S., Denson, L.A., Hyams, J.S., Friedman, J.R., Desai, P.T., Ko, H.M., Laface, I., Akturk, G., Schadt, E.E., Salmon, H., Gnjatic, S., Rahman, A.H., Merad, M., Cho, J.H. and Kenigsberg, E. (2019) 'Single-Cell Analysis of Crohn's Disease Lesions Identifies a Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy', Cell, 178(6):1493-1508.

Mary, J. and Modigliani, R. (1989) 'Development and validation of an endoscopic index of the severity for Crohn's disease: a prospective multicentre study', Gastrointestinal Endoscopy, 30:983–989.

Mayer, J.U., Brown, S.L., MacDonald, A.S. and Milling, S.W. (2020) 'Defined Intestinal Regions Are Drained by Specific Lymph Nodes That Mount Distinct Th1 and Th2 Responses Against Schistosoma mansoni Eggs', Frontiers in Immunology, 11:2745.

McKenzie, G.J., Bancroft, A., Grencis, R.K. and McKenzie, A.N.J. (1998) 'A distinct role for interleukin-13 in Th2-cell-mediated immune responses', Current Biology, 8(6):339–342.

McLarren, K.W., Cole, A.E., Weisser, S.B., Voglmaier, N.S., Conlin, V.S., Jacobson, K., Popescu, O., Boucher, J.L. and Sly, L.M. (2011) 'SHIP-Deficient Mice Develop Spontaneous Intestinal Inflammation and Arginase-Dependent

Fibrosis', The American Journal of Pathology, 179(1):180-188.

McMurchy, A.N., Gillies, J., Gizzi, M.C., Riba, M., Garcia-Manteiga, J.M., Cittaro, D., Lazarevic, D., Di Nunzio, S., Piras, I.S., Bulfone, A., Roncarolo, M.G., Stupka, E., Bacchetta, R. and Levings, M.K. (2013) 'A novel function for FOXP3 in humans: intrinsic regulation of conventional T cells', Blood, 121(8), pp. 1265–1275.

MFMER (2022) Inflammatory bowel disease (IBD) - Symptoms and causes - Mayo Clinic. Available at: https://www.mayoclinic.org/diseases-conditions/inflammatory-bowel-disease/symptoms-causes/syc-20353315 (Accessed: 5 July 2022).

Mills, C., Shearer, J., Evans, R. and Caldwell, M. (1992) 'Macrophage arginine metabolism and the inhibition or stimulation of cancer', Journal of Immunology, 149(8):2709–14.

Mills, C.D., Kincaid, K., Alt, J.M., Heilman, M.J. and Hill, A.M. (2000) 'M-1/M-2 Macrophages and the Th1/Th2 Paradigm', The Journal of Immunology, 164(12):6166–6173.

Mills, S., Trego, A.C., Lens, P.N.L., Ijaz, U.Z. and Collins, G. (2021) 'A Distinct, Flocculent, Acidogenic Microbial Community Accompanies Methanogenic Granules in Anaerobic Digesters', Microbiology Spectrum, 9(3):e0078421.

Mitselos, I. V., Katsanos, K.H., Tatsioni, A., Skamnelos, A., Eliakim, R., Tsianos, E. V. and Christodoulou, D.K. (2018) 'Association of clinical and inflammatory markers with small bowel capsule endoscopy findings in Crohn's disease', European journal of gastroenterology & hepatology, 30(8):861–867.

Mizoguchi, E. (2006) 'Chitinase 3-Like-1 Exacerbates Intestinal Inflammation by Enhancing Bacterial Adhesion and Invasion in Colonic Epithelial Cells'. Gastroenterology, 130:398-411.

Moeller, A.H., Li, Y., Ngole, E.M., Ahuka-Mundeke, S., Lonsdorf, E. V., Pusey, A.E., Peeters, M., Hahn, B.H. and Ochman, H. (2014) 'Rapid changes in the gut

microbiome during human evolution', Proceedings of the National Academy of Sciences of the United States of America, 111(46):16431–16435.

Mogensen, T.H. (2019) 'IRF and STAT transcription factors - From basic biology to roles in infection, protective immunity, and primary immunodeficiencies', Frontiers in Immunology, 10:3047.

Molnár, T., Tiszlavicz, L., Gyulai, C., Nagy, F. and Lonovics, J. (2005) 'Clinical significance of granuloma in Crohn's disease', World Journal of Gastroenterology, 11(20):3118.

Monteleone, G., Monteleone, I., Fina, D., Vavassori, P., Del Vecchio Blanco, G., Caruso, R., Tersigni, R., Alessandroni, L., Biancone, L., Naccari, G.C., Macdonald, T.T. and Pallone, F. (2005) 'Interleukin-21 Enhances T-Helper Cell Type I Signaling and Interferon-Production in Crohn's Disease'. Gastronenterology, 128(3):687-694.

Morain, C.O., Segal, A.A., Walker, D. and Levi, A.J. (1981) 'Abnormalities of neutrophil function do not cause the migration defect in Crohn's disease.', Gut, 22(10):817-822.

Morhardt, T.L., Hayashi, A., Ochi, T., Quirós, M., Kitamoto, S., Nagao-Kitamoto, H., Kuffa, P., Atarashi, K., Honda, K., Kao, J.Y., Nusrat, A. and Kamada, N. (2019) 'IL-10 produced by macrophages regulates epithelial integrity in the small intestine', Scientific Reports, 9(1):1–10.

Morishima, N., Owaki, T., Asakawa, M., Kamiya, S., Mizuguchi, J. and Yoshimoto, T. (2005) 'Augmentation of Effector CD8+ T Cell Generation with Enhanced Granzyme B Expression by IL-27', The Journal of Immunology, 175(3):1686–1693.

Morley, J.J. and Kushner, I. (1982) 'SERUM C-REACTIVE PROTEIN LEVELS IN DISEASE*', Annals of the New York Academy of Sciences, 389(1):406–418.

Müller, S., Lory, J., Corazza, N., Griffiths, G.M., Z'graggen, K., Mazzucchelli, L., Kappeler, A. and Mueller, C. (1998) 'Activated CD4+ and CD8+ cytotoxic cells are

present in increased numbers in the intestinal mucosa from patients with active inflammatory bowel disease.', The American Journal of Pathology, 152(1):261–268.

Murch, S. (2021) 'Gastrointestinal Mucosal Immunology and Mechanisms of Inflammation', Pediatric Gastrointestinal and Liver Disease:40-52.e3.

Muzes, G., Molnár, B., Tulassay, Z. and Sipos, F. (2012) 'Changes of the cytokine profile in inflammatory bowel diseases', World Journal of Gastroenterology : WJG, 18(41):5848-5861.

Nathan, C. and Shiloh, M.U. (2000) 'Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens', Proceedings of the National Academy of Sciences of the United States of America, 97(16):8841–8848.

Nathan, C.F., Murray, H.W., Wlebe, I.E. and Rubin, B.Y. (1983) 'Identification of interferon-gamma as the lymphokine that activates human macrophage oxidative metabolism and antimicrobial activity', The Journal of Experimental Medicine, 158(3):670-689.

Neubauer, K., Woźniak-Stolarska, B. and Krzystek-Korpacka, M. (2018) 'Peripheral Lymphocytes of Patients with Inflammatory Bowel Disease Have Altered Concentrations of Key Apoptosis Players: Preliminary Results', BioMed Research International, 2018:4961753.

Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D., Diehn, M. and Alizadeh, A.A. (2019a) 'Determining cell type abundance and expression from bulk tissues with digital cytometry', Nature Biotechnology, 37(7):773–782.

Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D., Diehn, M. and Alizadeh, A.A. (2019b) 'Determining cell type abundance and expression from bulk tissues with digital cytometry', Nature Biotechnology, 37(7):773–782.

Ng, S.C., Bernstein, C.N., Vatn, M.H., Lakatos, P.L., Loftusjr, E. V, Tysk, C., O'morain, C., Moum, B., Siew, D. and Ng, C. (2013) 'Geographical variability and environmental risk factors in inflammatory bowel disease'. Gut, 62(4):630-649.

Ng, S.C., Tang, W. and Leong, R.W. (2015) 'Environmental risk factors in inflammatory bowel disease: a population-based case-control study in Asia-Pacific', Gut, 64(7):1063–1071.

Ngoh, E.N., Weisser, S.B., Lo, Y., Kozicky, L.K., Jen, R., Brugger, H.K., Menzies, S.C., McLarren, K.W., Nackiewicz, D., Van Rooijen, N., Jacobson, K., Ehses, J.A., Turvey, S.E. and Sly, L.M. (2016) 'Activity of SHIP, Which Prevents Expression of Interleukin 1ß, Is Reduced in Patients with Crohn's Disease', Gastroenterology, 150(2):465–476.

Nguyen, P., Melzer, M., Beck, F., Krasselt, M., Seifert, O., Pierer, M., Rothe, K. and Wagner, U. (2022) 'Expansion of CD4+CD8+ double-positive T cells in rheumatoid arthritis patients is associated with erosive disease', Rheumatology, 61(3):1282–1287.

NICE (2017) Corticosteroids - oral, Clinical Knowledge Summaries. Available at: https://cks.nice.org.uk/corticosteroids-oral#!scenario (Accessed: 8 March 2019).

NICE (2022) Recommendations | Crohn's disease: management | Guidance | NICE. Available at: https://www.nice.org.uk/guidance/ng129/chapter/Recommendations (Accessed: 19 July 2022).

Nishikomori, R., Usui, T., Wu, C.-Y., Morinobu, A., O'Shea, J.J. and Strober, W. (2002) 'Activated STAT4 Has an Essential Role in Th1 Differentiation and Proliferation That Is Independent of Its Role in the Maintenance of IL-12Rß2 Chain Expression and Signaling', The Journal of Immunology, 169(8):4388–4398.

Noble, A., Durant, L., Hoyles, L., McCartney, A.L., Man, R., Segal, J., Costello, S.P., Hendy, P., Reddi, D., Bouri, S., Lim, D.N.F., Pring, T., O'Connor, M.J., Datt, P., Wilson, A., Arebi, N., Akbar, A., Hart, A.L., Carding, S.R. and Knight, S.C. (2020) 'Deficient Resident Memory T Cell and CD8 T Cell Response to

Commensals in Inflammatory Bowel Disease', Journal of Crohn's & Colitis, 14(4):525-537.

O'Brien, C.L., Pavli, P., Gordon, D.M. and Allison, G.E. (2014) 'Detection of bacterial DNA in lymph nodes of Crohn's disease patients using high throughput sequencing', Gut, 63(10):1596–1606.

Ogawa, C., Tone, Y., Tsuda, M., Peter, C., Waldmann, H. and Tone, M. (2014) 'TGF-β–Mediated Foxp3 Gene Expression Is Cooperatively Regulated by Stat5, Creb, and AP-1 through CNS2', The Journal of Immunology, 192(1):475–483.

Ogura, Y., Bonen, D.K., Inohara, N., Nicolae, D.L., Chen, F.F., Ramos, R., Britton, H., Moran, T., Karaliuskas, R., Duerr, R.H., Achkar, J.P., Brant, S.R., Bayless, T.M., Kirschner, B.S., Hanauer, S.B., Nũez, G. and Cho, J.H. (2001) 'A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease', Nature 2001 411:6837, 411(6837):603–606.

Okou, D.T. and Kugathasan, S. (2014) 'The Role of Genetics In Pediatric Inflammatory Bowel Disease', Inflammatory bowel diseases, 20(10):1878-1884.

Okumura, R. and Takeda, K. (2017) 'Roles of intestinal epithelial cells in the maintenance of gut homeostasis', Experimental & molecular medicine, 49:e338.

Olaisen, M., Flatberg, A., Granlund, A.V.B., Røyset, E.S., Martinsen, T.C., Sandvik, A.K. and Fossmark, R. (2021) 'Bacterial Mucosa-associated Microbiome in Inflamed and Proximal Noninflamed Ileum of Patients With Crohn's Disease', Inflammatory Bowel Diseases, 27(1):12-24.

Oliveros, J.C. (2007) VENNY. An interactive tool for comparing lists with Venn Diagrams. Available at: https://bioinfogp.cnb.csic.es/tools/venny_old/venny.php (Accessed: 28 July 2022).

Omenetti, S. and Pizarro, T.T. (2015) 'The Treg/Th17 axis: A dynamic balance regulated by the gut microbiome', Frontiers in Immunology, 6:639.

Oostenbrug, L.E., Drenth, J.P.H., De Jong, D.J., Nolte, I.M., Oosterom, E., Van Dullemen, H.M., Van Der Linde, K., Te Meerman, G.J., Van Der Steege, G., Kleibeuker, J.H. and Jansen, P.L.M. (2005) 'Association between Toll-like receptor 4 and inflammatory bowel disease', Inflammatory bowel diseases, 11(6):567–575.

Ouellette, A.J. (2005) 'Paneth cell α-defensins: Peptide mediators of innate immunity in the small intestine', Springer Seminars in Immunopathology, 27(2):133–146.

Pagán, A.J. and Ramakrishnan, L. (2018) 'The Formation and Function of Granulomas', Annual Review of Immunology, 36:639-665.

Palmer, N.P., Silvester, J.A., Lee, J.J., Beam, A.L., Fried, I., Valtchinov, V.I., Rahimov, F., Kong, S.W., Ghodoussipour, S., Hood, H.C., Bousvaros, A., Grand, R.J., Kunkel, L.M. and Kohane, I.S. (2019) 'Concordance between gene expression in peripheral whole blood and colonic tissue in children with inflammatory bowel disease', PLoS ONE, 14(10):e0222952.

Panaccione, R. (2013) 'Mechanisms of Inflammatory Bowel Disease', Gastroenterology & Hepatology, 9(8):529-532.

Papay, P., Ignjatovic, A., Karmiris, K., Amarante, H., Milheller, P., Feagan, B., D'Haens, G., Marteau, P., Reinisch, W., Sturm, A., Steinwurz, F., Egan, L., Panés, J., Louis, E., Colombel, J.F. and Panaccione, R. (2013) 'Optimising monitoring in the management of Crohn's disease: A physician's perspective', Journal of Crohn's and Colitis, 7(8):653–669.

Pardoux, C., Asselin-Paturel, C., Chehimi, J., Gay, F., Mami-Chouaib, F. and Chouaib, S. (1997) 'Functional interaction between TGF-beta and IL-12 in human primary allogeneic cytotoxicity and proliferative response.', The Journal of Immunology, 158(1):136-143.

Parronchi, P., Romagnani, P., Annunziato, F., Sampognaro, S., Becchio, A., Giannarini, L., Maggi, E., Pupilli, C., Tonelli, F. and Romagnani, S. (1997) 'Type 1 T-helper cell predominance and interleukin-12 expression in the gut of

patients with Crohn's disease.', The American Journal of Pathology, 150(3), p. 823-832.

Pearce, E.L., Mullen, A.C., Martins, G.A., Krawczyk, C.M., Hutchins, A.S., Zediak, V.P., Banica, M., DiCioccio, C.B., Gross, D.A., Mao, C.A., Shen, H., Cereb, N., Yang, S.Y., Lindsten, T., Rossant, J., Hunter, C.A. and Reiner, S.L. (2003) 'Control of Effector CD8+ T Cell Function by the Transcription Factor Eomesodermin', Science, 302(5647):1041–1043.

Pelletier, M., Maggi, L., Micheletti, A., Lazzeri, E., Tamassia, N., Costantini, C., Cosmi, L., Lunardi, C., Annunziato, F., Romagnani, S. and Cassatella, M.A. (2010) 'Evidence for a cross-talk between human neutrophils and Th17 cells', Blood, 115(2):335–343.

Pène, J., Chevalier, S., Preisser, L., Vénéreau, E., Guilleux, M.-H., Ghannam, S., Molès, J.-P., Danger, Y., Ravon, E., Lesaux, S., Yssel, H. and Gascan, H. (2008) 'Chronically Inflamed Human Tissues Are Infiltrated by Highly Differentiated Th17 Lymphocytes', The Journal of Immunology, 180(11):7423–7430.

Penna, F.G.C.E., Rosa, R.M., Sales Da Cunha, P.F., Silva De Souza, S.C. and De Lourdes De Abreu Ferrari, M. (2020) 'Faecal calprotectin is the biomarker that best distinguishes remission from different degrees of endoscopic activity in Crohn's disease', BMC Gastroenterology, 20(1).

Perussia, B., Chen, Y. and Loza, M.J. (2005) 'Peripheral NK cell phenotypes: multiple changing of faces of an adapting, developing cell', Molecular immunology, 42(4):385–395.

Pesce, S., Squillario, M., Greppi, M., Loiacono, F., Moretta, L., Moretta, A., Sivori, S., Castagnola, P., Barla, A., Candiani, S. and Marcenaro, E. (2018) 'New miRNA Signature Heralds Human NK Cell Subsets at Different Maturation Steps: Involvement of miR-146a-5p in the Regulation of KIR Expression', Frontiers in immunology, 9:2360.

Peters, L.A., Perrigoue, J., Mortha, A., Iuga, A., Song, W.M., Neiman, E.M., Llewellyn, S.R., Di Narzo, A., et al. (2017) 'A functional genomics predictive

network model identifies regulators of inflammatory bowel disease', Nature genetics, 49(10):1437–1449.

Peters, P.J., Borst, J., Oorschot, V., Fukuda, M., Krahenbuhl, O., Tschopp, J., Slot, J.W. and Geuze, H.J. (1991) 'Cytotoxic T Lymphocyte Granules Are Secretory Lysosomes, Containing Both Perforin and Granzymes', Journal of Experimental Medicine, 173(5):1099–1109.

Peterson, D.A., Frank, D.N., Pace, N.R. and Gordon, J.I. (2008) 'Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases', Cell host & microbe, 3(6):417–427.

Pflanz, S., Timans, J.C., Cheung, J., Rosales, R., Kanzler, H., Gilbert, J., Hibbert, L., Churakova, T., Travis, M., Vaisberg, E., Blumenschein, W.M., Mattson, J.D., Wagner, J.L., To, W., Zurawski, S., McClanahan, T.K., Gorman, D.M., Bazan, J.F., De Waal Malefyt, R., Rennick, D. and Kastelein, R.A. (2002) 'IL-27, a heterodimeric cytokine composed of EBI3 and p28 protein, induces proliferation of naive CD4+ T cells', Immunity, 16(6):779–790.

Planell, N., Masamunt, M.C., Leal, R.F., Rodríguez, L., Esteller, M., Lozano, J.J., Ramírez, A., Ayrizono, M. de L.S., Coy, C.S.R., Alfaro, I., Ordás, I., Visvanathan, S., Ricart, E., Guardiola, J., Panés, J. and Salas, A. (2017) 'Usefulness of transcriptional blood biomarkers as a non-invasive surrogate marker of mucosal healing and endoscopic response in ulcerative colitis', Journal of Crohn's and Colitis, 11(11):1335–1346.

Platt, A.M. and Mowat, A.M.I. (2008) 'Mucosal macrophages and the regulation of immune responses in the intestine', Immunology letters, 119(1–2):22–31.

Ponchel, F, Burska, A., Hunt, L., Gul, H., Rabin, T., Parmar, R., Buch, M., Conaghan, P. and Emery, P. (2020) 'T-cell subset abnormalities predict progression along the Inflammatory Arthritis disease continuum: implications for management', Scientific Reports, 10(1):1–10.

Ponchel, Frederique, Burska, A.N., Hunt, L., Gul, H., Rabin, T., Parmar, R., Buch, M.H., Conaghan, P.G. and Emery, P. (2020) 'T-cell subset abnormalities

predict progression along the Inflammatory Arthritis disease continuum: implications for management', Scientific Reports, 10(1):1–10.

Posovszky, C., Pfalzer, V., Lahr, G., Niess, J.H., Klaus, J., Mayer, B., Debatin, K.M. and von Boyen, G.B.T. (2013) 'Age-of-onset-dependent influence of NOD2 gene variants on disease behaviour and treatment in Crohn's disease', BMC Gastroenterology, 13(1):77.

Prins, M.M., Verstockt, B., Ferrante, M., Vermeire, S., Wildenberg, M.E. and Koelink, P.J. (2021) 'Monocyte TREM-1 Levels Associate With Anti-TNF Responsiveness in IBD Through Autophagy and Fcγ-Receptor Signaling Pathways', Frontiers in Immunology, 12:627535.

Pulgaron, E.R. and Delamater, A.M. (2014) 'Obesity and Type 2 Diabetes in Children: Epidemiology and Treatment', Current diabetes reports, 14(8):508.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., et al. (2010) 'A human gut microbial gene catalog established by metagenomic sequencing', Nature, 464(7285):59-65.

Qiu, Q.C., Wang, L., Jin, S.S., Liu, G.F., Liu, J., Ma, L., Mao, R.F., Ma, Y.Y., Zhao, N., Chen, M. and Lin, B.Y. (2018) 'CHI3L1 promotes tumor progression by activating TGF-β signaling pathway in hepatocellular carcinoma', Scientific Reports 2018 8:1, 8(1):1–12.

Quezada, S.M., Steinberger, E.K. and Cross, R.K. (2013) 'Association of age at diagnosis and Crohn's disease phenotype', Age and Ageing, 42(1):102-106.

Ranasinghe, R. and Eri, R. (2018) 'Modulation of the CCR6-CCL20 Axis: A Potential Therapeutic Target in Inflammation and Cancer', Medicina, 54(5):88.

Rankin, G.B., Watts, H.D., Melnyk, C.S. and Kelley, M.L. (1979) 'National Cooperative Crohn's Disease Study: Extraintestinal manifestations and perianal complications', Gastroenterology, 77(4):914–920.

Rathmell, J.C., Farkash, E.A., Gao, W. and Thompson, C.B. (2001) 'IL-7

enhances the survival and maintains the size of naive T cells', Journal of immunology, 167(12):6869–6876.

Ratner, A. and Clark, W.R. (1993) 'Role of TNF-alpha in CD8+ cytotoxic T'. Journal of Immunology, 15(50):4303-4314.

Reiling, N., Hölscher, C., Fehrenbach, A., Kröger, S., Kirschning, C.J., Goyert, S. and Ehlers, S. (2002) 'Cutting edge: Toll-like receptor (TLR)2- and TLR4-mediated pathogen recognition in resistance to airborne infection with Mycobacterium tuberculosis', Journal of immunology, 169(7):3480–3484.

Rehli, M., Niller, H.H., Ammon, C., Langmann, S., Schwarzfischer, L., Andreesen, R. and Krause, S.W. (2003) 'Transcriptional Regulation of CHI3L1, a Marker Gene for Late Stages of Macrophage Differentiation', Journal of Biological Chemistry, 278(45):44058–44067.

Reinecker, H.C., Steffen, M., Witthoeft, T., Pflueger, I., Schreiber, S., MacDermott, R.P. and Raedler, A. (1993) 'Enhanced secretion of tumour necrosis factor-alpha, IL-6, and IL-1 beta by isolated lamina propria mononuclear cells from patients with ulcerative colitis and Crohn's disease.', Clinical and Experimental Immunology, 94(1):174-181.

Reinisch, W., Hung, K., Hassan-Zahraee, M. and Cataldi, F. (2018) 'Targeting Endothelial Ligands: ICAM-1/alicaforsen, MAdCAM-1', Journal of Crohn's & colitis, 12(suppl_2):S669–S677.

Richards, C.D. (2013) 'The Enigmatic Cytokine Oncostatin M and Roles in Disease', ISRN Inflammation, 2013:512103.

Richardson, M.D., Gray, C.A. and Shankland, G.S. (1991) 'Opsonic effect of C-reactive protein on phagocytosis and intracellular killing of virulent and attenuated strains of Candida albicans by human neutrophils', FEMS microbiology immunology, 3(6):341–344.

Rioux, J.D., Xavier, R.J., Taylor, K.D., Silverberg, M.S., Goyette, P., Huett, A., Green, T., Kuballa, P., Barmada, M.M., Datta, L.W., Shugart, Y.Y., Griffiths,

A.M., Targan, S.R., Ippoliti, A.F., Bernard, E.J., Mei, L., Nicolae, D.L., Regueiro, M., Schumm, L.P., Steinhart, A.H., Rotter, J.I., Duerr, R.H., Cho, J.H., Daly, M.J. and Brant, S.R. (2007) 'Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis', Nature genetics, 39(5):596–604.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C. and Müller, M. (2011) 'pROC: An open-source package for R and S+ to analyze and compare ROC curves', BMC Bioinformatics, 12(1):1–8.

Rogler, G., Singh, A., Kavanaugh, A. and Rubin, D.T. (2021) 'Extraintestinal Manifestations of Inflammatory Bowel Disease: Current Concepts, Treatment, and Implications for Disease Management', Gastroenterology, 161(4):1118–1132.

Rehli, M. et al. (2003) 'Transcriptional Regulation of CHI3L1, a Marker Gene for Late Stages of Macrophage Differentiation', Journal of Biological Chemistry. Elsevier, 278(45):44058–44067.

Roach, D.R., Bean, A.G.D., Demangel, C., France, M.P., Briscoe, H. and Britton, W.J. (2002) 'TNF regulates chemokine induction essential for cell recruitment, granuloma formation, and clearance of mycobacterial infection', *Journal of immunology,* 168(9):4620–4627.

Rohart, F., Gautier, B., Singh, A. and Le Cao, K.-A. (2017) 'mixOmics: An R package for 'omics feature selection and multiple data integration.', PLoS computational biology, 13(11):e1005752.

Rosenbloom, A.L., Joe, J.R., Young, R.S. and Winter, W.E. (1999) 'Emerging epidemic of type 2 diabetes in youth', Diabetes care, 22(2):345–354.

Rossy, J., Williamson, D.J. and Gaus, K. (2012) 'How does the kinase Lck phosphorylate the T cell receptor? Spatial organization as a regulatory mechanism', Frontiers in Immunology, 3:167.

Rousseaux, C., Lefebvre, B., Dubuquoy, L., Lefebvre, P., Romano, O., Auwerx, J., Metzger, D., Wahli, W., Desvergne, B., Naccari, G.C., Chavatte, P., Farce,

A., Bulois, P., Cortot, A., Colombel, J.F. and Desreumaux, P. (2005) 'Intestinal antiinflammatory effect of 5-aminosalicylic acid is dependent on peroxisome proliferator–activated receptor-γ', Journal of Experimental Medicine, 201(8):1205–1215.

rstudio-pubs-static (2000) All subset regression with leaps, bestglm, glmulti, and meifly. Available at: https://rstudio-pubs-static.s3.amazonaws.com/2897_9220b21cfc0c43a396ff9abf122bb351.html (Accessed: 20 July 2022).

Rubtsov, Y.P., Rasmussen, J.P., Chi, E.Y., Fontenot, J., Castelli, L., Ye, X., Treuting, P., Siewe, L., Roers, A., Henderson, W.R., Muller, W. and Rudensky, A.Y. (2008) 'Regulatory T cell-derived interleukin-10 limits inflammation at environmental interfaces', Immunity, 28(4):546–558.

Rudd, C.E., Taylor, A. and Schneider, H. (2009) 'CD28 and CTLA-4 coreceptor expression and signal transduction', Immunological reviews, 229(1):12–26.

Ryan, P., Kelly, R.G., Lee, G., Collins, J.K., O'Sullivan, G.C., O'Connell, J. and Shanahan, F. (2004) 'Bacterial DNA within granulomas of patients with Crohn's disease--detection by laser capture microdissection and PCR', *The American journal of gastroenterology*, 99(8):1539–1543.

Saad, A.M., Czul, F., Sakuraba, A., Rubin, D.T. and Cohen, R.D. (2016) 'Age of Diagnosis is Associated with Disease Presentation and Therapeutic Complications in Patients with Crohn's Disease', Inflammatory bowel diseases, 22(5):1027–1031.

Sallusto, F., Lenig, D., Förster, R., Lipp, M. and Lanzavecchia, A. (1999) 'Two subsets of memory T lymphocytes with distinct homing potentials and effector functions', Nature, 401(6754):708–712.

Sánchez-Muñoz, F., Fonseca-Camarillo, G., Villeda-Ramírez, M.A., Miranda-Pérez, E., Mendivil, E.J., Barreto-Zúñiga, R., Uribe, M., Bojalil, R., Domínguez-López, A. and Yamamoto-Furusho, J.K. (2011) 'Transcript levels of Toll-Like receptors 5, 8 and 9 correlate with inflammatory activity in Ulcerative Colitis',

BMC Gastroenterology, 11(1):138.

Sandborn, W.J., Feagan, B.G., Rutgeerts, P., Hanauer, S., Colombel, J.-F., Sands, B.E., Lukas, M., Fedorak, R.N., Lee, S., Bressler, B., Fox, I., Rosario, M., Sankoh, S., Xu, J., Stephens, K., Milch, C. and Parikh, A. (2013b) 'Vedolizumab as Induction and Maintenance Therapy for Crohn's Disease', New England Journal of Medicine, 369(8):711–721.

Sandborn, W.J., Gasink, C., Gao, L.-L., Blank, M.A., Johanns, J., Guzzo, C., Sands, B.E., Hanauer, S.B., Targan, S., Rutgeerts, P., Ghosh, S., de Villiers, W.J.S., Panaccione, R., Greenberg, G., Schreiber, S., Lichtiger, S. and Feagan, B.G. (2012) 'Ustekinumab Induction and Maintenance Therapy in Refractory Crohn's Disease', New England Journal of Medicine, 367(16), pp. 1519–1528. Available at: https://doi.org/10.1056/NEJMoa1203572.

Sandborn, W.J., Vermeire, S., Tyrrell, H., Hassanali, A., Lacey, S., Tole, S. and Tatro, A.R. (2020) 'Etrolizumab for the Treatment of Ulcerative Colitis and Crohn's Disease: An Overview of the Phase 3 Clinical Program', Advances in Therapy, 37(7):3417–3431.

Sands, B.E., Sandborn, W.J., Van Assche, G., Lukas, M., Xu, J., James, A., Abhyankar, B. and Lasch, K. (2017) 'Vedolizumab as Induction and Maintenance Therapy for Crohn's Disease in Patients Naïve to or Who Have Failed Tumor Necrosis Factor Antagonist Therapy', Inflammatory Bowel Diseases, 23(1):97–106.

Van De Sandt, C.E., Barcena, M., Koster, A.J., Kasper, J., Kirkpatrick, C.J., Scott, D.P., De Vries, R.D., Herold, S., Rimmelzwaan, G.F., Kuiken, T. and Short, K.R. (2017) 'Human CD81 T cells damage noninfected epithelial cells during influenza virus infection in vitro', American Journal of Respiratory Cell and Molecular Biology, 57(5):536–546.

Volck, B., Price, P.A., Johansen, J.S., Sorensen, O., Benfield, T.L., Nielsen, H.J., Calafat, J. and Borregaard, N. (1998) 'YKL-40, a mammalian member of the chitinase family, is a matrix protein of specific granules in human neutrophils.', Proceedings of the Association of American Physicians, 110(4):351–

360.

Sarr, D., Oliveira, L.J., Russ, B.N., Owino, S.O., Middii, J.D., Mwalimu, S., Ambasa, L., Almutairi, F., Vulule, J., Rada, B. and Moore, J.M. (2021) 'Myeloperoxidase and Other Markers of Neutrophil Activation Associate With Malaria and Malaria/HIV Coinfection in the Human Placenta', Frontiers in Immunology, 12:682668.

Sato, T., Nakai, T., Tamura, N., Okamoto, S., Matsuoka, K., Sakuraba, A., Fukushima, T., Uede, T. and Hibi, T. (2005) 'Osteopontin/Eta-1 upregulated in Crohn's disease regulates the Th1 immune response', Gut, 54(9):1254–1262.

Satsangi, J., Silverberg, M.S., Vermeire, S. and Colombel, J.F. (2006) 'The Montreal classification of inflammatory bowel disease: Controversies, consensus, and implications', Gut. 55(6):749–753.

Schenk, M., Bouchon, A., Seibold, F. and Mueller, C. (2007) 'TREM-1--expressing intestinal macrophages crucially amplify chronic inflammation in experimental colitis and inflammatory bowel diseases', The Journal of clinical investigation, 117(10):3097–3106.

Schoepfer, A.M., Beglinger, C., Straumann, A., Trummler, M., Vavricka, S.R., Bruegger, L.E. and Seibold, F. (2010a) 'Fecal calprotectin correlates more closely with the Simple Endoscopic Score for Crohn's disease (SES-CD) than CRP, blood leukocytes, and the CDAI', The American journal of gastroenterology, 105(1):162–169.

Schoepfer, A.M., Beglinger, C., Straumann, A., Trummler, M., Vavricka, S.R., Bruegger, L.E. and Seibold, F. (2010b) 'Fecal calprotectin correlates more closely with the Simple Endoscopic Score for Crohn's disease (SES-CD) than CRP, blood leukocytes, and the CDAI', The American journal of gastroenterology, 105(1):162–169.

Schridde, A., Bain, C.C., Mayer, J.U., Montgomery, J., Pollet, E., Denecke, B., Milling, S.W.F., Jenkins, S.J., Dalod, M., Henri, S., Malissen, B., Pabst, O. and McL Mowat, A. (2017) 'Tissue-Specific Differentiation of Colonic Macrophages

Requires TGFß Receptor Mediated Signalling', Mucosal immunology, 10(6), p. 1387–1399.

Schütze, S., Wiegmann, K., Machleidt, T. and Krönke, M. (1995) 'TNF-induced activation of NF-kappa B', Immunobiology, 193(2–4):193–203.

Senju, M., Hulstaert, F., Lowder, J. and Jewell, D.P. (1991) 'Flow cytometric analysis of peripheral blood lymphocytes in ulcerative colitis and Crohn's disease.', Gut, 32(7), p. 779–783.

Sheng, J., Ruedl, C. and Karjalainen, K. (2015) 'Most Tissue-Resident Macrophages Except Microglia Are Derived from Fetal Hematopoietic Stem Cells', Immunity, 43(2):382–393.

Shim, J.O. and Seo, J.K. (2014) 'Very early-onset inflammatory bowel disease (IBD) in infancy is a different disease entity from adult-onset IBD; one form of interleukin-10 receptor mutations', Journal of Human Genetics, 59(6):337–341.

Shouval, D.S., Biswas, A., Goettel, J.A., McCann, K., Conaway, E., Redhu, N.S., Mascanfroni, I.D., AlAdham, Z., Lavoie, S., et al. (2014) 'Interleukin-10 Receptor Signaling in Innate Immune Cells Regulates Mucosal Immune Tolerance and Anti-Inflammatory Macrophage Function', Immunity, 40(5):706–719.

Shrivastava, A.K., Singh, H. V., Raizada, A., Singh, S.K., Pandey, A., Singh, N., Yadav, D.S. and Sharma, H. (2015) 'Inflammatory markers in patients with rheumatoid arthritis', Allergologia et immunopathologia, 43(1):81–87.

Sidiq, T., Yoshihama, S., Downs, I. and Kobayashi, K.S. (2016) 'Nod2: A Critical Regulator of Ileal Microbiota and Crohn's Disease', Frontiers in Immunology, 7(SEP):367.

Sina, C., Kemper, C. and Derer, S. (2018) 'The intestinal complement system in inflammatory bowel disease: Shaping intestinal barrier function', Seminars in immunology, 37:66–73.

Skovdahl, H.K., Van Beelen Granlund, A., Østvik, A.E., Bruland, T., Bakke, I.,

Torp, S.H., Damas, J.K. and Sandvik, A.K. (2015) ‘Expression of CCL20 and Its Corresponding Receptor CCR6 Is Enhanced in Active Inflammatory Bowel Disease, and TLR3 Mediates CCL20 Expression in Colonic Epithelial Cells’, PLoS ONE, 10(11):e0141710.

Smids, C., Horje, C.S.H.T., Drylewicz, J., Roosenboom, B., Groenen, M.J.M., Koolwijk, E. van, van Lochem, E.G. and Wahab, P.J. (2018) ‘Intestinal T Cell Profiling in Inflammatory Bowel Disease: Linking T Cell Subsets to Disease Activity and Disease Course’, Journal of Crohn’s and Colitis, 12(4):465–475.

Smythies, L.E., Sellers, M., Clements, R.H., Mosteller-Barnum, M., Meng, G., Benjamin, W.H., Orenstein, J.M. and Smith, P.D. (2005) ‘Human intestinal macrophages display profound inflammatory anergy despite avid phagocytic and bacteriocidal activity’, The Journal of Clinical Investigation, 115(1):66–75.

Smythies, L.E., Shen, R., Bimczok, D., Novak, L., Clements, R.H., Eckhoff, D.E., Bouchard, P., George, M.D., Hu, W.K., Dandekar, S. and Smith, P.D. (2010) ‘Inflammation anergy in human intestinal macrophages is due to Smad-induced IkappaBalpha expression and NF-kappaB inactivation’, The Journal of biological chemistry, 285(25):19593–19604.

Solem, C.A., Loftus, E. V., Tremaine, W.J., Harmsen, W.S., Zinsmeister, A.R. and Sandborn, W.J. (2005) ‘Correlation of C-reactive protein with clinical, endoscopic, histologic, and radiographic activity in inflammatory bowel disease’, Inflammatory bowel diseases, 11(8):707–712.

Sorini, C., Cardoso, R.F., Gagliani, N. and Villablanca, E.J. (2018) ‘Commensal bacteria-specific CD4+T cell responses in health and disease’, Frontiers in Immunology. 9:2667.

Sowerwine, K.J., Holland, S.M. and Freeman, A.F. (2012) ‘Hyper-IgE syndrome update’, Annals of the New York Academy of Sciences, 1250:25-32.

Spits, H. (2002) ‘Development of αß T cells in the human thymus’, Nature Reviews Immunology, 2(10):760–772.

Staab, J.F., Lemme-Dumit, J.M., Latanich, R., Pasetti, M.F. and Zachos, N.C. (2020) 'Co-Culture System of Human Enteroids/Colonoids with Innate Immune Cells', Current protocols in immunology, 131(1):e113.

Stein, M., Keshav, S., Harris, N., Experimental, S.G.-T.J. of and 1992, U. (1992) 'Interleukin 4 potently enhances murine macrophage mannose receptor activity: a marker of alternative immunologic macrophage activation.', Journal of Experimental Medicine, 176(1):287–292.

Stout, R.D., Jiang, C., Matta, B., Tietzel, I., Watkins, S.K. and Suttles, J. (2005) 'Macrophages Sequentially Change Their Functional Phenotype in Response to Changes in Microenvironmental Influences', The Journal of Immunology, 175(1):342–349.

Svolos, V., Hansen, R., Nichols, B., Quince, C., Ijaz, U.Z., Papadopoulou, R.T., Edwards, C.A., Watson, D., Alghamdi, A., Brejnrod, A., Ansalone, C., Duncan, H., Gervais, L., Tayler, R., Salmond, J., Bolognini, D., Klopfleisch, R., Gaya, D.R., Milling, S., Russell, R.K. and Gerasimidis, K. (2018) 'Treatment of Active Crohn's Disease With an Ordinary Food-based Diet That Replicates Exclusive Enteral Nutrition Q5', Gastroenterology, 156(6):1354-1367.e6.

Swain, S. (1994) 'Generation and in vivo persistence of polarized Th1 and Th2 memory cells', Immunity, 1(7):543–552.

Swanson, K. V., Deng, M. and Ting, J.P.Y. (2019) 'The NLRP3 inflammasome: molecular activation and regulation to therapeutics', Nature Reviews Immunology, 19(8):477–489.

Sýkora, J., Pomahačová, R., Kreslová, M., Cvalínová, D., Štych, P. and Schwarz, J. (2018) 'Current global trends in the incidence of pediatric-onset inflammatory bowel disease', World Journal of Gastroenterology, 24(25):2741-2763.

Szabo, S.J., Kim, S.T., Costa, G.L., Zhang, X., Fathman, C.G. and Glimcher, L.H. (2000) 'A Novel Transcription Factor, T-bet, Directs Th1 Lineage Commitment', Cell, 100(6):655–669.

Tabaqchali, S., O'Donoghue, D.P. and Bettelheim, K.A. (1978) 'Escherichia coli antibodies in patients with inflammatory bowel disease.', Gut, 19(2):108–113.

Tai, Y., Wang, Q., Korner, H., Zhang, L. and Wei, W. (2018) 'Molecular mechanisms of T cells activation by dendritic cells in autoimmune diseases', Frontiers in Pharmacology, 9:642.

Takayama, T., Kamada, N., Chinen, H., Okamoto, S., Kitazume, M.T., Chang, J., Matuzaki, Y., Suzuki, S., Sugita, A., Koganei, K., Hisamatsu, T., Kanai, T. and Hibi, T. (2010) 'Imbalance of NKp44(+)NKp46(-) and NKp44(-)NKp46(+) natural killer cells in the intestinal mucosa of patients with Crohn's disease', Gastroenterology, 139(3):882-892.

Takeda, K. and Akira, S. (2005) 'Toll-like receptors in innate immunity', International Immunology, 17(1):1–14.

Targońska-Stępniak, B., Zwolak, R., Piotrowski, M., Grzechnik, K. and Majdan, M. (2020) 'The Relationship between Hematological Markers of Systemic Inflammation (Neutrophil-To-Lymphocyte, Platelet-To-Lymphocyte, Lymphocyte-To-Monocyte Ratios) and Ultrasound Disease Activity Parameters in Patients with Rheumatoid Arthritis', Journal of clinical medicine, 9(9):1–11.

Teixeira-Pinto, A. (2021) Machine Learning for Biostatics. Available at: https://bookdown.org/tpinto_home/Regularisation/best-subset-selection.html (Accessed: 20 July 2022).

Tester, A.M., Cox, J.H., Connor, A.R., Starr, A.E., Dean, R.A., Puente, X.S., López-Otín, C. and Overall, C.M. (2007) 'LPS Responsiveness and Neutrophil Chemotaxis In Vivo Require PMN MMP-8 Activity', PLOS ONE, 2(3):e312.

Therrien, A., Chapuy, L., Bsat, M., Rubio, M., Bernard, G., Arslanian, E., Orlicka, K., Weber, A., Panzini, B.P., Dorais, J., Bernard, E.J., Soucy, G., Bouin, M. and Sarfati, M. (2019b) 'Recruitment of activated neutrophils correlates with disease severity in adult Crohn's disease', Clinical and Experimental Immunology, 195(2):251–264.

Thomas, M.G., Bayliss, C., Bond, S., Dowling, F., Galea, J., Jairath, V., Lamb, C., Probert, C., Timperley-Preece, E., Watson, A., Whitehead, L., Williams, J.G., Parkes, M., Kaser, A. and Raine, T. (2019) 'Trial summary and protocol for a phase II randomised placebo-controlled double-blinded trial of Interleukin 1 blockade in Acute Severe Colitis: the IASO trial', BMJ Open, 9(2):e023765.

Thornton, A.M., Korty, P.E., Tran, D.Q., Wohlfert, E.A., Murray, P.E., Belkaid, Y. and Shevach, E.M. (2010) 'Expression of Helios, an Ikaros Transcription Factor Family Member, Differentiates Thymic-Derived from Peripherally Induced Foxp3+ T Regulatory Cells', The Journal of Immunology, 184(7):3433–3441.

Topchyan, P., Xin, G., Chen, Y., Zheng, S., Burns, R., Shen, J., Kasmani, M.Y., Kudek, M., Yang, N. and Cui, W. (2021) 'Harnessing the IL-21-BATF Pathway in the CD8+ T Cell Anti-Tumor Response', Cancers, 13(6):1–15.

Townsend, A.R.M., Gotch, F.M. and Davey, J. (1985) 'Cytotoxic T cells recognize fragments of the influenza nucleoprotein', Cell, 42(2):457–467.

Travassos, L.H., Carneiro, L.A.M., Ramjeet, M., Hussey, S., Kim, Y.G., Magalhes, J.G., Yuan, L., Soares, F., Chea, E., Le Bourhis, L., Boneca, I.G., Allaoui, A., Jones, N.L., Nũez, G., Girardin, S.E. and Philpott, D.J. (2010) 'Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry', Nature immunology, 11(1):55–62.

Tsianos, E. V., Katsanos, K.H. and Tsianos, V.E. (2012) 'Role of genetics in the diagnosis and prognosis of Crohn's disease', World Journal of Gastroenterology : WJG, 18(2):105-118.

Vasileva, D. and Badawi, A. (2019) 'C-reactive protein as a biomarker of severe H1N1 influenza', Inflammation Research, 68(1):39-46.

Vazeille, E., Buisson, A., Bringer, M.A., Goutte, M., Ouchchane, L., Hugot, J.P., de Vallée, A., Barnich, N., Bommelaer, G. and Darfeuille-Michaud, A. (2015) 'Monocyte-derived Macrophages from Crohn's Disease Patients Are Impaired in the Ability to Control Intracellular Adherent-Invasive Escherichia coli and Exhibit Disordered Cytokine Secretion Profile', Journal of Crohn's and Colitis, 9(5):410–

420.

Vermeire, S., Schreiber, S., Sandborn, W.J., Dubois, C. and Rutgeerts, P. (2010) ‘Correlation Between the Crohn’s Disease Activity and Harvey-Bradshaw Indices in Assessing Crohn’s Disease Severity’, Clinical Gastroenterology and Heterology, 8:357–363.

Verstockt, B., Mertens, E., Dreesen, E., Outtier, A., Noman, M., Tops, S., Schops, G., Van Assche, G., Vermeire, S., Gils, A. and Ferrante, M. (2020) ‘Influence of Drug Exposure on Vedolizumab-Induced Endoscopic Remission in Anti-Tumour Necrosis Factor [TNF] Naïve and Anti-TNF Exposed IBD Patients’, Journal of Crohn’s and Colitis, 14(3):332–341.

Verstockt, B., Verstockt, S., Blevi, H., Cleynen, I., De Bruyn, M., Van Assche, G., Vermeire, S. and Ferrante, M. ‘TREM-1, the ideal predictive biomarker for endoscopic healing in anti-TNF-treated Crohn’s disease patients?’, Gut, 68(8):1531–1533.

Verstockt, B., Verstockt, S., Dehairs, J., Ballet, V., Blevi, H., Wollants, W.-J., Breynaert, C., Van Assche, G., Vermeire, S. and Ferrante, M. ‘Low TREM1 expression in whole blood predicts anti-TNF response in inflammatory bowel disease.’, EBioMedicine, 40:733–742.

Verstockt, S., Verstockt, B., MacHiels, K., Vancamelbeke, M., Ferrante, M., Cleynen, I., De Hertogh, G. and Vermeire, S. (2021) ‘Oncostatin M Is a Biomarker of Diagnosis, Worse Disease Prognosis, and Therapeutic Nonresponse in Inflammatory Bowel Disease’, Inflammatory Bowel Diseases, 27(10):1564–1575.

Vieira, A., Fang, C. Bin, Rolim, E.G., Klug, W.A., Steinwurz, F., Rossini, L.G.B. and Candelria, P.A. (2009) ‘Inflammatory bowel disease activity assessed by fecal calprotectin and lactoferrin: correlation with laboratory parameters, clinical, endoscopic and histological indexes’, BMC Research Notes, 2:221.

Vietinghoff, S. von and Ley, K. (2008) ‘Homeostatic regulation of blood neutrophil counts’, Journal of immunology, 181(8):5183.

Volck, B. et al. (1998) ‘YKL-40, a mammalian member of the chitinase family, is a matrix protein of specific granules in human neutrophils.’, Proceedings of the Association of American Physicians, 110(4):351–360.

Voss, E., Wehkamp, J., Wehkamp, K., Stange, E.F., Schrö Der ‡, J.M. and Rgen Harder, J. (2005) ‘NOD2/CARD15 Mediates Induction of the Antimicrobial Peptide Human Beta-defensin-2 *’, Journal of Biological Chemistry, 281:2005–2011.

Vroom, T.M., Scholte, G., Ossendorp, F. and Borst, J. (1991) ‘Tissue distribution of human gamma delta T cells: no evidence for general epithelial tropism.’, Journal of Clinical Pathology, 44(12):1012–1017.

Wandall, J.H. (1985) ‘Function of exudative neutrophilic granulocytes in patients with Crohn’s disease or ulcerative colitis’, Scandinavian journal of gastroenterology, 20(9):1151–1156.

Wang, C., Baer, H.M., Gaya, D.R., Nibbs, R.J.B. and Milling, S. (2019) ‘Can molecular stratification improve the treatment of inflammatory bowel disease?’, Pharmacological Research, 148:10442.

Wang, C., Kang, S.G., Lee, J., Sun, Z. and Kim, C.H. (2009) ‘The roles of CCR6 in migration of Th17 cells and regulation of effector T-cell balance in the gut’, Mucosal Immunology, 2(2):173–183.

Wang, G.B., Li, C.R., Yang, J., Wen, P.Q. and Jia, S.L. (2011) ‘A regulatory polymorphism in promoter region of TNFR1 gene is associated with Kawasaki disease in Chinese individuals’, Human Immunology, 72(5):451–457.

Wang, Y., Van Boxel-Dezaire, A.H.H., Cheon, H., Yang, J. and Stark, G.R. (2013) ‘STAT3 activation in response to IL-6 is prolonged by the binding of IL-6 receptor to EGF receptor’, Proceedings of the National Academy of Sciences of the United States of America, 110(42):16975–16980.

Wang, Y., Yin, Y., Chen, X., Zhao, Y., Wu, Y., Li, Y., Wang, X., Chen, H. and Xiang, C. (2019) ‘Induction of Intestinal Th17 Cells by Flagellins From Segmented Filamentous Bacteria’, Frontiers in Immunology, 10:2750.

Watanabe, T., Kitani, A. and Strober, W. (2005) 'NOD2 regulation of Toll-like receptor responses and the pathogenesis of Crohn's disease', Gut, 54(11):1515.

Weersma, R.K., Crusius, J.B.A., Roberts, R.L., Koeleman, B.P.C., Palomino-Morales, R., Wolfkamp, S., Hollis-Moffatt, J.E., Festen, E.A.M., et al. (2010) 'Association of FcgR2a, but not FcgR3a, with inflammatory bowel diseases across three Caucasian populations†', Inflammatory Bowel Diseases, 16(12):2080–2089.

Weersma, R.K., Stokkers, P.C.F., van Bodegraven, A.A., van Hogezand, R.A., Verspaget, H.W., de Jong, D.J., van der Woude, C.J., Oldenburg, B., Linskens, R.K., Festen, E.A.M., van der Steege, G., Hommes, D.W., Crusius, J.B.A., Wijmenga, C., Nolte, I.M., Dijkstra, G. and Dutch Initiative on Crohn and Colitis (ICC) (2009) 'Molecular prediction of disease risk and severity in a large Dutch Crohn's disease cohort', Gut, 58(3):388–395.

Wei, L., Laurence, A., Elias, K.M. and O'Shea, J.J. (2007) 'IL-21 is produced by Th17 cells and drives IL-17 production in a STAT3-dependent manner', The Journal of biological chemistry, 282(48):34605–34610.

Weiser, M., Simon, J.M., Kochar, B., Tovar, A., Israel, J.W., Robinson, A., Gipson, G.R., Schaner, M.S., Herfarth, H.H., Sartor, R.B., McGovern, D.P.B., Rahbar, R., Sadiq, T.S., Koruda, M.J., Furey, T.S. and Sheikh, S.Z. (2016) 'Molecular classification of Crohn's disease reveals two clinically relevant subtypes', Gut, 67(1):36–42.

West, B.C., Rosenthal, A.S., Gelb, N.A. and Kimball, H.R. (1974) 'Separation and Characterization of Human Neutrophil Granules', The American Journal of Pathology, 77(1):41-66.

West, N.R., Hegazy, A.N., Owens, B.M.J., Bullers, S.J., Linggi, B., Buonocore, S., Coccia, M., Görtz, D., et al. (2017) 'Oncostatin M drives intestinal inflammation in mice and its abundance predicts response to tumor necrosis factor-neutralizing therapy in patients with inflammatory bowel disease', Nature medicine, 23(5):579-589.

Westmeier, J., Paniskaki, K., Karaköse, Z., Werner, T., Sutter, K., Dolff, S.,

Overbeck, M., Limmer, A., Liu, J., Zheng, X., Brenner, T., Berger, M.M., Witzke, O., Trilling, M., Lu, M., Yang, D., Babel, N., Westhoff, T., Dittmer, U. and Zelinskyy, G. (2020) 'Impaired Cytotoxic CD8 T Cell Response in Elderly COVID-19 Patients'. American Society for Microbiology, 11:e02243-20.

Wilks, S. (1859) 'Morbid appearances in the intestine of Miss Bankes', London Medical Gazette, (2):264.

Wolfson, D.M, Sachar, D. B., Cohen, A., Goldberg, J., Styczynski, R., Greenstein, A.J., I M Gelernt, H.D.J. (1982) 'Granulomas do not affect postoperative recurrence rates in Crohn's disease', *Gastroenterology*, 83(2):405–9.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X. and Guan, Y. (2021) 'clusterProfiler 4.0: A universal enrichment tool for interpreting omics data.', The Innovation, 2(3):100141.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X. and Yu, G. (2012) 'clusterProfiler: an R package for comparing biological themes among gene clusters.', OMICS: A Journal of Integrative Biology, 16(5), pp. 284–287.

Wu, Z., Kim, H.-P., Xue, H.-H., Liu, H., Zhao, K. and Leonard, W.J. (2005) 'Interleukin-21 receptor gene induction in human T cells is mediated by T-cell receptor-induced Sp1 activity', Molecular and cellular biology, 25(22):9741–9752.

Xu, M., Cen, M., Chen, X., Chen, H., Liu, X. and Cao, Q. (2019) 'Correlation between Serological Biomarkers and Disease Activity in Patients with Inflammatory Bowel Disease', BioMed Research International, 2019:651749.

Yamamoto, S., Shimizu, S., Kiyonaka, S., Takahashi, N., Wajima, T., Hara, Y., Negoro, T., Hiroi, T., Kiuchi, Y., Okada, T., Kaneko, S., Lange, I., Fleig, A., Penner, R., Nishi, M., Takeshima, H. and Mori, Y. (2008) 'TRPM2-mediated Ca2+ influx induces chemokine production in monocytes that aggravates inflammatory neutrophil infiltration', Nature Medicine, 14(7):738–747.

Yang, R., Mele, F., Worley, L., Langlais, D., Rosain, J., Benhsaien, I., Elarabi, H., Croft, C.A., et al. (2020) ‘Human T-bet Governs Innate and Innate-like Adaptive IFN-γ Immunity against Mycobacteria’, Cell, 183(7):1826–1847.e31.

Yang, X.O., Panopoulos, A.D., Nurieva, R., Seon, H.C., Wang, D., Watowich, S.S. and Dong, C. (2007) ‘STAT3 regulates cytokine-mediated generation of inflammatory helper T cells’, The Journal of biological chemistry, 282(13):9358–9363.

Yazar, S., Can, G., İzzet, A., Üniversitesi, B., Fakültesi, T., Dalı, G.B., Gürler, M., Can, H., Tezel, A., Ünsal, G., Soylu, A.R., Celalettin Ümit, H. and Kayhan, A. (2020a) ‘The relationship between hematological parameters and disease activity in inflammatory bowel disease’, Anatolian Current Meddical Journal, 2(3):68–74.

Yazar, S., Can, G., İzzet, A., Üniversitesi, B., Fakültesi, T., Dalı, G.B., Gürler, M., Can, H., Tezel, A., Ünsal, G., Soylu, A.R., Celalettin Ümit, H. and Kayhan, A. (2020b) ‘The relationship between hematological parameters and disease activity in inflammatory bowel disease’, Anatiolian Current Medical Journal, 2(3):68–74.

Yim, J.J., Kim, H.J., Kwon, O.J. and Koh, W.J. (2008) ‘Association between microsatellite polymorphisms in intron II of the human Toll-like receptor 2 gene and nontuberculous mycobacterial lung disease in a Korean population’, Human Immunology, 69(9):572–576.

Yim, J.J., Lee, H.W., Lee, H.S., Kim, Y.W., Han, S.K., Shim, Y.S. and Holland, S.M. (2006) ‘The association between microsatellite polymorphisms in intron II of the human Toll-like receptor 2 gene and tuberculosis among Koreans’, Genes and immunity, 7(2):150–155.

Yona, S., Kim, K.W., Wolf, Y., Mildner, A., Varol, D., Breker, M., Strauss-Ayali, D., Viukov, S., Guilliams, M., Misharin, A., Hume, D.A., Perlman, H., Malissen, B., Zelzer, E. and Jung, S. (2013) ‘Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis’, Immunity, 38(1):79–91.

Yoshimoto, T., Bendelac, A., Watson, C., Hu-Li, J., Paul, W.E., Barton, C.H., White, J.K., Roach, T.I., Blackwell, J.M., Ahmed, R. and Gray, D. (1996) 'Immunological memory and protective immunity: Understanding their relation', Science, 272(5258):54–60.

Yu G, Wang L, Han Y, H.Q. (2012) 'clusterProfiler: an R package for comparing biological themes among gene clusters', OMICS: A Journal of Integrative Biology, 16(5):284–287.

Yu, Q., Zhang, S., Wang, H., Zhang, Y., Feng, T., Chen, B., He, Y., Zeng, Z. and Chen, M. (2016) 'TNFAIP6 is a potential biomarker of disease activity in inflammatory bowel disease', Biomarkers in medicine, 10(5):473–483.

Zhang, S., Takaku, M., Zou, L., Gu, A., Chou WC., Zhang, G., Bing, W., Kong, Q., K, Thomas, S.Y., Serody, JS., Chen, X., Xu, X., Wade, P.A., Cool, D.N, Ting, J.P.Y and Wan, Y.Y. (2017) 'Reversing SKI-SMAD4-mediated suppression is essential for T_H17 cell differentiation', Nature, 551(7678):105-109.

Zhao, T., Su, Z., Li, Y., Zhang, X. and You, Q. (2020) 'Chitinase-3 like-protein-1 function and its role in diseases', Signal Transduction and Targeted Therapy, 5:201.

Zhao, W., Wang, P., Jia, H., Chen, M., Gu, X., Liu, M., Zhang, Z., Cheng, W. and Wu, Z. (2017) 'Lymphocyte count or percentage: which can better predict the prognosis of advanced cancer patients following palliative care?', BMC Cancer, 17:514.

Zheng, W.P. and Flavell, R.A. (1997) 'The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells', Cell, 89(4):587–596.

Zhou, Y., He, C.H., Herzog, E.L., Peng, X., Lee, C.M., Nguyen, T.H., Gulati, M., Gochuico, B.R., Gahl, W.A., Slade, M.L., Lee, C.G. and Elias, J.A. (2015) 'Chitinase 3-like-1 and its receptors in Hermansky-Pudlak syndrome-associate

Zhou, Z., Ding, M., Huang, L., Gilkeson, G., Lang, R. and Jiang, W. (2016) 'Toll-

like receptor-mediated immune responses in intestinal macrophages; implications for mucosal immunity and autoimmune diseases', Clinical immunology, 173:81.

Zhou, Z., He, H., Wang, K., Shi, X., Wang, Yupeng, Su, Y., Wang, Yao, Li, D., Liu, W., Zhang, Y., Shen, Lianjun, Han, W., Shen, Lin, Ding, J. and Shao, F. (2020) 'Granzyme A from cytotoxic lymphocytes cleaves GSDMB to trigger pyroptosis in target cells', Science, 368(6494).

Zhu, Y., Jiang, H.G., Chen, Z.H., Lu, B.H., Li, J. and Shen, X.N. (2020) 'Genetic association between IL23R rs11209026 and rs10889677 polymorphisms and risk of Crohn's disease and ulcerative colitis: evidence from 41 studies', Inflammation Research, 69(1):87–103.

Zingone, F., Barberio, B., Compostella, F., Girardin, G., D'Incà, R., Marinelli, C., Marsilio, I., Lorenzon, G. and Savarino, E.V. (2020) 'Good efficacy and safety of vedolizumab in Crohn's disease and ulcerative colitis in a real-world scenario', Therapeutic Advances in Gastroenterology, 13:1-9.

Zuo, S., Wang, B., Liu, J., Kong, D., Cui, H., Jia, Y., Wang, C., Xu, X., Chen, G., Wang, Y., Yang, L., Zhang, K., Ai, D., Du, J., Shen, Y. and Yu, Y. (2021) 'ER-anchored CRTH2 antagonizes collagen biosynthesis and organ fibrosis via binding LARP6', The EMBO journal, 40(16): e107403.