CHARACTERISATION OF THE FIBROTIC MICROENVIRONMENT ASSOCIATED WITH STRICTURES IN PATIENTS WITH CROHN'S DISEASE

By

SARAH AKIROR

A thesis submitted to the University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

Institute of Immunity and Immunotherapy

College of Medical and Dental Science

University of Birmingham

May 2023

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive
e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

Crohn's disease is an inflammatory driven condition in which 50% of patients develop fibrostenosing strictures in the course of disease progression, this remains a clinical unmet need. The study aimed to characterise the molecular and cellular fibrotic microenvironment in fibrotic (stricture) vs non- fibrotic (proximal and distal) segments derived from resected tissue of CD patients. Bulk RNA and single cell sequencing was performed on stricture, proximal and distal segments. Biomarkers of interest were identified to include Gremlin 1, HDAC1 and FGF2 whose expression patterns were studies in relation to localisation of fibroblasts using multispectral imaging. Studies on the role on Gremlin 1 in regulation of the TGF-β/BMP signalling pathways were conducted. The readouts included fibroblast activation, proliferation and expression of extracellular matrix (ECM) related genes including *COL1A1*, *ACTA2*, *FN1* and *TIMP1*. The data showed a distinct expression pattern associated with stricturing including upregulation of stromal cell linked genes, GREM1 expression was localised to the cell population identified as stromal cells. Knock down of *GREM1* using siRNA showed a downregulation in expression of proliferation associated gene *MKI67* and myofibroblast activation genes including *ACTA2*.

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisors Dr Christopher Weston and Dr Evaggelia Liaskou, for their patience, encouragement, and guidance. I will forever be grateful that you saw me through some of the most difficult times especially when carrying out lab work during the pandemic.

Thank you to Prof. Subrata Ghosh for the opportunity given to me to undertake this PhD project and all the members of the Ghosh lab including Dr Louisa Jeffery, Dr Stuart Hunter, Dr Uday Shivaji and Miss Jagjeet Sondh for their help with collection of resected samples, experimental design, planning and results interpretation. I am also indebted to Dr Animesh Acharjee and Dr Csilla Varnai. Much appreciation to members of the Weston lab in particular Marco, Paul and Celine and to everyone at the CLGR (Liver labs) for creating a great work environment.

I would like to thank my dear friends including Mohamed Zaakouk, Celine Hernandez, Thin Luu for the continued support and encouragement. Finally, to my family who have always believed in me much more than I believed in myself. Thank you for your unwavering love and support.

LIST OF FIGURES

Figure 1.1 Representative graph showing global trends in IBD cases……4
Figure 1.2 Schematic representation showing location and inflammation pattern of the affected areas in UC……10
Figure 1.3 Schematic representation showing location and inflammation pattern of the affected areas in CD……28
Figure 1.4 Schematic diagram showing cross-sectional structure of the small intestinal wall……...40
Figure 1.5 Representative image showing the process of endoscopic balloon dilation on an anastomosis stricture in a CD patient……..42
Figure 1.6 Cellular mechanisms that drive the fibrotic phenotype in intestinal fibrosis..60
Figure 1.7 Canonical TGF-β/SMAD signaling pathway……..63
Figure 2.1 Representative images showing location of sampled segments……..70
Figure 2.2 Distribution of gut resected tissue (proximal, stricture and distal) and blood for various experimental procedures……71
Figure 3.1 Schematic diagram that shows workflow involved in bulk RNAseq and scRNAseq processes……..109
Figure 3.2 An outline showing steps involved in bulk RNA-seq data analysis…….....112
Figure 3.3 Transcriptome analysis showed differential gene expression profiles in the stricture, proximal and distal segments of resected gut……114
Figure 3.4 Supervised PLS-DA of DEGs showed variation in gene expression between segments and within samples from the same segment……...116
Figure 3.5 Transcriptome analysis showed differential gene expression profiles in the stricture, proximal and distal segments of resected gut……..119
Figure 3.6 Hierarchical clustering of 81 genes to classify gene expression patterns in the stricture and proximal segments……121
Figure 3.7 Enriched term connections visualised in circular network……124
Figure 3.8 DEGs identified as significant (FDR adjusted $p<0.05$) for the segment pairwise comparisons……..……125-126
Figure 3.9 Validation of DEGs by qPCR……129
Figure 3.10 Figure shows calculations to determine sample size and power analysis…..131
Figure 3.11 Quantity of mitochondrial content being released by cells from each segment……...133
Figure 3.12 Multiple cell populations identified in both the stricture and proximal segments……135
Figure 3.13 UMAP plots showing gene expression patterns of markers that were used to define 14 clusters that represent individual cell populations……138
Figure 3.14 Even cell distribution across clusters……140
Figure 3.15 Violin plots show activation markers with increased expression across immune cell clusters……...142
Figure 3.16 Sub clustering of cells based on gene expression to identify stromal cells with a low transcription profile……..…145

Figure 3.17 Expression of target genes in scRNA-seq data from proximal and stricture segments. ……...147
Figure 3.18 GREM1 expression identified in intestinal stromal cells……...149
Figure 3.19 HDAC1 expression seen across sub-clustered cell populations in both the proximal and stricture segments…….150
Figure 3.20 FGF2 expression localised to clusters identified as endothelial/smooth muscle cell cluster and fibroblasts……152
Figure 4.1 Schematic diagram showing layers of the GIT……160
Figure 4.2 Histological appearance and structure of tissue in resected CD gut sections……165
Figure 4.3 Vimentin+ve cells identified in the proximal, stricture and distal segments……...168
Figure 4.4 Expression of α-SMA spread across the deeper layers of the gut in the stricture segment……...170
Figure 4.5 Expression of Desmin localised to the submucosa and muscularis layers in all the three segments……..173
Figure 4.6 Protein expression of stromal cell markers in the stricture segment……..175
Figure 4.7 Co-expression of fibroblast markers using immunofluorescence……..178
Figure 4.8 FGF2 protein expression in the proximal, stricture and distal segments of resected gut…….181
Figure 4.9 Expression of HDAC1 in the proximal, stricture and distal segments of resected gut……..…184
Figure 4.10 Comparison of two peroxidase blocking solutions for gremlin 1 IHC optimisation……187
Figure 4.11 Expression levels of gremlin 1 in the proximal, stricture and distal segments of resected gut…….189
Figure 5.1 Schematic representation of regulation of the TGF-β/BMP signalling…..…198
Figure 5.2 Isolation of primary fibroblasts from explanted liver tissue……..200
Figure 5.3 Organ culture method for intestinal fibroblast isolation……201
Figure 5.4 Isolation of intestinal fibroblasts using the tissue enzymatic digestion protocol……...203
Figure 5.5 Isolation of intestinal fibroblasts from gut resected tissue using the hybrid protocol……...205
Figure 5.6 Schematic summary of mesenchymal stromal cell identification markers…209
Figure 5.7 Human intestinal derived fibroblast cell lines in culture at 370C 5% CO2...211
Figure 5.8 Flow cytometry gating to show identification of fibroblast like phenotype in CCD-18Co…….213
Figure 5.9 CCD-18Co expressed α-SMA, PDGFRβ, CD90 and vimentin by immunofluorescence…….....215
Figure 5.10 CCD-18Co respond to profibrotic mediator stimulation with increased mRNA expression of ECM related genes……..217
Figure 5.11 CCD-18Co have a higher expression of ECM related genes at 48hrs upon stimulation with TGF-β……220
Figure 5.12 Increase in proliferation of CCD-18Co in the presence of profibrotic mediators…….223

Figure 5.13 : Presence of profibrotic mediators increase proliferation in primary fibroblasts from PSC explant liver……..225
Figure 5.14 CCD-18Co have similar expression patterns of BMP receptors at 24 and 48 hrs……...227
Figure 5.15 Unstimulated CCD-18Co express GREM1 at transcript level……...229
Figure 5.16 CCD-18Co express GREM1 in response to TGF-β stimulation at 24hrs…..230
Figure 5.17 CCD-18Co secrete gremlin-1 upon stimulation with profibrotic mediators..232
Figure 5.18 Preparation of GREM1-GFP tagged and control lentiviral particles…….234
Figure 5.19 Transduction efficiency of GREM1 lentivirus particles into CCD-18Co…..236
Figure 5.20 Gating strategy for sorting of CCD-GFP positive cells…….238
Figure 5.21 Assessing the knock down effect of siRNAs to suppress GREM1 expression when using lipofectamine as a transfection reagent……240
Figure 5.22 Assessing different siRNAs efficiency to suppress GREM1 expression when using HiPerFect as a transfection reagent……...242
Figure 5.23 Assessing siRNA1 efficiency at 20nM vs 60nM to suppress GREM1 expression……244
Figure 5.24 GREM1 mRNA expression was reduced in CCD-18Co at 72 hrs following transfection of 60 nM siRNA……246
Figure 5.25 GREM1 mRNA expression maintained 48 hrs post transfection…….248
Figure 5.26 TGF-β stimulation does not rescue GREM1 expression in transfected CCD-18Co 48hrs post knockdown…….249
Figure 5.27 GREM1 siRNA knockdown decreases ACTA2 expression in both unstimulated and TGF-β stimulated cells…….251
Figure 5.28 GREM1 siRNA knockdown decreases proliferation in CCD-18Co……253
Figure 6.1 Regulation of TGF-β/BMP signalling pathways……..267

LIST OF TABLES

Table 1.1 Differences between incidence and prevalence …….2
Table 1.2 Cells involved in intestinal fibrosis and their markers of identification……..59
Table 2.1 Summary of gut resected samples delivered to the lab for in vitro experiments……..72
Table 2.2 Summary of steps used during optimisation protocol for isolation of gut derived primary fibroblasts …….77
Table 2.3 Culture media and reagents used in cell isolation and culture……77
Table 2.4 List of antibodies used for flow cytometry staining for immune cell, cytokine and stromal cell panels ……79
Table 2.5 Reagents used for cDNA synthesis …….84
Table 2.6 Reagents used for qRT-PCR……85
Table 2.7 Protocol for qRT-PCR……..…85
Table 2.8 List of genes and primer sequences for targets used in qPCR……..…86
Table 2.9 Sequences of the four synthetic siRNAs targeting GREM1……94
Table 2.10 Reagents used for qRT-PCR……96
Table 2.11 Protocol for semi-quantitative PCR…….96
Table 2.12 List of antibodies used for IHC, IF and mIF……..101
Table 3.1 Patient demographics and characteristics……..107
Table 3.2 Showing concentration, absorbance and RIN values for RNA samples……111
Table 3.3 PLS model validation; model variance (R2) and predictive ability (Q2) of each of the comparisons…….118
Table 3.4 Table showing AUC for individual genes and a combined score……..128
Table 3.5 Quality control was performed on the cells for data analysis……134
Table 3.6 Genes selected for cluster identification of the different cell populations….137
Table 4.1 Common markers for identification of mesenchymal stromal cells……...166
Table 5.1 FCD cohort samples, experiments performed……207

ABBREVIATIONS

IBD	Inflammatory Bowel Disease
CD	Crohn’s disease
UC	Ulcerative Colitis
SMAD	Small mothers against decapentaplegic
BMP	Bone morphogenetic protein
TGF-β	Transforming growth factor-beta
α-SMA	α-smooth muscle actin
ECM	Extracellular matrix
EMT	Epithelial-to-mesenchymal transition
FN	Fibronectin
MSC	Mesenchymal stromal cell
TGF-β	Transforming growth factor
CD:	Cluster of differentiation
ANOVA	Analysis of variance
BSA	Bovine serum albumin
cDNA	Complimentary Deoxyribonucleic acid
Ct	Cycle threshold
DAPI	4′,6-diamidino-2-phenylindole
DMEM	Dulbecco's modified eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase I	Deoxyribonuclease I
IHC	Immunohistochemistry
FFPE	Formalin Fixed Paraffin Embedded sections
NOD2	Nucleotide binding oligomerization domain containing 2
ATG16L1	Autophagy related 16 like 1
IRGM	Immunity related GTPase M
CTGF	Connective tissue growth factor

EMT	Epithelial to mesenchymal transition
GWAS	Genome wide association study
HDAC	Histone deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IL	Interleukin
TIMP-1	Tissue inhibitor of metalloproteinase
MMP	Matrix metalloproteinase
PDGF	Platelet-derived growth factor
FGF	Fibroblast growth factor
FITC	Fluorescein isothiocyanate
KEGG	Kyoto encyclopedia of genes and genomes
GO	Gene ontology
AUC	Area under the curve
ROI	Region of interest
PCA	Principal component analysis
IPF	Idiopathic pulmonary fibrosis)
TNFα	Tumour necrosis factor alpha
Treg	T regulatory cell
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EMT	Epithelial mesenchymal transition
FBS	Fetal bovine serum
MTS	3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium)
OD	Optical density
PBS	Phosphate buffered saline
PSC	Primary Sclerosing Cholangitis
PCR	Polymerase chain reaction
PI	Propidium iodide

qPCR	Real time polymerase chain reaction
RNA	Ribonucleic acid
SD	Standard deviation
VEGF	Vascular endothelial growth factor
QC	Quality control
Limma	Linear models for microarray
Vs.	Versus
PECAM-1	Platelet endothelial cell adhesion molecule
ICAM-1	intercellular adhesion molecule 1
PBMCs:	Peripheral blood mononuclear cells
UMAP	Uniform manifold approximation and projection
HBSS	Hanks' Balanced Salt Solution
HEPES	4-2-hydroxyethyl-1-piperazineethanesulfonic acid
EMEM	Eagle’s minimum essential medium
TBE	Tris-Borate-EDTA buffer
VEO-IBD	very early onset of IBD
FOXP3	Forkhead box P3 (FOXP3)
IKBKG	Inhibitor of nuclear factor kappa B kinase regulatory subunit gamma

TABLE OF CONTENTS

CHAPTER 1 1
1.1 Inflammatory bowel disease 2
1.1.1 Aetiology of IBD 4
1.1.2 Extraintestinal manifestations in IBD 5
1.2 Ulcerative colitis 8
1.2.1 Aetiology of UC 10
1.3 Crohn's disease 26
1.3.1 Aetiology of CD 30
1.4 Intestinal fibrosis 38
1.4.1 Intestinal structure 39
1.4.2 Stricture formation in patients with Crohn's disease 41
1.4.3 Histological features of strictures 42
1.5 Cellular mechanisms involved in intestinal stricture formation 43
1.5.1 Immune cells regulate the inflammatory response that initiates stricture formation 44
1.5.2 Mesenchymal stromal cells 50
1.6 Molecular mediators that drive stricture formation in intestinal fibrosis 61
1.7 Gaps in current knowledge of intestinal fibrosis and the unmet need 64
1.8 Aims for the study; 67
CHAPTER 2 68
2.1 Sample collection, transportation and processing 69
2.1.1 Ethical approval for use of resected tissue samples 69
2.1.2 Sample collection and transportation to the lab 69
2.1.3 Sample receipt and processing in the lab 70
2.1.4 Resected gut tissue samples received and analysed in the whole study 71
2.2 Primary cell isolation 72
2.2.1 Isolation of Peripheral Blood Mononuclear Cells (PBMCs) from blood 72
2.2.2 Isolation of primary fibroblasts 72
2.2.3 Isolation of primary gut fibroblasts 73
2.3 Flow cytometry 78
2.3.1 Surface staining 78
2.3.2 Intracellular staining 78
2.4 In vitro cell culture assays 80

2.4.1 Cell lines used in this study 80
2.4.2 Gene expression assays 82
2.5 Bulk RNA sequencing 87
2.5.1 RNA extraction 87
2.5.2 RNA quality assessment 87
2.5.3 cDNA library preparation for sequencing 88
2.5.4 Analysis of bulk RNA-seq data 88
2.6 Single cell RNA sequencing 89
2.6.1 Isolation of single cells from resected gut samples for RNA sequencing 89
2.6.2 Enrichment of viable cells by Fluorescence Activated Cell Sorting (FACS) 90
2.6.3 10X Genomics Chromium single cell sequencing 90
2.6.4 Analysis of single cell sequencing data 91
2.7 CCD-18Co transfection with small interfering RNA (siRNA) 92
2.8 CCD-18Co lentiviral transduction 95
2.9 Semi- quantitative PCR 95
2.10 CCD-18Co proliferation using the MTS assay 96
2.11 Protein expression assays 97
2.11.1 ELISA 97
2.11.2 Haematoxylin and Eosin staining (H & E) 97
2.11.3 Immunohistochemistry 98
2.11.4 Immunofluorescence 100
2.12 Statistical analysis 102
2.12.1 Graphpad prism 102
CHAPTER 3 103
3.1 Introduction 104
3.2 Aims of this chapter: 105
3.3 Results 106
3.3.1 Patient demographic information and characteristics 106
3.3.2 Sequencing as a tool to study transcriptome profiles “the old versus the new” 108
3.3.3 Single cell RNA-seq to determine unique cell phenotypes in stricture and proximal tissues 131
3.3.4 Fibroblasts and biomarkers of fibrosis 146
3.4 Discussion 152
CHAPTER 4 157
4.1 Introduction 158

4.2 Aims of this chapter 162
4.3 Results 163
4.3.1 Structure and architecture of gut tissue 163
4.3.2 Histology of gut resected tissue 163
4.3.3 Fibroblasts identified in the proximal, stricture and distal segments of resected gut 165
4.3.4 Expression of vimentin in the proximal, stricture and distal segments 166
4.3.5 Expression of α-SMA in the proximal, stricture and distal segments 169
4.3.6 Expression of desmin in the proximal, stricture and distal segments 171
4.3.7 Expression patterns of FGF2, HDAC1 and Gremlin1 179
4.3.8 Optimisation of gremlin 1 detection by IHC staining 186
4.4 Discussion: 190
CHAPTER 5 193
5.1 Introduction 194
5.2 Aims of this chapter 197
5.3 Results 199
5.3.1 Isolation of primary gut fibroblasts 199
5.3.2 Characterisation of fibroblasts and myofibroblasts 208
5.3.3 Culture and expansion of colon fibroblast cell lines 210
5.3.4 Characterisation of CCD-18Co 212
5.3.5 Gremlin 1 and its role in BMP/ TGF-β signalling 228
5.3.6 Over expression of GREM1 in CCD-18Co using lentivirus 3-vector protocol 233
5.3.7 Optimisation of siRNA transfection to knock down GREM1 in CCD-18Co 239
5.4 Discussion 253
CHAPTER 6 260
6.1 Overview 261
6.2 A unique profibrotic signature in the stricture segment of CD resected samples 261
6.3 Localisation of stromal cells within the proximal, stricture and distal segments 263
6.4 Role of Gremlin 1 in stricture formation 264
6.5 Conclusions 265
6.6 Limitations of the study 268
6.7 Future work 269

CHAPTER 1

INTRODUCTION

1.1 Inflammatory bowel disease

Inflammatory Bowel Diseases (IBD) are chronic inflammatory disorders of the gastrointestinal tract that can lead to morbidity and mortality in severe cases (1,2). Epidemiological data suggests a global rise in prevalence and incidence of IBD cases, prevalence and incidence are epidemiological measurements used to determine disease occurrence and burden in a population (3). Prevalence is the number of individuals diagnosed with a particular disease in a population at a given time point or over a period of time, it includes both old and new cases. Incidence is the number of new cases of a disease within a population, this parameter measures the rate of occurrence (4). The differences between prevalence and incidence are outlined in table 1.1;

Table 1.1: Differences between incidence and prevalence (EUPATI, 2018)

Incidence	Prevalence
Measures the rate of occurrence of a disease in a population within a specific time period	Measures the proportion of a population found to be affected by a particular disease over a period of time
Only newly diagnosed cases are used to establish the incidence rates	Both new and old cases within a population are used to determine the prevalence
Denominator: Population at risk	Denominator: Whole population
Preferred measure for studying cause and effect	Preferred for estimating population based disease burden
More reliable in determining the risk of a specific disease in the population	Indicates how wide spread the disease is

The highest prevalence rates of IBD have been reported in Europe, North America and Oceania (5). In the past decade, there have been increasing cases in newly industrialised countries from Africa, Asia and South America, thus IBD is now considered an emerging global disease (6). An estimated 6.8 million people are living with IBD globally, with over 2 million of those in Europe and 1.5 million in North America (7,8). The global rise in IBD has psycho-social and

economic implications to society in terms of healthcare costs, absenteeism both at school and work, patients living with IBD also suffer from a poor quality of life. In Europe alone, an estimated direct healthcare cost per annum is 4.6–5.6 billion Euros (9,10).

The increase in IBD cases particularly in developing countries has been linked to a growing similarity in lifestyle and diets around the world (11). Figure 1.1 gives a global overview of disease progression in incidence and prevalence, the emergence of IBD in the western world was much earlier (1750-1950) compared to that of newly industrialised countries (1950-2000). This fuels the hypothesis that IBD is associated with factors linked to lifestyle such as diet and alteration in the intestinal microbiome that could trigger its onset. The concept of 'hygiene hypothesis' has been suggested as a contributing factor to higher prevalence in urban compared to rural areas (12), the effects of vitamin D may also be a factor as regions experiencing colder climates have a higher prevalence of IBD compared to those with a warmer climate (13). The overall trend in prevalence and incidence of IBD in developing countries exhibits a similar pattern to that previously seen in the western world during disease emergence stage (1750-1950); it is projected that between 2020-2050, there will be a sharp increase in IBD cases globally (14,15).

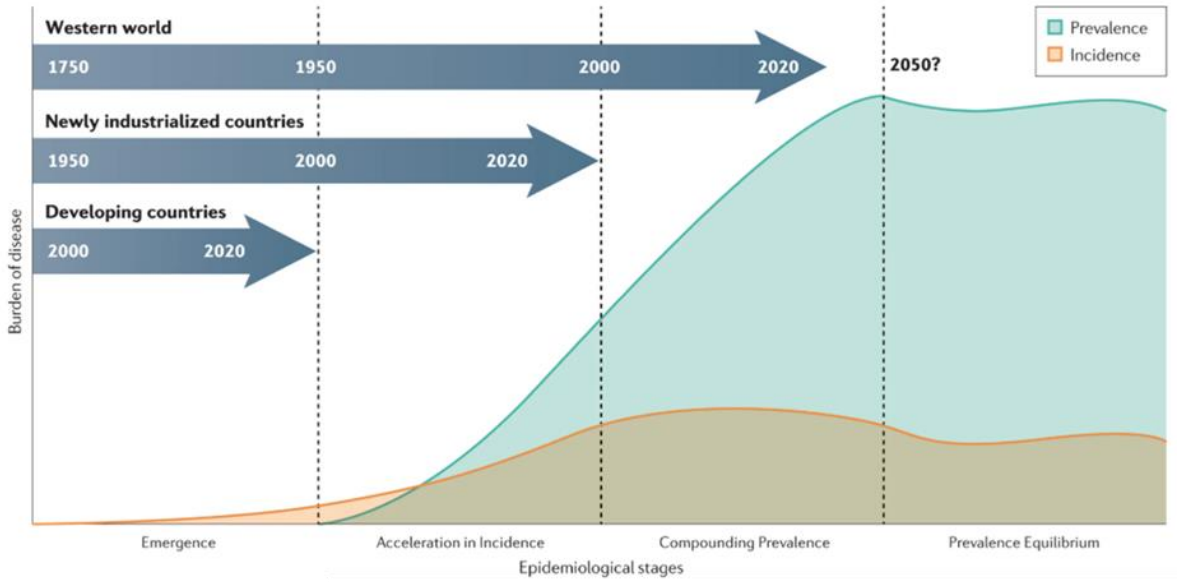


Figure 1.1: Representative graph showing global trends in IBD cases. Global IBD trends from 1750 to future projections in 2050. Data from 1750s show the early emergence of IBD in western countries (with high economies) to a projected global plateau in prevalence and incidence rates in 2050. Data from 2000-2020 shows developing countries going through an emergence stage with diagnosis of new cases, newly industrialised countries are in the acceleration stage where both prevalence and incidence are rising, the western world countries are at compounding stage with a steep rise in prevalence and steady incidence rates. (*Image obtained from Kaplan and Windsor, 2021)*

1.1.1 Aetiology of IBD

IBD constitutes of two main conditions, Crohn's disease (CD) and ulcerative colitis (UC). Disease classification is based on age of disease onset, disease location, pathological and clinical features (16,17). The aetiology of IBD is complex and not fully understood; studies have attributed it to a complex interaction of genetic susceptibility of the individual, epigenetic alterations, dysbiosis in the microbiome, environmental triggers and a dysregulated immune response (18). In a genetically susceptible host, an imbalance between commensal and pathogenic bacteria triggers a dysregulated immune response leading to chronic inflammation (19,20). Damage to the intestinal barrier impairs the mucosal epithelium through which pathogens, luminal antigens and toxins access the intercellular spaces within the tissue. This triggers activation of the innate immune response consisting of tissue resident cells such as

intestinal epithelial cells (IECs), fibroblasts, dendritic cells and macrophages to release pro-inflammatory mediators. Inflammatory mediators produced include cytokines, chemokines and growth factors such as TNF-α, IL-1β, IL-12, IL-6, IL-23, IL-10 and intercellular adhesion molecule 1 (ICAM-1), mucosal vascular addressin cell adhesion molecule 1 (MADCAM-1) (21,22). These in turn activate the adaptive immune response including cytotoxic T cells, helper T cells, regulatory T cells and B cells; the main immune mediators that further drive the inflammatory response to restore homeostasis include IL-12, IL-10, IL-17, IFN-γ, TGF-β, IL-23 and IL-4 (23,24). In a physiological state, this process is well regulated; however, in IBD the inflammatory response fails to "switch on" its regulatory mechanisms leading to chronic inflammation characterised in both UC and CD (25,26).

The advancement in high quality sequencing techniques has led to a high number of genetic studies such as the genome wide association studies (GWAS) that have progressed a better understanding into the role of genetics in the pathogenesis of IBD (27). Currently, GWAS have identified hundreds of susceptibility loci linked to IBD. Studies have also shown that genetic factors only account for 20–25% of disease variance (28), this suggests a key role of other risk factors including environment, epigenetic modulation, the microbiome and immune response in the pathogenesis of IBD (29,30). Approximately 25-40% of patients diagnosed with IBD also develop extraintestinal manifestations (EIMs) (31). Details on the proposed mechanisms of action of EIMs in IBD are covered in section 1.1.2.

1.1.2 Extraintestinal manifestations in IBD

Extraintestinal manifestations (EIMs) are inflammatory pathologies in patients with IBD that occur outside the gut, EIMs are seen in 25-40% of IBD patients (32–34). They manifest in various organs including the skin in the form of erythema nodosum and pyoderma gangrenosum, joint manifestations present as peripheral arthritis, eye manifestations include

episcleritis and uveitis while hepatobiliary EIMs manifest in the form of primary sclerosing cholangitis (PSC). EIMs greatly affect the patient's quality of life and complicate management of IBD (34,35).

The pathogenesis of EIMs is dependent on extension and translocation of immune responses or independent inflammatory events perpetuated by IBD immune responses, it may also result from shared environmental or genetic predisposition between the gastrointestinal tract and other affected organs (33,35) . The mechanisms through which EIMs occur in IBD are not clearly defined; a commonly reported hypothesis is ectopic expression of gut-specific chemokines and adhesion molecules including MAdCAM-1 and CCL-25. Expression of these molecules is restricted to intestinal tissue, their expression has been shown as upregulated in IBD patients. In addition, there is increased frequency of activated T cells expressing their cognate receptors α4β7 and CCR9 in IBD (31,36). Studies have shown ectopic expression of MAdCAM-1 and CCL-25 in the vascular endothelium of the portal tract in PSC; frequency of infiltrating activated T cells expressing α4β7 and CCR9 in explant livers of PSC patients has also been reported. The hypothesis has been supported in mice studies in which MAdCAM-1/ α4β7 signalling has been shown to exacerbate liver inflammation. This suggests an impaired recruitment of activated T cells to the liver as a contributing factor to the pathogenesis in PSC-IBD (37,38).

On the other hand, a recent study by Graham (2022) have refuted the "gut homing" hypothesis as not specific to PSC. The study showed that MAdCAM-1 and CCL-25 expression at mRNA and protein level was upregulated in tissue of other chronic liver diseases (CLDs) such as non-alcoholic steatohepatitis (NASH) and Alcoholic liver disease (ALD) at similar frequencies to PSC. The study showed similar expression patterns of their cognate receptors α4β7 and CCR9 in CLDs and PSC compared to donor livers used as healthy controls, the inconsistencies in the

literature have been attributed to investigations carried out with a focus on PSC-IBD with limited studies on the gut homing hypothesis in other CLDs (39).

Other mechanisms include T cell trafficking due to activation of non-specific adhesion molecules, an example is endothelial vascular adhesion protein 1 (VAP-1) which has been shown to recruit gut leukocytes in IBD patients to the synovial membrane which could be a contributing factor in the pathogenesis of peripheral arthritis (40). Microbial antigens derived from the gut microbiota or circulating antibodies reactive to colonic proteins have also been suggested to translocate to extraintestinal sites eliciting an impaired inflammatory response that could trigger onset of EIMs (33,34,36). Alternative mechanisms have been linked to shared genetic and environmental factors between IBD and other organs affected by EIMs. IBD is characterised by a reduction in gut permeability referred to as a "leaky gut", increased expression of pro-inflammatory mediators such as IL-6, TNF-α, IFN-γ and bacterial LPS (41). These pro-inflammatory mediators could promote further activation of immune cells in extraintestinal sites with shared genetic and environmental predisposing factors as reported in IBD; this could lead to a dysregulated immune response in extraintestinal sites.

Other co-morbidities in IBD patients include rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis, type 1 diabetes and psoriasis (42). GWAS have identified genetic links and shared susceptibility loci between IBD and other immune mediated diseases (43,44). Co-morbidities may have a strong association to either UC or CD; for example, studies have shown that PSC, a chronic disease of the liver characterised by inflammation and fibrosis of the biliary tree has a strong association to UC. Approximately, 14-21% of patients with UC will develop PSC; meanwhile, the prevalence of UC among PSC patients is approximately 70-80% (45). Therefore, despite both forms of IBD being inflammation driven, there are some differences in the pathophysiology, pattern of inflammation, disease progression and

management of the two conditions (46). Overall, progress has been made in understanding the mechanisms involved in the pathogenesis of EIMs and their links to IBD; however, further understanding of the pathogenesis and genetics of EIMs in relation to IBD will provide more efficacious treatment options, improved patient management and better outcomes of therapy (37).

1.2 Ulcerative colitis

UC is an immune-mediated disease of unknown aetiology characterised by chronic inflammation. UC affects individuals of any age; however, disease onset is predominant between the ages of 20-40 years with more females being diagnosed compared to males (47,48). There is evidence that indicates a rise in incidence and prevalence of very early onset of IBD (VEO-IBD) in recent years, this is in children as young as <6 years. Data shows a rise in cases from 1.3 to 2.1 per 100 000 children from 1994 to 2009 respectively; of the cases classified as VEO-IBD, 35%–59% are classified as UC (49). Studies have attributed the rise of VEO-IBD to changes in the environment and the microbiome which mostly occurs from birth to the age of 3-4 years. Interestingly, this coincides with the developmental stage of the child's immune system. A child with genetic predispositions undergoing changes to the microbiome and exposed to various environmental factors could trigger a dysregulated immune response resulting in VEO-IBD. In addition, sequencing has identified single gene (monogenic) defects in VEO-IBD patients linked to children with underlying primary immune deficiencies (50). Sequencing data on samples from VEO-IBD patients have revealed mutations on genes that are important in maintaining intestinal epithelial barrier function, phagocyte bacterial killing and immune regulation; examples include *IL10*, *IL10R*, a disintegrin and metalloprotease 17 (*ADAM17*), forkhead box P3 (*FOXP3*) and inhibitor of nuclear factor kappa B kinase regulatory subunit gamma (*IKBKG*). A loss in epithelial barrier function and a dysregulated immune response are contributing factors in the pathogenesis of VEO-IBD. Further research

and understanding into factors and mechanisms in VEO-IBD will improve management and identification of therapies for this cohort of patients (49–52).

A key differential diagnostic feature in IBD is the pattern of inflammation; in UC, inflammation affects the mucosal and submucosal regions of the colon. In 95% of patients diagnosed with UC, inflammation starts from the rectum and extends to the proximal region of the colon in a continuous pattern (53,54). Clinical examination and histopathological findings on tissue from patients diagnosed with UC show inflamed regions that are ulcerated with diffuse lesions beginning in the rectum and extending proximally without skip areas. The inflamed regions appear edematous, granular and vulnerable as illustrated in figure 1.2 (55,56). Biopsies from patients with UC show distorted tissue integrity and structure, infiltration of inflammatory cells such as neutrophils and the presence of crypt abscesses (57–59).

The signs and symptoms experienced by most UC patients include bloody diarrhoea, faecal urgency, weight loss, rectal pain and bleeding (54,60). Upon diagnosis, the course of treatment aims at resolving the inflammatory response and restore mucosal healing. Current first line medication includes 5-aminosalicylic acid (5-ASA) agents, corticosteroids, thiopurines, biologics such as anti-tumour necrosis factor (anti-TNF-α) and small molecules such as Janus kinase (JAK) inhibitors; the choice of medication depends on the severity of disease. Total proctocolectomy is regarded as the last resort and only if other treatment options fail to manage UC symptoms (61).

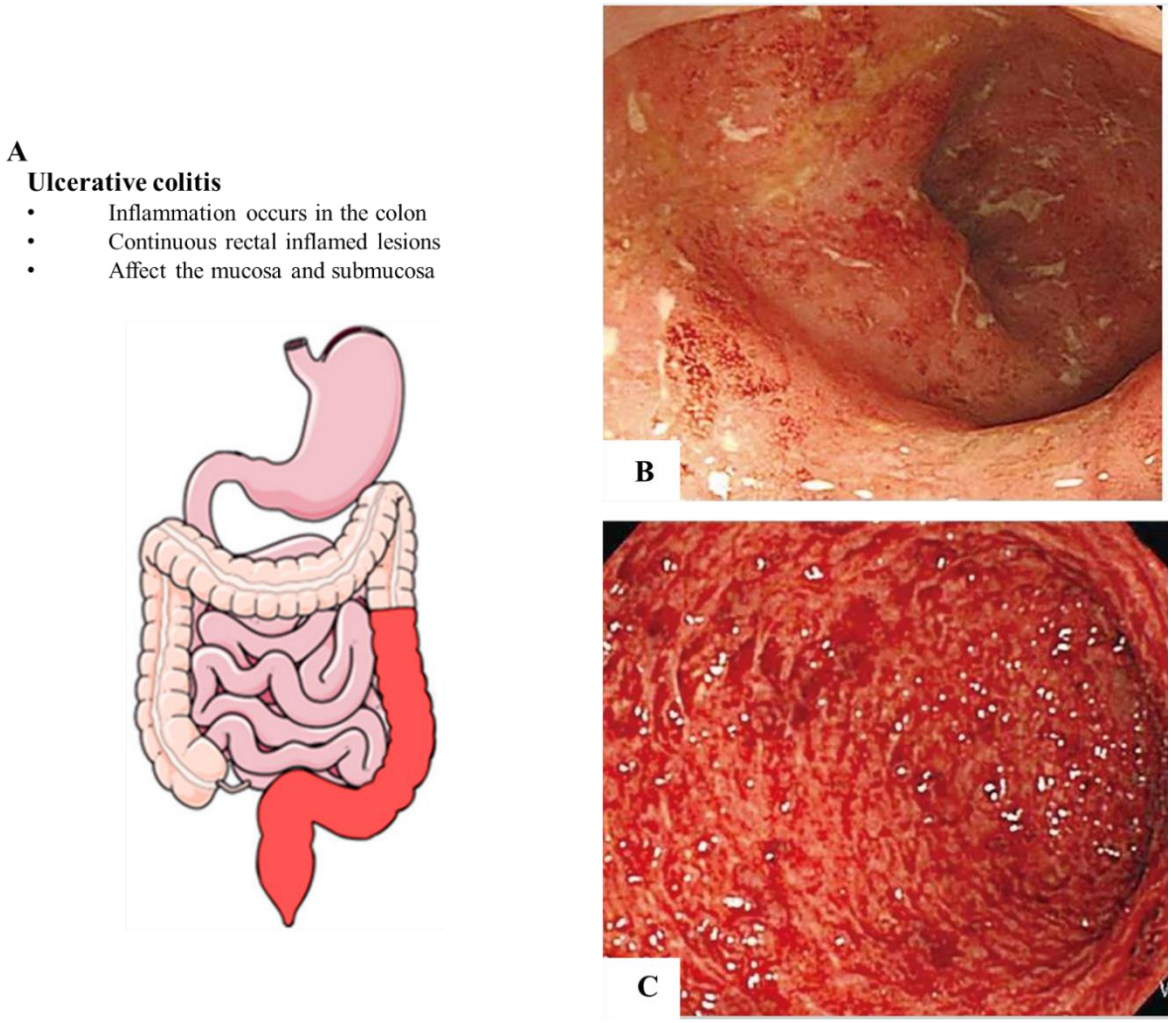


Figure 1.2: Schematic representation showing location and inflammation pattern of the affected areas in UC. **A)** Inflammation in the colon that starts in the rectum and extends proximally (red-coloured region in gastrointestinal tract). **B)** Endoscopic features of UC showing moderate inflammation with marked erythema, loss of vascular marking and erosions. **C)** Severe inflammation with spontaneous bleeding. *Images obtained from Lee and Lee, 2016; Hoter and Naim, 2019.*

1.2.1 Aetiology of UC

1.2.1.1 Genes linked to epithelial barrier function impaired in UC

The aetiology of UC is multifactorial, various studies have linked genetics as one of the key factors involved in its pathogenesis (62,63). Familial studies have shown that the prevalence of first-degree relatives developing UC is 0.3-2.7%. Of the patients diagnosed with UC, data

showed that 8-14% had a family history of IBD. Twin studies have also shown concordance rates of 16% in monozygotic twins and 4% in dizygotic twins (64–66). Genome wide association studies (GWAS) have generated data that strongly links genetics to the pathogenesis of UC. To date, a total of 240 risk loci have been identified and linked to the pathogenesis of IBD (67–70).

Of the identified 240 susceptibility genes, 47 have been associated with UC (71,72). UC specific susceptibility mutations have been identified on genes linked to human leukocyte antigen (HLA class II) region near HLA-DRA with a strong association to *HLA-DR-B1*15*, *DRB1*15:01*, *DRB1*15:02* and *DRB1*15:03*, the B1 genes contain the β chain and are the most polymorphic. It is not clear why these loci are predisposed in UC; however, HLA class II genes modulate both innate and adaptive immune responses through recognition of microbes, lymphocyte activation, cytokine signalling and maintenance of intestinal epithelial barrier function (73). Mutations on genes that are important in these biological processes increase the risk of developing UC. For example, susceptibility loci were identified on genes that modulate epithelial organisation and barrier function, these include *HNF4A*, *CDH1* and *LAMB1* (74,75). GWAS have also identified susceptibility loci that are linked to increased risk of developing severe UC, these include *TNFSF15* (*TL1A*), *MDR1*, *HSP70* and loci on genes associated within the HLA-DRA regions (76–78). GWAS have identified susceptibility genes with an overlap between UC and CD, these genes play a key modulatory role in innate and adaptive immune responses. These include autophagy genes (*PARK7* and *DAP*), oxidative stress (*HSPA6*, *DLD*), immune tolerance (*IL1R1*, *IL1R2*, *IL10*) and T cell activation and regulation (*IFNG*, *JAK2*, *TNFSF8*, *TNFSF15*, *ICOSLG*, *IL12B*, *IL23R*, *IL17REL* (71,79–81).

Another important factor in the pathogenesis of UC is the dysregulation in recognition of pathogens, Caspase recruitment domain-containing protein 9 (*CARD9*) is a member of the

CARD protein family that acts as an adaptor protein (82). *CARD9* is expressed in myeloid immune cells, including dendritic cells, macrophages, and neutrophils that are important in antigen recognition. Cells of the innate immune system recognise antigens through pattern recognition receptors (PRRs) such as C-type lectin receptor (CLR) and toll-like receptor (TLR). These receptors couple with signaling adaptor proteins such as *CARD9* to activate their functions and initiate defence against invading pathogens (83,84). GWAS have identified polymorphisms on *CARD9*, protein interaction domains on *CARD9* perform a dual function of activating or suppressing other *CARD*-containing proteins, hence driving the pathogenesis of IBD. There is a significant increase in gene expression of *CARD9* in the UC group compared with CD ($P < 0.001$). The data showed a strong disease association of *CARD9* to UC and a moderate association with CD (85–87).

The study identified high genetic risk alleles in IBD, these included rs10870077, rs4077515 and rs10781499 while the protective alleles included rs141992399 and rs200735402. The dual role has been attributed to *CARD9* variants having different mechanisms of pathogenesis that contribute to the differences in disease susceptibilities; the predisposing risk variants were associated with impairment in recognition of pathogens that results in a dysregulated immune response in a genetically predisposed host. The protective variants may have lost the biological function of *CARD9;* a study was conducted using a total and conditional *CARD9* knock-out in a dextran sulphate sodium (DSS) induced mouse model of colitis and an adoptive transfer model of colitis to assess the cell type that plays a key role in the *CARD9* protective phenotype (88). The data showed a protective role in *CARD9* expressing neutrophils; in the absence of *CARD9*, mitochondrial dysfunction increases mitochondrial reactive oxygen species (ROS) production leading to a premature death of neutrophils through apoptosis. The reduction in functional neutrophils in tissues explains an increased susceptibility to intestinal inflammation. This presents opportunities to target neutrophils as potential therapeutic targets in IBD (88,89).

Overall, the current data from various studies indicates that genetics is a key factor in the pathogenesis of UC.

1.2.1.2 Epigenetic changes can impair intestinal barrier integrity in UC

Epigenetic mechanisms are reversible changes that occur causing an alteration in the structure and function of DNA without changing the DNA sequence. Alteration to the DNA may influence gene transcription and alter gene expression hence impacting disease progression and overall outcome (90,91). The most common epigenetic mechanisms include DNA methylation in which a methyl group(s) is added to a DNA molecule, histone modification which is a post transcriptional modification (PTM) which exerts alterations to histone proteins and changes due to non-coding RNA associated gene silencing (ncRNAs) (92). In the last decade, there have been significant advancement in understanding epigenetics and its role in disease through epigenome-wide association studies (EWAS). In UC, epigenetic modifications have been identified in genes linked to the intestinal barrier function (93,94). The intestinal barrier is made up of a thick internal lining that consists of specialised intestinal epithelial cells (IECs). These include enterocytes, paneth cells, goblet cells, endocytes and microfold cells; each with a distinct function to maintain epithelial homeostasis. Paneth cells secrete anti-microbial peptides, such as α-defensin, to limit microbial entry to the intestinal lumen. Goblet cells secrete mucous that is key in defence and repair of the epithelium (95–97) while endocytes regulate incoming antigens and present bacterial antigens to dendritic cells. These cells establish a tight barrier using adherens junctions and desmosomes, hence maintaining the intestinal barrier integrity (96).

A study by Saito (2011) on colon biopsies showed hypermethylation of loci on CDH1 and GDNF in colon mucosa of active UC vs inactive. GDNF encodes for Glial cell-derived neurotrophic factor, the protein plays a key role in promoting maturation of intestinal epithelial

cells (IECs) (99). *CDH1* encodes for E-cadherin, E-cadherin is an intercellular adhesive protein and a component of adherens junctions found within IECs. The altered gene expression of CDH1 and GDNF due to hypermethylation results in a reduced expression of E-cadherin or Glial cell-derived neurotrophic factor in the small intestine and colon resulting in an impaired intestinal homeostasis and barrier function (100,101).

In another study, next-generation sequencing was used on treatment-naïve UC mucosal biopsies, the use of treatment naïve patients was to rule out the effect of medication on DNA modification. The study showed that genes such as proline-rich acidic protein 1 (*PRAP1*) and a member of the solute carrier protein family (*SLC3A1*) were hypermethylated with a significantly reduced expression in UC mucosa (102). Both genes encode for proteins that maintain homeostasis in epithelial cell and protect the intestinal barrier from oxidative insult (103,104).

Other important genes include hyperpigmentation progressive 1 (*HPP1*) and myogenic differentiation 1 (*MYOD1*) which showed a hypermethylation profile in UC patient samples with actively inflamed mucosa compared to quiescent colon derived samples, this indicates a strong link between DNA modification and disease phenotype in UC (98). Studies have also shown hypermethylation in promoter sites of genes such as multidrug resistance 1(*MDR1*), myogenic differentiation 1 (*MYOD1*), *CXCL5*, *IL12RB2*, *IL2RA*, intercellular adhesion molecule 1 (*ICAM)*, *CD40*, tumor necrosis factor receptor superfamily members (*TNFRSF4* and *TNFRSF5*); DNA modifications leads to changes in gene expression and cellular processes such as proliferation, intracellular signalling, apoptosis and DNA repair; which in turn drives the pathogenesis of UC affecting clinical phenotype, reduced age for disease on set (<20 years old), higher number of hospitalisations and severe disease outcomes (98,105–107).

Another study explored the role of hypomethylation in tissue collected from treatment-naïve severe UC patients (108). The data showed hypomethylation in some differentially expressed genes (DEGs) in patients with severe UC. These included genes such as *SIGLEC5*, *NKG7*, *ITGB2*, *GNAI2*, *TNFSF8* and *IL10*; which are typically involved in anti-inflammatory responses. The study suggests that hypomethylation of genes that are important in anti-inflammatory responses impairs the immune system's regulatory processes in resolving inflammation which is typical in severe UC patients; the impaired immune mechanisms are further discussed in section 1.2.1.5. Hypomethylated genes specific to severe UC could also be used as biomarkers of disease severity, providing a more targeted treatment approach (108).

Non-coding RNAs (ncRNAs) such as microRNAs (miRNAs) have also become of interest as modulators of the epigenome. These are small, endogenous molecules that target mRNA at the post-transcription phase causing it to degrade (109). A study by Min (2014) showed an upregulation of miR-155 in mucosa of UC patients with active disease corresponding to a decrease in expression of *FOXO3*, a gene that encodes Forkhead box O3 protein (FOXO3a). FOXO3a is an important transcription factor that drives multiple homeostatic processes such as apoptosis, immune response and growth. Intestinal inflammation was more severe in a *FOXO3a* knockout mice (Foxo3a −/−), a DSS induced mouse model of colitis as compared to the wild type This indicates a regulatory role for FOXO3a during the inflammatory response in UC (111,112).

Further investigation in the HT29 cell line showed that upregulation of miR-155 has a negative effect on expression of FOXO3a and an increase in expression of pro-inflammatory mediators including IL-8 and TNF-α (110). Although various miRNA such as miR-103-2, miR-362-3p, miR-532-3p, miR-20b, miR-98 and miR-125b-1 have been shown as upregulated in the colonic mucosa of UC patients, there is still need for further research to identify what mechanisms they

regulate and how best these can be targeted as potential therapies in the management of UC (109,113).

1.2.1.3 Environmental factors such as smoking and appendectomy are protective in UC

Data has shown an increase in prevalence and stable incidence of IBD cases in economically developed countries whereas there is an increase in developing countries. The rise in prevalence is partly associated with IBD being a chronic condition with low mortality rates (5,15). These trends led to further investigations on the role of environment in the pathogenesis of IBD. Environmental risk factors such as western diet and lifestyle that were associated with developed countries had been implicated in the pathogenesis of IBD. However, the world is fast becoming uniform in human behavioural patterns resulting in developing countries adapting to similar diets and lifestyles as in developed countries (114). This trend coincides with the increase in prevalence and incidence of IBD cases in developing countries in the last 20 years (14).

One of the main environmental risk factors in IBD is cigarette smoking, smoking is protective in UC with low incidence rates of UC in smokers. Studies have also shown an overall positive disease outcome in UC patients that start smoking after diagnosis (115). The protective mechanisms have been attributed to nicotine and its metabolites, such as cotinine that mediate immunomodulatory effects, this includes activation of nicotinic receptors α7 (α7-nAChR) on macrophages and dendritic cells. Activation of α7-nAChR receptor exerts immunosuppressive function of T regulatory cells and reduces activation of NF-κB and IL-2 production (116).

Studies have also shown expression of α7-nAChR receptors on endothelial cells whose activation results in a reduced production of chemokines and adhesion molecules that would otherwise modulate leukocyte recruitment (115,117,118). The mechanisms highlighted reduce chronic inflammation, the protective effect is specific to UC due to differences in expression

of receptors and molecules on immune cells residing in different parts of the gastrointestinal tract (GIT) affected in UC or CD. For example, aryl hydrocarbon receptor (AhR) stimulation by dioxin from tobacco combustion has anti-inflammatory and protective properties, AhR expression is upregulated in inflamed UC as compared to inflamed CD tissue (119). Although smoking has been shown as protective in UC, it has been reported as a high-risk factor in conditions such as rheumatoid arthritis, Crohn's disease, and cardiovascular disease among others. The deleterious effects of smoking outweigh its benefits in UC, it is therefore good clinical practice to recommend that UC patients stop smoking as part behavioural change to minimise co-morbidities (120–122).

Another environmental risk factor in UC is appendectomy, a retrospective study in a Swedish cohort of UC patients from 1964-2010 compiled by Myrelid (2017) showed a lower risk of colectomy, reduced frequencies of hospital admissions and/or a mild disease phenotype in patients that undergo appendectomy before the age of 20 years prior to their UC diagnosis; this was reported at an odds ratio ([OR], 0.44; 95% confidence interval [CI], 0.30-0.64; $P < .0001$). The protective effect was seen in patients that had appendectomy surgery due to inflammation driven conditions such as appendicitis or mesenteric lymphadenitis. There was loss in protection for UC patients that had appendectomy surgery above 20 years of age and those that received appendectomy for non-inflammation related conditions (123).

On the contrary, patients with established UC undergoing appendectomy because of appendicitis showed a worse disease course with an increased rate of subsequent colectomy (124,125). The data suggests an immunomodulatory role of the appendix; studies have shown abnormal appendicular infiltration of CD4+ T-cell in UC patients with active and inactive disease. The suggested mechanisms involve the appendix acting as a priming site for CD4+ T cells preceding to the development of UC in genetically susceptible individuals (126,127).

Acute appendicitis may activate pro and anti-inflammatory pathways that are ready to resolve any triggers that may initiate UC associated inflammation leading to protection against developing UC or a mild disease phenotype. In a murine model, there is increased expression of T regulatory cells (CD25+FoxP3+) in appendicitis although this is age limited occurring mostly in younger animals. The data aligns with previous studies in UC patients that showed loss in protection from development of UC when appendectomy is performed after 20 years of age (123,127). A comprehensive systematic review and meta-analysis by Welsh (2022) have also shown that appendectomy in established UC was associated with higher rates of subsequent colorectal cancer (CRC) and high-grade dysplasia (HGD). Overall, the role of the appendix in UC is complex and multifactorial; further studies are required to understand the role of appendicectomy in left-sided treatment-refractory UC (128).

Diet is another risk factor of interest with consumption of high sugar, soft drinks and lack of vegetables showing an increased risk of developing UC, the data in this case is not robust enough to draw conclusions (129). The hygiene hypothesis has also been proposed as a risk factor for both UC and CD, with people living in rural locations and developing countries shown to have a low risk of developing IBD as compared to those in urban and industrialised geographical locations (130,131). Trends show a rise in incidence of IBD particularly in developing countries coinciding with improvement in hygiene over the 20th century; these include access to clean water, a hot water tap, less crowding due to smaller family sizes and consumption of non-contaminated food (132). The hygiene hypothesis is based on the premise that improved hygiene over protects a child from exposure to common environmental infectious agents at an early age. If the child is exposed to pathogenic infectious agent later in life, the immune system is not primed to respond to these pathogens hence mounting a dysregulated immune response that is characteristic of IBD (133,134).

A counter argument has been made as data from the World Health Organisation shows very poor availability of data on basic demographic information, causes of death and their registration especially in developing countries. The rise in incidence rates of IBD in developing countries has been attributed to improved public infrastructure, provision of proper diagnostic tools and improved reporting of cases for conditions such as IBD. This has raised doubts on the role of the hygiene hypothesis in the pathogenesis of IBD (135,136). Other risk factors include use of antibiotics, breast feeding, nonsteroidal anti-inflammatory drugs and air pollution; however, the data is inconsistent and inconclusive. Further research is required to establish the actual risk of developing IBD upon exposure to these factors (137).

1.2.1.4 Intestinal microbial dysbiosis is a key mediator in the pathogenesis of UC

The human GIT is composed of commensal microbes such as bacteria, fungi, archaea and viruses located on mucosal surfaces (138). The early childhood microbiome is a key modulator of health and development, host-microbiome cross-talks influence metabolic and immunological homeostasis of the host through facilitating nutrient uptake, promoting immune maturation, development of tolerance and clearance of pathogens (139,140). The microbiome undergoes multiple changes until it stabilises at the age of 3-4 years (141). Several factors contribute to early childhood alterations in the microbiome; they include mode of delivery, breastfeeding, early exposure to antibiotics, presence of siblings, family members, pets, genetics and the local environment (142).

During delivery, microbiota is transferred from mother to child; babies born through natural means will contain microbes common to the mother's birth canal such as Lactobacillus, Prevotella, and Sneathia while the microbiome of babies born through caesarean section is colonised by skin microbes, for example Staphylococcus, Propionibacterium, and Corynebacterium (143–145). Breast feeding is another contributing factor to early childhood

microbiome development, there is a clear difference in the microbiota of babies that are breastfed and those that are formula-fed. Babies that are breastfed have a microbiome that is predominant of Bifidobacterium and Lactobacillus spp which originate predominantly from the milk ducts, mothers' nipple and surrounding skin while those that are formula fed tend to have higher proportion of Clostridiales and Proteobacteria which originate from the equipment and water used for suspension of the formula milk (146,147).

Other factors include diet and the use of antibiotics, exposure to antibiotics alters the microbiome but particularly in neonates, a systematic review was performed on 48 studies where findings showed reduced gut microbial diversity associated with prolonged antibiotic treatment. Studies also showed a reduction in colonisation rates of protective commensal anaerobic bacteria and an increased risk of developing antibiotic resistance (148). Overall, the data presents a strong argument for minimal use of antibiotics in neonates as the alternative could lead to alteration of the microbiome into disease-promoting gut microbiota (149).

Host genetics has also been discussed as a potential contributing factor, studies have shown similarity in microbiota of monozygotic twins under the age of 10 years as compared to dizygotic twins. However, due to the complex nature of genetic factors, more studies are required to understand the association between genetics and infant microbiome (150). In a state of homeostasis, the microbiome is predominantly made up of probiotics; these are non-pathogenic bacteria such as Bifidobacterium, Bacteroides and Peptococcus living in the gut mucosa in a symbiotic relationship with the host and are important in immune regulatory mechanisms (151).These organisms establish a colony in the intestine through a process called colonisation resistance that resists new bacterial growth that might be harmful (152). Other inhabitants of the gut mucosa are the facultative anaerobes such as Enterococcus and Enterobacter, they inhabit the host and are harmless when intestinal

homeostasis is maintained; however, they are opportunistic pathogens in dysbiosis (153). Increase in pathogenic bacteria populations in the gut leads to increased levels of enterotoxins that weaken the intestinal epithelium to what is known as a “leaky gut”. Luminal antigens can then easily gain access past the damaged epithelial barrier prompting activation of an immune response (154).

Intestinal microbes perform immunomodulatory functions through microbial metabolites of dietary substrates such as bile acids (BAs) and short-chain fatty acids (SCFAs) (155). BAs are metabolites derived from cholesterol metabolism, they perform a critical role in facilitating digestion and absorption of lipids from the diet but also perform a role in regulating intestinal mucosal immune responses (156). Primary bile acids (PBAs) such as cholic acid (CA) and chenodeoxycholic acid (CDCA) are synthesised in the liver by hydroxylation catalysed by cytochrome P450 enzyme cholesterol 7α-hydroxylase (CYP7A1) or an alternative pathway (CYP27A1); these are conjugated to glycine or taurine to form conjugated BAs. Conjugated BAs are transported to the intestine where they are metabolised and converted into secondary bile acids (SBAs), for example deoxycholic acid (DCA) and lithocholic acid (LCA) by gut microbes such as Clostridium and Eubacterium species, the SBAs then get excreted in faeces (157–159).

BA metabolism is a well-regulated process that is necessary in maintaining gut homeostasis; however, gut dysbiosis in IBD results in an imbalanced BA metabolism, multiple studies in cohorts that include UC, CD and paediatric IBD patients have shown a defect in BA metabolism showing increase in PBAs and/or their conjugated forms and a reduction in SBAs (160,161). In addition, a study by Sinha (2020) showed a significant reduction in SBAs in the ileal pouches of colectomy-treated UC patients compared to pouches of colectomy-treated Familial Adenomatous Polyposis (FAP) patients (162). The shift in balance between PBA and

SBA may results in pro-inflammatory intestinal mucosal responses through differential effects on epithelial and immune cells. BAs act on BA-activated receptors which include farnesoid X receptor (FXR), Takeda G-protein receptor 5 (TGR5), pregnane X receptor (PXR) and vitamin D receptor (VDR). This are highly expressed on epithelial cells of the gut as well as immune cells including type 3 innate lymphoid cells (ILCs), Th17 cells, dendritic cells (DCs), natural killer cells (NKCs), and monocytes/macrophages (156,163,164).

SCFAs such as butyrate, propionate, and acetate are microbial metabolites produced when proteolytic enzymes digest polysaccharides; SCFAs play a key role in gut homeostasis through anti-inflammatory, anti-microbial properties and modulating intestinal barrier function (165). Butyrate promotes homeostasis within the colonic mucosa by stimulation of mucous secretion, increase of vascular flow, absorption of sodium and water. Butyrate signals through G-protein coupled receptors GPR41 and GPR109A expressed on intestinal epithelial cells and to a lesser extent in macrophages and neutrophils (166,167). Butyrate is taken up by IECs and binds its cognate receptors, intracellular signalling exerts anti-inflammatory effects through inhibition of histone deacetylase (HDAC) that suppresses activation of nuclear factor kappaB (NF-κB) hence a reduced inflammatory response which is protective in UC (168).

A study performed on faecal samples obtained from UC patients compared to healthy controls showed a reduction in R hominis and F prausnitzii, both bacteria belong to Firmicutes phylum and producers of Butyrate. Data also showed a reduction in Butyrate in UC patients' faecal samples (169). On the contrary, pathogenic Proteobacteria such as E. coli and Shigella are increased in UC patients. A healthy gut usually has a population of E. coli and Shigella of <0.5% of the gut microbiota, a value of 3.9% has been reported in UC patient samples (169). Therefore, studies have been conducted to explore the use of butyrate treatment as a therapy in UC. A study performed by Varela (2013) showed re-colonisation and increase in F. prausnitzii

in UC patients under remission as compared to those with active disease. One of the main promising therapies focused on restoring homeostasis in ulcerative colitis through re-colonisation is the use of fecal microbiota transplantation (FMT) (171).

In a study using a DSS induced mouse model of colitis showed that administration of butyrate significantly reduced neutrophil infiltration in the colon mucosa and improved the inflammatory profile including activated macrophages, B and T lymphocytes. The data showed that butyrate improved mucosal lesions and attenuated the inflammatory profile in the intestinal mucosa, local draining lymph nodes and Peyer's patches of the experimental group compared to controls (172). A randomised controlled trial (RCT) FMT study has shown clinical remission in 42.1% of patients with UC as compared to the control group. This data presents FMT as a promising and safe potential therapy in the management of UC; however, further research is required to optimise the FMT procedure in larger cohorts (171,173).

1.2.1.5 T helper 2 (Th2) immune response is predominant in UC

Ulcerative colitis is characterised by decrease in microbial diversity and a reduced production of mucosal surface mucus (169). This results in exposure of the gut epithelium to pathogens, toxins and weakening of tight junction proteins (174). In response to epithelium damage and invasion of pathogens, intestinal epithelial cells (IECs) secrete defensins and signals such as IL-18, IL-33 and IL-37, which are important in activation and recruitment of other innate immune cells (175). Luminal antigens are recognised by innate immune cells through PRRs such as Toll-like receptors and NOD-like receptors. PRRs are expressed on IECs but predominantly on tissue resident antigen presenting cells (APCs), these include macrophages and dendritic cells (DCs). IECs express toll-like receptors (TLRs) 2 and 4 for antigen recognition, expression of TLRs 2 and 4 is significantly increased in patients with UC (175,176). Impairment of the intestinal epithelial barrier results in activation of resident APCs

for example macrophages located in the lamina propria. Macrophages detect, phagocytose and neutralise pathogens intracellularly through opsonisation. The DSS induced mouse model of colitis is a well-established animal model of mucosal inflammation that has been used to study the pathogenesis of UC. A study by Zhu (2014) in a DSS induced mouse model of colitis showed interplay between pro-inflammatory (M1) and anti-inflammatory (M2) macrophages. In the initial phase of the immune response, percentages of M1 and M2 were nearly equivalent; however, the balance shifted as M1 increased (from 14.08% ± 1.83% to 38.58% ± 4.29%) while M2 reduced (from 15.49% ± 1.21% to 11.54% ± 1.49%, p50.01). There was also increased secretion of pro-inflammatory mediators such as TNF-α, IL-1β, IL-6, IL-8, IL-12, IL-18, IL-23, CXCL9 and CXCL10 (177).

A subsequent parasitic infection in this colitis mouse model initiated a Th2 immune response, associated with an increase in M2 macrophages with increased secretion of IL-10 and TGF-β, there was increase in frequencies of FOXP3+ T regulatory cells, suppressing inflammation and promoting tissue remodelling and repair (177). Other studies performed in biopsies collected from both UC and CD patients in comparison to healthy controls confirmed a major role for macrophages in the pathogenesis of IBD, single cell sequencing data showed an increase in expression of genes linked to T cell recruitment and activation such as CXCL9, CXCL10, CXCL11 and CD40 within the macrophage population. It suggests a key role that macrophages play in the recruitment of T cells into inflamed mucosa hence driving a pro-inflammatory microenvironment (178–180). Overall, the balance between M1 and M2 macrophages is a key factor in chronic inflammation and drives pathogenesis in UC.

Neutrophils are also recruited to damaged intestinal tissue in response to chemokine signals such as CXCL8 (IL-8). In a regulated immune response, neutrophils release reactive oxygen species (ROS) and antimicrobial peptides that destroy invading pathogens. Neutrophils release

extracellular traps referred to as Neutrophil extracellular traps (NETs) by a process called Netosis to trap pathogens. In UC, there is a significant increase in neutrophil frequencies in inflamed tissue, increased activity of neutrophils results in release of damage-associated molecular patterns which in turn facilitate chronic inflammation (181,182).

In response to impaired intestinal epithelium and invading luminal pathogens, activated innate immune cells release mediators that prime the adaptive immune response through antigen presentation; the main cells that initiate this process are dendritic cells (DCs). DCs phagocytose antigen, process it intracellularly and migrate to the draining lymph nodes where they present and stimulate naïve CD4+ T cells that differentiate into effector phenotypes depending on signalling mediators in the inflamed microenvironment. Multiple effector cells are activated in UC, however a study by Nemeth (2017) using bulk RNA-sequencing on resected tissue samples collected from UC, CD and non-IBD (control) revealed an upregulated mRNA cytokine profile within the T cell population that included *IL5*, *IL13*, *IL15* and to a lesser extent *IL4* in UC samples when compared to CD and controls, suggesting Th2 driven inflammatory response in UC (183). Furthermore, a study by Li (2016) on mucosal biopsies showed increased frequencies of lamina propria GATA-3+ CD4+ T-cells secreting IL-5 and IL-13 in UC patients as compared to CD and controls (184). Another study has shown increased expression of IL-33 at both transcript and protein level compared to CD and controls. IL-33 expression was upregulated in colonocytes of UC patients and so was its corresponding T1/ST2 receptor, this suggested an active Th2 immune response (185). In addition, Ngo (2021) used a DSS induced mouse model of colitis to show a key role of IL-33 in expansion of Type 2 innate lymphoid cells that have a similar phenotype to Th2 immune cells. The data show a promising trend of targeting the IL-33/ST2 receptor pathway as potential therapeutic target for UC (185,186). Although earlier studies had considered the immune response in UC to be Th2 driven, it has been shown that after blocking the expression of Th2 cytokines such as IL-13 in

a clinical trial, there is no improvement in clinical outcomes of UC patients (187,188). This was attributed to inflammation driven by other immune cells such as natural-killer T (NKT) cells, Th17 cells that produce IL-17, and Th1 cells that produce IFN-γ and IL-12.

The pathogenesis in IBD involves epithelial damage, aberrant lymphocyte activation and infiltration. Integrins are important molecules in T cell trafficking and gut homing, Lamb (2017) used sequencing studies to show that CD4+ T cells that expressed αEβ7 had lower expression of genes associated with regulatory function including *FOXP3, IL10, CTLA4* and *ICOS* as compared to CD4+αEβ7- phenotype. In UC, the CD4+αEβ7+ T cell subset expressed higher levels of IFN-γ, TNF-α, and enriched in T cells that expressed both IL-17 and IFN-γ (189,190). The gut homing integrin αEβ7 is important for T cell mucosal homing as it signals through its interaction with E-cadherin. The data showed that colonic CD4+αEβ7+ T cells in UC patients expressed pro-inflammatory cytokines associated with Th1, Th17 and Th1/Th17 with a reduced expression of those associated with T regulatory cells. This suggests that colonic CD4+αEβ7+ T cell phenotype as a key cell population in driving inflammation and UC pathogenesis (189).

Overall, the immune response in UC is complex and varies across patients, although Th2 cells have previously been reported as the predominant cell type that drives disease pathogenesis in UC, multiple cell types are involved too including cells of both the innate and adaptive immune systems such as DCs, macrophages, neutrophils, IECs, Th1, Th2, Th17 and T regulatory cells. Further studies are required to identify therapeutic targets with inhibitory effects across cell types while limiting off-target effects (183,185,191).

1.3 Crohn's disease

CD is a chronic inflammatory condition of the gastrointestinal tract (GIT) that affects both males and females usually between 15-29 years of age. Recent studies have shown disease

onset at a later age of 55-59 years (192,193). Inflammation in CD can occur in any part of the GIT including both the colon and small bowel as shown in figure 1.3 A; however, the ileocecal regions are predominantly affected. The pattern of inflammation includes discontinuous inflammatory lesions that are transmural extending to the deeper layers of the GIT with the full bowel wall affected (194,195). The inflamed region appear as longitudinal ulcers of ≥4 to 5 cm (Japanese criteria) as shown in figure 1.3B (46) . The deep ulcerations also form features that appear as cobblestone seen in figure 1.3C. Other features that are unique to CD include sparing of the rectum, presence of perianal disease, and occurrence of strictures and fistulas (46,196)

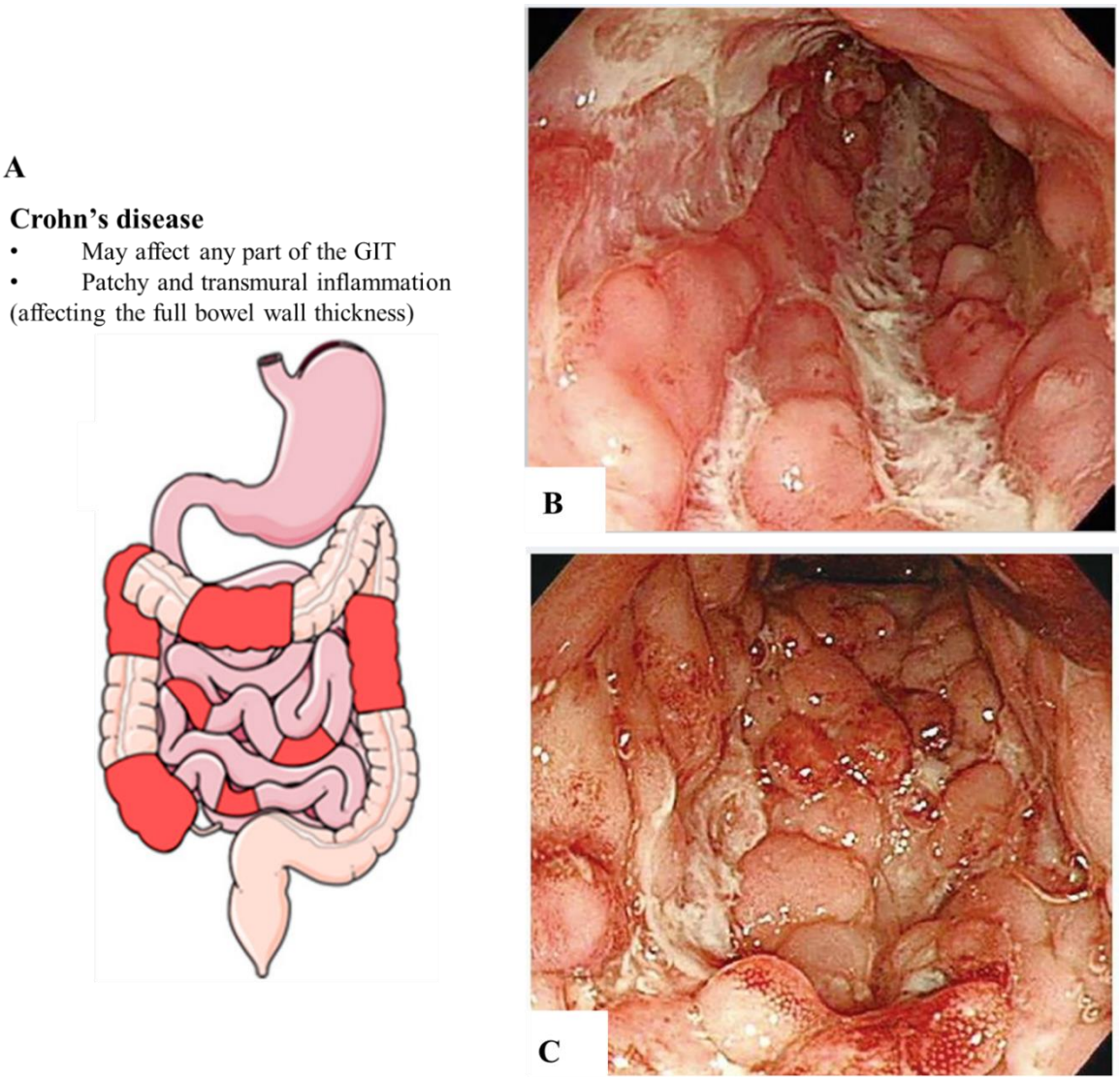


Figure 1.3: Schematic representation showing location and inflammation pattern of the affected areas in CD. A) Inflammation in the colon and small bowel that is discontinuous and patchy (red-coloured region in gastrointestinal tract). **B)** Endoscopic features of CD showing longitudinal ulcers **C)** Severe inflammation with cobblestone appearance. *Images obtained from Lee and Lee, 2016; Hoter and Naim, 2019*

Histological findings include clustering of immune cells referred to as granulomas, fat deposition, fistulae and abscess (197,198). Severe cases of CD show a chronic inflammatory phenotype with full bowel thickening resulting in stricturing disease and bowel obstruction, this is common along the terminal ileum (199,200). CD diagnosis is based on physical presentation of the patient, laboratory tests and imaging. Patients typically present with

diarrhoea, weight loss, fever, perianal disease, signs of malnutrition and growth failure in children. The aim in management of CD is to suppress an overly active intestinal inflammatory response, the current standard of care treatment involves steroids, immunomodulators such as methotrexate and thiopurines (azathioprine and 6-mercaptopurine), biologics therapy such as anti-TNF-α agents (Infliximab, Adalimumab), anti-IL-12/IL-23 (Ustekinumab) and anti- α4β7 (Vedolizumab). Immunomodulators are effective therapies in the management of CD; however, data has shown that they are less effective than biologic therapies or combinations of biologics (201,202).

In the SONIC trial25, infliximab monotherapy, azathioprine monotherapy, and the two drugs were combined in moderate-to-severe CD treatment for naïve patients. Of the patients receiving combination therapy, 56.8% were in free clinical remission and mucosal healing at week 26 (the primary end point) as compared to 44.4%) receiving infliximab alone and patients (30.0%) receiving azathioprine alone. A similar numerical trend was found at week 50 (203–205). Overall, a combination of immunomodulators and biologics have been effective in the management of CD. Despite the success of the available therapies in attenuating inflammation, approximately 40% of patients lose response to biologics therapy. Loss of response has been attributed to immunogenicity such as formation of antibodies that either neutralise the drug or hasten the drug's clearance; both actions limit the drug's efficacy (206). As indicated above, approximately 50% of CD patients development intestinal strictures that will require surgery; there are currently no anti-fibrotic therapies for management of strictures, the standard of care treatment is surgical resection which presents a major unmet need (207,208).

CD presents in multiple forms including inflammatory, stricturing or fistulising disease phenotypes. Fistulas are formed due to a connection between two epithelial surfaces following chronic inflammation. The three main classifications include; entero-enteric in which a fistula

is formed when a segment from the small intestine joins another segment either from the small intestine or colon, entero-cutaneous fistula occur when a segment of the small intestine connects to the skin, while entero-vesical fistula are formed as a result of an abnormal connection between the small intestine and the urinary bladder (209–212). Overall, accurate diagnosis of disease phenotype is important as it informs on the treatment strategy to manage disease-specific symptoms (213).

1.3.1 Aetiology of CD

1.3.1.1 Genetic mutations on the NOD2 gene increase the risk of CD

The aetiology of CD is multifactorial, genetic predisposition is one of the main factors involved in pathogenesis and a major determinant in disease outcome. An editorial on twin studies compiled from 1988-2010 showed that there was a strong genetic association in CD with concordance rate of 30.3% and 3.6% on monozygotic and dizygotic twins respectively. The data showed that 5.2-22.5% of people diagnosed with CD have history of IBD in their family (64,214–216). Familial and twin studies, data obtained through linkage studies, GWAS and sequencing in the last 20 years have identified risk loci on genes important in intestinal immune regulatory mechanisms. To date, a total of 240 risk loci have been identified in IBD (217), with over 200 of those linked to increased risk of CD (218). The earliest and strongest susceptibility locus to be identified specific for CD is on the *NOD2* (*CARD15*) gene. *NOD2* encodes for *NOD2* (nucleotide-binding oligomerization domain 2) protein (219), one of the PRRs expressed on macrophages, dendritic cells, epithelial and Paneth cells of the gut epithelium. *NOD2* plays a key role in intracellular microbial recognition of muramyl dipeptide (MDP) that makes up part of the bacterial wall. Innate immune cells can recognise tissue invading pathogenic bacteria through expression of *NOD2*, and signal through the NF-κB pathway to initiate transcription of target genes that encode for secretion of pro-inflammatory mediators such as IL-8, TNF-α, and IL-1β (220). Impairment of NOD2 signalling in innate immune cells

results in a dysregulated immune response leading to chronic inflammation which is characteristic of CD (221). Other notable polymorphisms have been identified on TLR4 (222,223). The key role that TLRs play as part of the PRRs used in antigen recognition has been well documented whereby innate immune cells express TLRs through which they recognise luminal antigens that could be derived from microbes, such as unmethylated double-stranded DNA, single-stranded RNA, lipopolysaccharide (LPS), lipoproteins, and flagellin. The antigens are ingested and processed then presented to activate an appropriate adaptive immune response. Mutations in genes linked to antigen presentation may result in a dysregulated immune response which has been noted in CD patients (84,224).

The adaptive immune system is essential in clearance of pathogens and regulating the immune response. However, risk loci on genes associated with the adaptive immune response have been identified in CD patients, these include *IL23R, JAK2, STAT3, CCR6, ICOSLG* (218). In particular, the expression of the heterodimeric complex *IL23R* and *IL12RB1* regulate the activity of IL-23 and IL-17 cytokine axis, this is a key pathogenic mechanism that mediates development and progression of inflammation in Th17 cells. Therefore, gene mutations within the IL23R locus may result in an impaired immune response associated with Th17 cells. Th17 cells play a key role in clearance of extracellular bacteria and other pathogenic organisms; impairment in their signalling drives a pro-inflammatory phenotype which has been found in CD. This is supported by studies that have shown increased IL-17 levels in both intestinal mucosa and serum of CD patients (225,226). Other important risk loci have been identified on genes involved in maintaining epithelial barrier function, autophagy and antigen presentation. The genes include *ATG16L1, IRGM, PTGER4, ITLN1, DMBT1, XBP1, HLA*-region, *TNFSF15, PTPN2, PTPN22* (227). Overlap in risk loci that have been identified as predisposing factors that apply to CD and other diseases include; *HLADR1:03, IL10, IL23R* and *IL12B* which have been reported as risk genes for both CD and UC; *PTPN22* and *CCR6* are shared risk loci

between CD and rheumatoid arthritis, coeliac disease, multiple sclerosis and many others. Individuals with mutations on these genes are predisposed to CD and there is overwhelming evidence of a genetics being a strong contributor in the pathogenesis of CD (228,229).

1.3.1.2 Epigenetics mechanisms are influenced by environmental factors in CD

Genetics account for 13.6% of disease variability suggesting an influence by other factors (230). Epigenome wide association studies (EWAS) have identified genes that have undergone epigenetic modifications that may contribute to the pathogenesis of CD (231). DNA methylation has been the main epigenetic mechanism identified in CD (232). Epigenetic changes can be triggered by environmental factors such as smoking, diet, and the microbiome, these factors have already been indicated as risk factor for CD (233,234). A study by Kim & Cheon (2020) using whole blood and biopsies from CD patients showed differentially hyper methylated genes, subsequent pathway analysis revealed hypermethylation in genes associated with development of the cell structure and cytoskeleton. These included protein localisation to cell periphery (*EMP2, ANXA2*); connective tissue development (*RUNX3, COL21A1, TRPM4*) and regulation of actin filament-based processes (*PPM1F, NUMA1, BCAS3*). Hyper methylation of these genes resulted in their transcriptional silencing therefore impairing key biological processes and in turn contributing to the pathogenesis of CD (235). On the other hand, hypomethylated genes were found to be involved in processes important in pro-inflammation such as signalling by interleukins (*IL21R, TNFRSF18, IL12RB1, FGF17*), lymphocyte proliferation (*IDO1, ITGB7, LEF1*), and cytokine production (*TLR1, LEF1, IL22RA2, HDAC7*) suggesting that this mechanism regulates expression of pro-inflammatory genes as part the immune response in CD (236,237).

Other epigenetic modulators of CD include miRNAs; a study conducted on endoscopic biopsies from treatment naïve adult and paediatric CD patients gave an indication of the role

of miRNA in disease progression. Aberrant expression changes in miR-31 may play a role in the pathogenesis of CD; a study by Olaru (2011) identified factor inhibiting hypoxia inducible factor 1 (FIH1), an inhibitor of hypoxia inducible factor 1, as a direct target of miR-31. Targeting FIH1 leads to repression of survival in endothelial cells, thus promoting cell apoptosis and growth arrest, this could be a factor in driving a pro-inflammatory microenvironment (238). In adults diagnosed with CD, low levels of miR-31 at the time of resection surgery indicated a worse postoperative prognosis. Paediatric patients with low levels of miR-31 at the time of diagnosis had an increased risk of development of strictures. These findings indicated that miR-31 could potentially be used as a predictive marker for disease progression and outcome in CD (239).

MiR-31 has also been reported to modulate activation of T cells by inhibiting expression of molecules that control T-helper function, such as CD40L. MiR-31 is down-regulated in follicular helper T cells (TFHs), TFHs are a key component of B cell activation in production of antibodies during the adaptive immune response. Differentiation and function of TFHs is mediated by BCL6; BCL6 suppresses miR-31 expression in TFHs by binding to its promoter. Reduced expression of miR-31 leads to a decrease in CD40:CD40L driven costimulatory mechanism of T cell activation (240). Another study conducted with samples from CD and UC showed unique expression signatures that were specific to each condition. The data showed miR-19a, miR-31, miR-142-3p, miR-375, and miR-494 as statistically significant and upregulated specifically in CD (241). MiRNAs play key roles in modulating various immune processes, however, there is need for further studies to understand their specific roles in disease mediating mechanisms; miRNA present opportunities for identification of novel targets in various immune mediated diseases including CD.

1.3.1.3 Environmental factors such as smoking is associated with worse CD

Interaction of the environment with the immune system in a genetically predisposed host is implicated in the pathogenesis of CD. Environmental factors that have been studied as potential risk factors in development of CD include smoking, diet, medications, stress and exercise (242,243). Smoking has a contradictory effect in CD as compared to UC, with smokers having an increased risk to develop CD. CD patients that smoke experience a severe inflammatory response, patients with a history of smoking had with CD were associated with an increased proportional surgery rate (45.8% vs. 37.8%, P = 0.045) as compared to non-smokers, smokers had increased recurrence in patients that have undergone surgery and poor outcomes post-surgery.

The mechanisms that result in detrimental effects of smoking in CD are not clearly defined; the data therefore strongly encourages patients with CD to quit smoking (244). Previous studies have shown an increased risk of CD in women on contraception; a study in a UK cohort of women on combined oral contraceptive pills (COCPs) was conducted to measure their risk of developing CD. The data showed an odds ratio (OR) of 1.6, with the ratio being a measure of association between women on COCPs and development of CD. The odds ratio was lower in UC at 1.3, there was no disease association in women that medicated with the progesterone contraception only. This suggests that the possible risk factor could be the oestrogen in the COCPs (245). The mechanisms are not clear; however, oestrogen has been shown to enhance cellular proliferation and humoral immunity, modify colonic barrier function and contribute to thrombosis. These factors may become triggers in onset of CD particularly in a genetically disposed individual (246–248). Other factors such as stress and use of nonsteroidal inflammatory medications (NSAIDs) may contribute to severe disease phenotype although there is no risk involved (249). In general, studies show that external environmental factors play an important role in the pathogenesis of CD in combination with other factors.

1.3.1.4 Microbiome in CD

Previous studies have shown key roles of commensal microbes in maintenance of homeostasis in the gut mucosa. Dysbiosis between "healthy" inhabitant bacteria of the gut mucosa and pathogenic bacteria has been implicated in the pathogenesis of CD (250). The healthy human gut is colonized by species that belong predominantly to phyla Bacteroidetes, Firmicutes and to a lesser extent Proteobacteria, Actinobacteria (251). Studies have shown a 25% reduction in microbial transcriptome diversity in CD in comparison to controls (252). This has mainly been attributed to reduction in Firmicutes in the gut; these species include *Ruminococcus albus*, R. *callidus,* R. *bromii*, C. *leptum* and F. *prausnitzii*. The reduction of bacteria that established a colony and a symbiotic relationship within the tissue enables invasion of facultative bacteria. An example of protective mechanisms by the Firmicutes group is exhibited by C. *leptum*. In a healthy gut, C. *leptum* produces butyrate, a SCFA described earlier with immunomodulatory effects in processes such as intestinal barrier function, cell proliferation and differentiation. Butyrate also acidifies the intestinal lumen making it an unfavourable environment for pathogenic bacteria (253). In contrast, species belonging to the Proteobacteria phylum are increased in CD (254). The data suggests how changes in microbiota can increase CD risk, it also shows that the microbiota is important in modification of disease phenotype and determining the course of CD progression.

1.3.1.5 Th1/Th17 plasticity is predominant in CD immune response

The immune response in CD is characterised by chronic inflammation in response to tissue injury, driven by a combination of environmental factors and a dysbiosis of the gut microbiota in a genetically predisposed individual which determines the extent and phenotype of immune response (255). Luminal antigens are recognised by innate immune cells such as dendritic cells, macrophages and neutrophils that are the first responders to invading pathogens. Innate

immune cells produce pro-inflammatory cytokines such as TNF-α, IL-6, IL-18 and IL-1β, these cytokines modulate pro-inflammatory mechanisms in response to tissue injury. They perform their function through neutrophil recruitment, stromal cell activation and proliferation (22). Similar to UC, antigen recognition is achieved via PRRs expressed on APCs, mainly macrophages and DCs. One of the main PRRs is the NOD-like receptors (NLRs), which is the first risk locus implicated in the pathogenesis of CD (256). *NOD2* encodes for the NLRs expressed on APCs; studies have shown impaired function of these receptors in CD patients, this can impact CD phenotype and determine disease severity (257).

One of the main PRRs is the nucleotide-binding domain and leucine rich repeat containing family (NLRs), which is the first risk locus implicated in the pathogenesis of CD (256). NLRs are a family of intracellular PRRs whose stimulation results in activation of the caspace-1 signalling complex (inflammasome), mediating cleavage of pro-IL-18 and pro-IL-1β or activation of NF-kB signaling. *NOD2*, a member of the NLR family encodes for NOD2 protein that is expressed on APCs. As a pathogen-recognition molecule for muramyl dipeptide (MDP), NOD2 modulates mechanisms associated with innate and adaptive signalling such as regulation of cytokines, chemokines and antimicrobial peptides production. Therefore, individuals with genetic mutations on the NOD2 locus as are predisposed to abnormal bacterial sensing and development of inflammatory disorders which is characterised in CD. This also presents a potential therapeutic approach in CD involving restoring of NOD2 function (219,221,258).

As described previously DCs are the main APCs in the gut, which phagocytose pathogens and migrate to the lymph node where they present antigen to naïve CD4+ T cells. Specifically, the immune response in CD is characterised by Th1 and Th17 plasticity. In the presence of IL-12, Th1 effector cells differentiate from CD4+ T cells. Th1 cells secrete IFN-γ, IL-1β, IL-6, IL-8

and TNF-α (259,260). Studies have shown increased frequencies of CD4+ T-cells expressing T-box transcription factor (T-bet); T-bet is an important factor in the development of Th1 cells and is key in transactivation of IFN-γ. Interestingly, the CD4+ T-bet+ phenotype also showed increased levels of IFN-γ in mucosal biopsies from CD patients (261,262). CD4+ T cells differentiate into Th17 cells in the presence of IL-23 and TGF-β, and several studies have implicated Th17 cells in the pathogenesis of IBD (263–265).

Research has shown that in CD, there is a CD4+ T cell phenotype that expresses markers for both Th1 and Th17, and these cells are characterised as Th17 secreting IL-17, IL-22, IFN-γ and expressing CXCR3 (266). Th17 plasticity is further highlighted in a Th17 transfer model of colitis in which Th17 cells were found to transition into "Th1 like" cells that lost expression of IL-17 and increased expression of IFN-γ. Th17 cells were found to act as precursors of pathogenic Th1 cells that drove the inflammatory immune response in CD (267). Another study by Veny (2010) also showed an increased percentage of Th1/Th17 cells in the lamina propria of CD patients especially at the later stages of disease (268). This indicates that the immune response in CD is predominantly driven by Th1/Th17 inflammation. IL-12 and IL-23 associate with Th1 and Th17 populations respectively and share the p40 subunit. The dynamics of the two cell subsets have been of interest as a potential target for anti-fibrotic therapy particularly looking to dampen inflammation by targeting the IL-12/IL-23 axis (269,270). Other cell types such as Th2 cells are also important contributors in CD but most importantly are Tregs that play a suppressive role to regulate effector T cells during inflammation. Studies have shown a reduction in Treg frequencies and function in peripheral blood and tissue collected from CD patients therefore impairing their suppressive role (271,272).

Overall, pathogenesis in both UC and CD is attributed to a combination of factors; the immune response in both conditions is complex however there are disease specific risk factors that

influence the course of disease progression. One of the main outcomes of severe disease in both UC and CD is tissue fibrosis. Colonic strictures have been reported in 1% to 11.2% of UC patients (273), and over 50% of CD patients develop strictures (274). The next sections focus on the various factors that influence formation of strictures associated with CD.

1.4 Intestinal fibrosis

Intestinal fibrosis is the end-stage process resulting from a dysregulated immune response affecting the GIT, over 50% of CD patients develop strictures within 10 years of diagnosis (199). Intestinal fibrosis results from chronic inflammation and an imbalance between deposition of extracellular matrix (ECM) proteins and their degradation. The ECM is a non-cellular component providing essential physical scaffolding and signals required for tissue formation, differentiation, and homeostasis (275,276). ECM proteins include structural proteins such as crosslinking collagens and elastin, adhesive proteins such as laminin and fibronectin, ground substances within the mesenchyme such as hyaluronan and glycoproteins, matrix metalloproteinases (MMPs) and TIMPs tissue inhibitors of metalloproteinases (TIMPs) (277,278). MMPs are enzymes that play a major role in degradation of the ECM, while tissue inhibitors of metalloproteinases (TIMPs) inhibit the actions of MMPs.

Mesenchymal stromal cells (MSCs) are the main producers of ECM proteins (279,280), there are three main types of mesenchymal cells that are resident in each tissue: fibroblasts, myofibroblasts and smooth muscle cells (281). Tissue integrity is maintained when ECM proteins are deposited, these are constantly remodelled by MMPs whose activity is regulated by TIMPs hence maintaining a controlled deposition of matrix and tissue homeostasis. In chronic inflammation, myofibroblasts acquire a pathogenic phenotype with an increased production of ECM proteins, a downregulation of MMPs and an upregulation of TIMPs. The imbalance in MMPs and TIMPs results in formation of a highly cross-linked fibrotic structure

which leads to formation of fibrotic strictures. Previous studies revealed a dysregulation in MMPs and TIMPs in CD patients presenting with intestinal strictures, resulting in excess deposition of ECM proteins with an impaired degradation mechanism (282–284). Restoring a balance between tissue MMPs and TIMPs is a key therapeutic target in the management of fibrosis (285).

1.4.1 Intestinal structure

The intestinal mesenchyme is separated into distinct layers, these include the mucosa, sub-mucosa, muscular propria and serosa (Figure 1.4), with the structure and cellular distribution of each layer being dependent on its function. The mucosa is the inner most layer facing the lumen, it is made up of the gut epithelium, lamina propria (villi and intestinal crypt) (286,287) . The mucosal lining is covered with mucus secreted by goblet cells that make up part of the epithelium. The mucus and epithelial barrier are the first line of defence to invading pathogens and toxins. Epithelial cells are located in the gut epithelium, while intestinal villi and crypts form part of the mucosa that plays an important role in absorption of nutrients and the maintenance of a homeostatic mucosal surface (288,289). The submucosa is the next inner most layer and is made up of tissue resident cells such as fibroblasts, lymphocytes, basophils, mast cells. It is highly vascularised and made up of the lymphatic drainage. Next is the muscularis propria which mainly consists of smooth muscle cells, this layer also contains abundant muscle fibers that regulate movement of products through the gut by peristalsis. The gut serosa is the outer most layer that plays a major role in tissue protection (290–294).

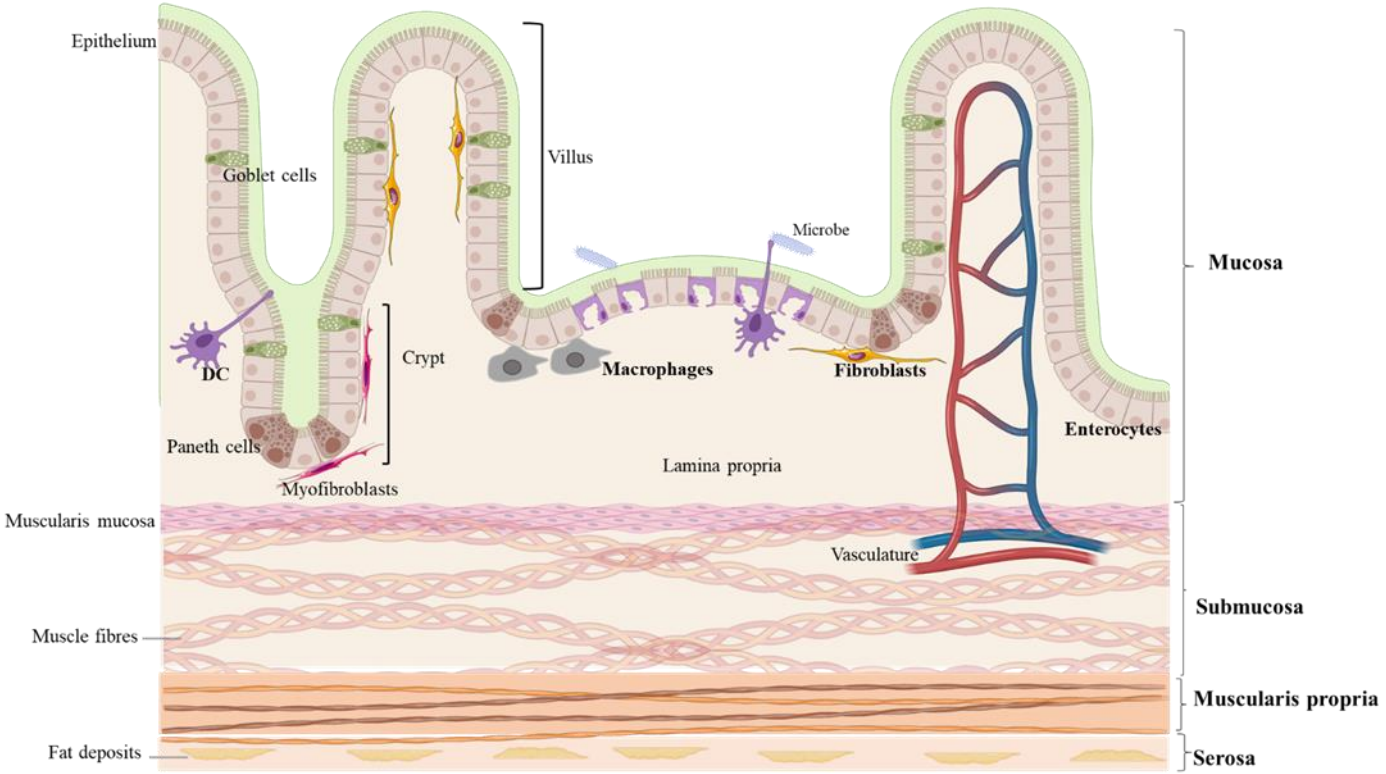


Figure 1.4: Schematic diagram showing cross-sectional structure of the small intestinal wall. The four layers of the gut include; mucosa (villi, epithelium, crypts, enterocytes, goblet cells, Paneth cells and fibroblasts). Submucosa contains durable and flexible muscle fibres and vasculature, Muscularis propria is made up of well-developed muscle fibres that are able to produce churning movements required for mechanical digestion. Serosa contains a layer of simple squamous epithelium and fat deposits. ***Image was obtained from BioRender and modified***.

1.4.2 Stricture formation in patients with Crohn's disease

The mechanisms involved in the formation of intestinal strictures are unclear, however it has been attributed to chronic inflammation and a dysregulated wound healing process (295,296). Strictures are characterised by excess deposition of ECM proteins predominantly type I collagen and increased levels of pathogenic myofibroblasts that have contractile properties and an increased rate of deposition of ECM (297–299). Figure 1.5 shows a stricture in a CD patient characterised with thickening of the bowel wall extending into the lumen leading to intestinal blockage (Figure 1.5 A), an endoscopic balloon dilation (EBD) technique is used to open up the lumen (Figure 1.5 B-C), and Figure 1.5 D shows the lumen opened up post balloon dilation procedure (300). The procedure does not automatically resolve the stricture, there are high incidences of recurrence which may eventually result in resection surgery. Treatment options for management of strictures are based on the specific phenotype. Intestinal stricture phenotypes include inflammatory, fibrotic or combined depending on the histological features. Inflammatory strictures are usually treated using anti-inflammatory drugs while fibrotic strictures are managed mechanically either by stenotic dilation or resection surgery (301,302).

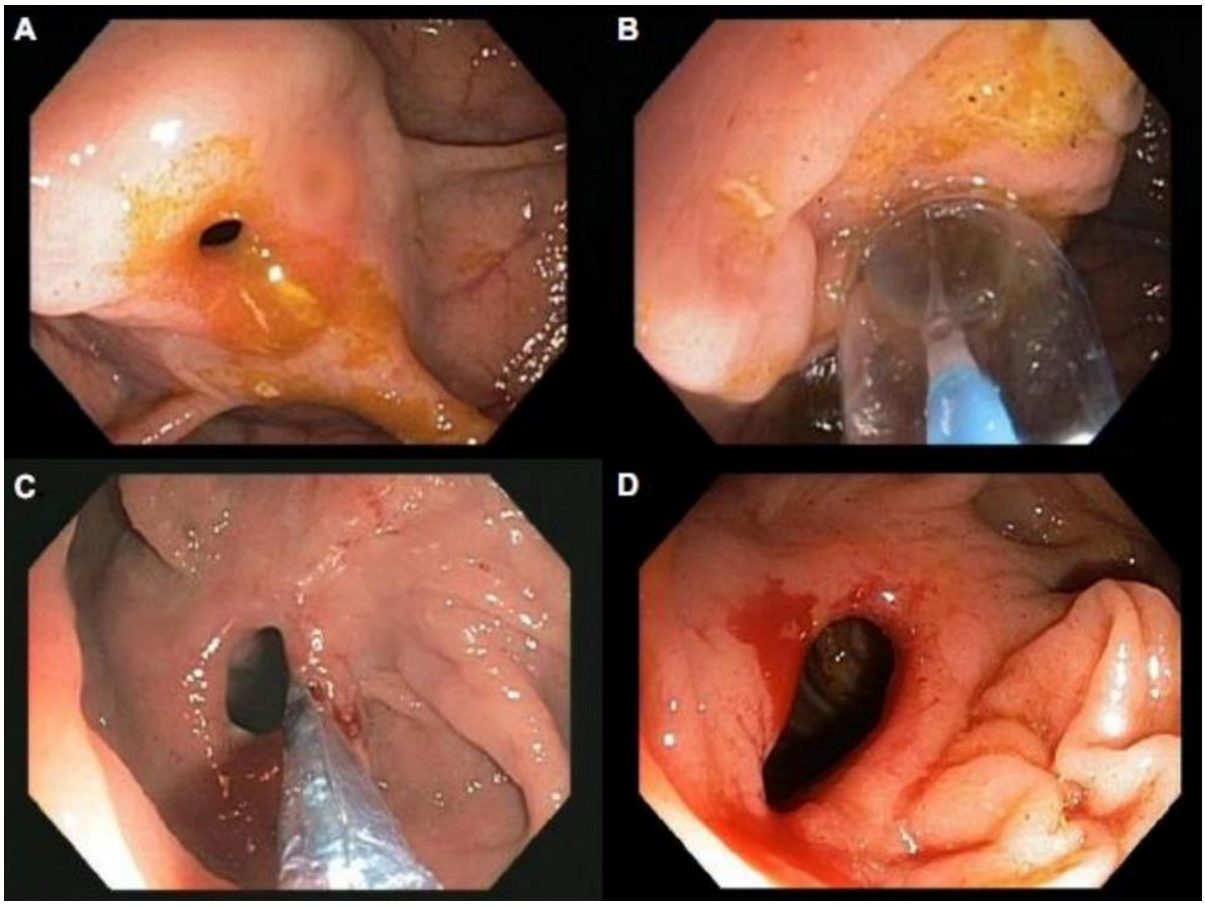


Figure 1.5: Representative image showing the process of endoscopic balloon dilation on an anastomosis stricture in a CD patient. A) Stricture showing thickening of the bowel wall and narrowing of the lumen causing blockage. **B-C)** The process of endoscopic balloon dilation **D)** Post balloon dilation showing clear access to the rest of the intestinal lumen. ***Image from (Keihanian and Moss, 2016)***

1.4.3 Histological features of strictures

Intestinal strictures are commonly located in the terminal ileum and ileocolonic region of CD patients (199,302). A key feature of CD is chronic transmural inflammation across all layers of the bowel; these include the mucosa, submucosa, muscularis propria and serosa. Chronic inflammation results in formation of penetrating ulcerations, crypt abscess and cryptitis (198,303). In the attempt to resolve inflammation, the inflamed regions undergo multiple rounds of deposition of collagen rich ECM (Collagens I, III and V) proteins. The balance between deposition and degradation of ECM which is regulated by MMPs and TIMPs is

impaired, resulting in bowel thickening and stiffness leading to formation of strictures (304–306).

A study was performed by members of our team at the University of Birmingham in which resection tissue samples from CD patients at the QE hospital, Birmingham, UK were histologically scored, and their findings showed that strictures are composed of multiple features including acute or chronic inflammation, muscular hyperplasia/hypertrophy, fibrosis, adipose hyperplasia and volume expansion or thickening of the bowel wall. These features are localised to specific gut layers and contribute to the narrowing of the bowel, for example inflammation was most detected within the mucosa while fibrosis was more prominent in the submucosa. Muscular hyperplasia and hypertrophy were detected across all the layers of the gut (302). Histology features are important in the classification and definition of intestinal strictures, this provides relevant information on a more individual patient focused treatment strategy.

1.5 Cellular mechanisms involved in intestinal stricture formation

Various immune and non-immune cells have been implicated in driving the chronic inflammatory phenotype in CD in the attempt to restore homeostasis. However, in genetically susceptible patients, the wound healing process is impaired, and unresolved chronic inflammation eventually results in development of intestinal strictures. The mechanisms involved in the formation of strictures are complex, however, it is clearly triggered by inflammation (307,308). Intestinal resident cells regulate homeostasis, they are activated in response to injury or invading pathogens secreting various cellular mediators and signals through which other cells are recruited to the inflamed site in the tissue. This is a physiological process of the immune system to restore homeostasis. However, in a host with a dysregulated immune response, the result is a vicious cycle of activated immune cells that release signals

resulting in the activation of mesenchymal cells and vice versa while maintaining a chronic inflamed microenvironment within the affected tissue site. During these processes, cells undergo changes in behaviour such as increased cell differentiation and proliferation; there is increased release of signalling mediators including cytokines, chemokines and growth factors that promote further inflammation and initiate or exacerbate fibrosis. Some of the main soluble mediators include IL-1, IL-4, IL-6, IL-13, IL-21, IL-10, IL-12, IFN-γ insulin-like growth factor I (IGF), epithelial growth factor (EGF), basic fibroblast growth factor (bFGF/FGF2), platelet-derived growth factor (PDGF), transforming growth factor beta (TGF-β) and connective tissue growth factor (CTGF) (307,309). The following subsections explore the role of the different cell types in formation of intestinal strictures.

1.5.1 Immune cells regulate the inflammatory response that initiates stricture formation

The intestinal fibrotic microenvironment is characterised by a heterogenous cell population of both activated stromal and immune cells; both resident and tissue infiltrating cells that in response to injury. These cells release soluble mediators that are pro-inflammatory and profibrotic, examples of which include TNF-α, IL-1β, IL-33, IL-4, IL-6, IL-13, IL-21 IGF, PDGF, TGF-β, CTGF, VEGF, CCL2, CCL3 and IL-12 (310). Neutrophils are first recruited to the tissue in response to chemotaxis signals. A combination of bulk RNA and single cell sequencing data by Friedrich (2021) showed increased expression of IL-1R on inflammatory fibroblasts which were stimulated by IL-1β signaling. In the early stages of the inflammatory response, IL-1β is secreted by activated macrophages which in turn signals through its cognate receptor IL-1R. The study showed an upregulation of IL-1R, a neutrophil chemoattractant expressed on myofibroblasts. This indicate the key role IL-1β signalling playing in recruitment of neutrophils and driving the inflammatory response (311). Interestingly, studies in preclinical models of colitis have shown a reduction in inflammation by blocking IL-1β signalling, this shows a potential for IL-1β blockade therapies (312–314). Another study by Schroder (2021)

of intestinal biopsies of CD patients showed a significant correlation between disease severity and increased levels of neutrophil extracellular traps (NETs). NETs produce hypochlorous acid (HOCl) that reacts with phagocytosed bacterial pathogens and kills them through oxidative damage. In a dysregulated inflammatory response, excess release of these HOCl can lead to intracellular damage of various cellular components leading to apoptosis which in turn exacerbates inflammation (311,315,316).

Macrophages of the lamina propria are another key innate immune cell implicated in the development of strictures. These cells differentiate into a pro-inflammatory phenotype referred to as M1 and anti-inflammatory (M2) macrophages. Macrophage plasticity is not clearly understood, but its phenotype is determined by the mediators in the inflamed tissue microenvironment. Differentiation of M1 macrophages occurs in the presence of IFN-γ and TNF-α, M1 macrophages typically secrete pro-inflammatory cytokines such as IL-6, IL-12, and TNF-α. Alternative activation gives rise to the M2 phenotype in the presence of IL-4 and IL-13, M2 macrophages exhibit anti-inflammatory and pro-fibrotic mechanisms by secreting factors such as IL-10, TGF-β and PDGF (317,318).

A study performed in a rat 2,4,6-trinitrobenzene sulfonic acid (TNBS) induced mouse model of colitis showed a persistent M2 macrophage population that was localised near intestinal smooth muscle cells. Both cell types secrete pro-fibrotic mediators that modulate formation of intestinal strictures (319,320). Studies performed in a $STAT6^{-/-}$ mouse model with chronic colitis showed contradicting findings with anti-fibrotic properties. STAT6 signalling is a vital pathway in the development and function of M2 phenotype. In this study, the $STAT6^{-/-}$ mouse model showed significantly reduced frequencies of M2 with a predominant population of $CD16^{+ve}$ macrophages which present a more M1-like phenotype. These macrophages also showed increased mRNA expression of profibrotic genes such as *COL1A1*, *MMP2*, *TIMP1* and

ACTA2 which exacerbated colonic fibrosis in the STAT6−/− mouse. The wild type mice showed increased frequencies of M2 ($CD206^{+ve}$) and no fibrogenic response. Interestingly, administration of IL4-treated WT macrophages to STAT6−/− mice reduced TNBS-induced fibrosis(321). This study suggests that M1 macrophages secrete factors that promote inflammation and fibrosis while M2 macrophages released factors that were anti-inflammatory and anti-fibrogenic, this contradicts what has been published in the literature suggesting M2 macrophages as pro-fibrogenic in tissues such as the liver and lung (322,323). Macrophages are a heterogenous cell type, these cells undergo cellular plasticity and the composition of the macrophage population can be specific to each individual human or animal model, data is subject to variability across different patients (324–326).

Once the resident stromal and immune cells have been activated, they prime and activate the adaptive immune cells that are recruited to the site of injury in response to signals from cytokines, chemokines and growth factors. Of the CD4+ T cell effector sub types, Th2 cells have been implicated in development of fibrosis (327). CD4+ T cells differentiate into Th2 cells in the presence of IL-4, Th2 cells secrete IL-13 and IL-4. IL-13 has been shown as an important mediator in driving the fibrotic phenotype. IL-13 levels are elevated in intestinal fibrosis with IL-13 shown to signal through the IL-13Rα2 expressed on fibroblasts (328). This signal regulates the secretion of proteins such as periostin which facilitate activation and proliferation of fibroblast into myofibroblasts, leading to increased production of ECM proteins (329).

Periostin is matricellular protein encoded by *POSTN* gene; it is involved in various inflammatory and fibrotic diseases. Its mechanism of action in intestinal fibrosis is not clear however, studies in idiopathic pulmonary fibrosis (IPF) have suggested that fibroblast derived periostin interacts with other soluble mediators such as TNF-α to activate the NF-κB signalling

pathway inducing the production of pro-inflammatory cytokines or chemokines (330). Another mechanism involves cross-talks between TGF-β and periostin via αVβ3 integrin converging into SMAD3; TGF-β/SMAD downstream signalling results in transcription of pro-fibrotic genes (328,331,332). Data generated by (333) on bronchial biopsies collected from asthma patients showed an increase in protein expression of type I and II procollagens, ECM remodelling proteins such as MMP-2, MMP-9 and TIMP-1 in the presence of IL-13. Various studies published in the literature show that IL-13 plays a major role in fibrosis and, therefore a viable target for anti-fibrotic therapy (328,333–335).

Other cellular mediators such as TNF-α modulate inflammation and fibrosis. TNF-α exhibits pro-inflammatory properties, however, it has also been shown to mediate fibrosis. A study in intestinal epithelial cells in which they were stimulated with TNF-α resulted in upregulation of TGF-β and TIMP-1 expression. A similar effect was observed when epithelial cell conditioned media was used to stimulate colonic sub-epithelial fibroblasts, leading to an increase in expression of MMP-9 and collagen. The data indicated that stromal cells respond to pro-inflammatory mediators to secrete pro-fibrogenic factors (336). Another family of cytokines reported to have profibrotic properties are members of the IL-1 family. IL-1 is a large family of cytokines including IL-1α, IL-1β, IL-18, IL-33, IL-36α, IL-36β and IL-36γ acting as agonist cytokines. Binding of these cytokines to their cognate receptors leads to intracellular signalling and activation of pathways such as MyD88, IRAK, and TRAF6, which in turn activate NF-κB, p38, JNK, and MAPK stimulating innate immune cells to secrete IL-6, IL-5, IL-4, IL-8 and MCP-1 (337). Downstream signalling of these pathways was illustrated in a TNBS mouse model of colitis; the data showed an increase in intestinal inflammation linked to NF-κB signalling, which eventually resulted in excess deposition of ECM related factors and colonic fibrosis. Inflammation and fibrosis were reduced upon NF-κB signalling blockade (338); this

was similar to the data observed in DSS induced mouse model of colitis in which p38 signalling was implicated as a key signalling pathway in fibrosis (339).

Certain cytokines from the IL-1 family have been implicated in intestinal fibrosis; elevated levels of IL-33 and its cognate receptor IL1RL1(ST2) have been reported in peri cryptal fibroblasts in samples collected from CD patients with active inflammation (175,340,341). Data suggests an indirect role of IL-33 in development of fibrosis in CD by driving activation of fibroblasts, smooth muscle cells and endothelial cells to secrete increased levels of profibrotic mediators such as TGFβ and pathogenic collagens (342). IL-36, another member of the IL-1 family, is secreted by epithelial cells and macrophages upon tissue injury; IL-36 signals through its cognate receptor IL1RL2 (IL36R). A study performed on mucosal and submucosal intestinal tissue of patients with fibrostenosing CD showed increased levels of IL-36, type I and type VI collagen in correlation with higher numbers of α-smooth muscle actin+ (α-SMA+) cells compared to UC and controls (343). In the same study, an IL-36R−/ − mouse model generated by homologous recombination as described in (344) was induced for colitis using chronic DSS and chronic TNBS mouse models of colitis.

In both models, defective IL36R signalling showed antifibrotic outcome; histological findings showed a reduction in chronic inflammation evidenced by reduced cell recruitment into the lamina propria of the experimental group as compared to wild-type controls. In addition, there was a reduction in deposition of types I and VI collagen (191,343). Interestingly, in vivo data in the experimental mouse models of fibrosis exhibiting intestinal inflammation showed significantly reduced fibrosis when IL-36/IL-36R signalling was neutralised by administration of anti-IL36R; this suggests the IL-36/IL-36R signalling axis as a mediator of intestinal fibrosis (343,345). Another interesting member of the IL-1 family is the IL-37 which has been shown to suppress innate and adaptive immune responses. A transgenic IL37 (IL-37tg) mouse model

showed was protected against carbon tetrachloride (CCl4) induced liver inflammation. In addition, profibrogenic gene expression was lower in hepatic stellate cells (HSC) and Kupffer cells (KC) isolated from IL-37tg and IL-37 overexpressing mice. The study suggests protective effects of IL-37 through interacting with SMAD3 resulting in a downregulation in activation of liver HSC and KC; this interferes with the TGF-β signaling cascade that modulates liver fibrogenesis (346). Overall, the data suggests potential in targeting members of the IL-1 family as targets for potential anti-fibrotic therapy in intestinal fibrosis.

Other mediators that have been implicated in the pathogenesis of intestinal fibrosis include cytokines such as IL-17A (IL-17), a glycoprotein produced by Th17 cells. It is well characterised as a pro-inflammatory cytokine but has also been reported as a mediator of fibrosis (347). In chronic inflammation, IL-17 activates fibroblasts, endothelial and epithelial cells and macrophages to secrete IL-1β, IL-6, GM-CSF, TNF-α, matrix metalloproteinases (MMPs) and chemokines such as CXCL1, CXCL8, CCL2, CCL7, and CCL20 further exacerbating inflammation which in a dysregulated immune response, results in fibrosis (348–350). IL-17 has also been shown to induce expression of heat shock protein 47 (HSP47), a molecule that controls the expression of collagen I by intestinal sub-epithelial myofibroblasts. Expression of both HSP47 and IL-17A genes has been elevated in intestinal tissue linking it to accumulation of collagen (351–353). A study by Mi (2011) showed IL-17 promotion of epithelial mesenchymal transition (EMT) in vitro in alveolar epithelial cells in a TGF-β dependent manner. There was increased expression of EMT positive markers such as Snail, Slug, α-SMA and a reduced expression of EMT negative markers such as E-cadherin and ZO-1.

In the same study, a bleomycin (BLM) induced mouse model of IPF showed increased expression of IL-17 and its associated signalling mediators such as IL-6, TGF-β and, IL-23 in

bronchoalveolar lavage fluid (BALF) of BLM-challenged mice; In addition, expression of RORγt and IL-17R were upregulated in fibrotic lung tissue of BLM-challenged mice. Neutralisation of IL-17 in vivo using anti IL-17 antibody showed a reduction in deposition of collagen, hydroxyproline, expression of α-SMA and activation of MMP-2; there was a reduction in acute inflammation and increased survival of the experimental group as compared to the controls (354,355). The data suggests IL-17 as a driver of both pro-inflammatory and pro-fibrotic mechanisms through TGF-β dependent mechanisms. Il-17 is also involved in indirect mechanism that drive a pro-fibrotic phenotype especially in the later stages of the inflammatory response, this presents opportunities in the development of anti-fibrotic therapy.

1.5.2 Mesenchymal stromal cells

Mesenchymal stromal cells (Stromal cells) are a heterogenous cell population found in most tissues and are part of the mesenchyme that supports tissue structure and architecture. Stromal cells are characterised by a negative expression of CD45 and positive expression of surface lineage markers such as CD90, CD105, CD44, CD29, CD73 and vimentin (356–358). Stromal cells exhibit multipotent properties and play a distinct role in maintenance of tissue homeostasis, upon tissue injury and in a chronic inflamed microenvironment, they differentiate into fibroblasts, myofibroblasts or smooth muscle cells depending on microenvironment stimuli. Recent interest has been on the potential use of stromal cells in tissue regeneration, and studies have shown that infusion of therapeutic MSCs into CD patients presenting with perianal fistula led to wound healing, related to a downregulation of pro-fibrogenic factors including IL-1β, IL-6 and IL-13 and an upregulation of IL-10, a known anti-fibrogenic mediator (359,360). These findings indicate that stromal cells play a key role in tissue homeostasis and regeneration, and understanding the distinct roles of tissue specific stromal cells presents opportunities for a cellular target for anti-fibrotic therapy. The next sections describe the role of intestinal stromal cells and their precursors in the pathogenesis of intestinal fibrosis.

1.5.2.1 Myofibroblasts

Intestinal fibrosis is a result of excess deposition of ECM proteins with a dysregulated degradation mechanism. The main producers of ECM are the activated intestinal mesenchymal cells mainly myofibroblasts. Myofibroblasts are identified by their expression of mesenchymal markers such as Vimentin and Thy1 (CD90). Myofibroblasts are differentiated from other stromal cells through their increased expression of alpha-smooth muscle actin (α-SMA) and enhanced contractile properties (361). Myofibroblasts are mainly localised in the sub-mucosal layer which is a common site for fibrosis (362). Studies have shown an increase in frequencies of myofibroblasts in tissue collected from CD patients with intestinal strictures (363). During chronic inflammation, myofibroblasts acquire a pathogenic phenotype which exhibits phenotypic changes such as increased proliferation, migration and differentiation, myofibroblasts also produce a myriad of ECM proteins including collagens I, III, IV, V, and VI; glycoproteins and proteoglycans such as fibronectin, laminin, and tenascin. Collagens I is the most abundant ECM produced by myofibroblasts (310,364,365). The primary source of myofibroblasts are resident fibroblasts; other sources of myofibroblasts include epithelial and endothelial cells through epithelial to mesenchymal transition (EMT), endothelial to mesenchymal transition EndoMT, pericytes, fibrocytes and smooth muscle cells (366,367). Myofibroblasts are derived from their precursors in response to inflammatory and fibrotic mediators including growth factors, cytokines, adipokines and chemokines (368,369).

Myofibroblasts modulate inflammation in other tissues as well. In the heart, cardiac myofibroblasts play an important role in immune response through cross-talks with inflammatory cells, and myofibroblasts release factors such as monocyte chemoattractant (MCP-1), CCL5 (RANTES), macrophage inflammatory protein (MIP-1) and Interferon gamma-induced protein 10 (IP-10) that recruit immune cells to the site of injury (350). In liver fibrosis, hepatic myofibroblasts regulate the migration and accumulation of lymphocytes

within inflamed tissue (370,371), with increased levels of myofibroblasts in fibrotic liver correlating with disease severity (372).

Another extensively studied fibrotic condition is lung fibrosis, Experimental studies from IPF patients have shown that bronchial myofibroblasts polarizing CD4+ T cells to Th17, these secrete IL-17 and IL-22 leading to a more pro-inflammatory environment (373). In summary, myofibroblasts are a pathogenic phenotype that drive fibrosis in various tissue types including the intestine, myofibroblasts could be a potential cellular target for anti-fibrotic therapy. However, further studies to gain in depth understanding of specific mechanisms through which myofibroblasts modulate pro-fibrotic pathways are required. There are multiple cell types that acquire a myofibroblast-like phenotype during chronic inflammation and tissue fibrosis, the following sections discuss the different myofibroblast precursors in intestinal fibrosis.

1.5.2.2 Fibroblasts

Fibroblasts are a spindle shaped cell type of stromal origin that make up part of the connective tissue of multiple organs, they are localised in the mucosa and submucosal layers of the intestinal mesenchyme. Characterisation of fibroblasts is challenging as there is no single robust marker for identification of fibroblasts, and typically a combination of markers has been used in the literature (374). The common surface and intracellular markers currently being used include fibroblast specific protein 1 (FSP-1) which is particularly expressed in quiescent fibroblasts, vimentin, Thy1.1 in humans and Thy1.2 in mice (CD90), platelet-derived growth factor receptor-alpha and -beta (PDGFRα/β), podoplanin and FAP (fibroblast activation protein) (375,376). Fibroblasts have been referred to as the main effector cells of fibrosis due to their plasticity, differentiating into myofibroblasts during chronic inflammation (363).

The inflamed intestinal microenvironment is characterised by the presence of cytokines and growth factors such as TGF-β, basic fibroblast growth factor (FGF2), connective tissue growth

factor (CTGF), epidermal growth factor (EGF), insulin-like growth factor I (IGF-I), platelet derived growth factor (PDGF) and pro-inflammatory cytokines, such as interleukins (IL-1β, IL-6, IL-13) and TNF-α, endothelins (ET-1, -2, -3), Angiotensin II and peroxisome proliferator-activated receptor gamma (PPARγ) among others. PPARγ is a nuclear receptor that regulates genes involved in energy metabolism, cell development and differentiation. PPARγ is highly expressed in the intestine and its activation has been shown to downregulate TGF-β/SMAD signalling pathway. Administration of a PPARγ agonist into a DSS mouse model of colitis that had developed colonic fibrosis resulted in a pro-fibrotic outcome.

Macroscopic and microscopic evaluation showed a reduction in expression of pro-fibrotic genes including *ACTA2*, *COL1A1* and *FN1* by 1.48-folds ($P < 0.05$), 1.93-folds ($P < 0.005$), and 1.03-fold ($P < 0.05$). In addition, there was reduction in proteins levels of α-SMA, collagens I and II. The PPARγ activity also decreased expression of IL-13 and connective tissue growth factor (CTGF) by 1.89-folds ($P < 0.05$) and 2.2-folds ($P < 0.005$) respectively. In this study, PPARγ modulation significantly reduced expression of pro-fibrotic markers in TGF-β stimulated human fibroblast cell line (CCD-18Co) and intestinal epithelial cell line (HT29). The study showed anti-fibrotic properties both in vitro and in vivo in which activation of PPARγ inhibited TGF-β/SMAD signalling (377). Fibroblasts have also been reported to undergo activation due to mechanical stress caused by changes in the matrix (378). Overall, pro-inflammatory and pro-fibrotic signalling predominantly mechanisms modulated by TGF-β signaling pathway promote fibroblast to myofibroblast transition (FMT) hence exacerbating tissue fibrosis (366,379).

1.5.2.3 Smooth muscle cells

Intestinal smooth muscle cells (SMCs) are located in the muscularis propria layer of the gut, and these cells replenish the muscle fibres which are important in tissue contraction and

relaxation. Smooth muscle cells express contractile proteins such as alpha-smooth muscle actin (α-SMA), smooth muscle myosin heavy chain (MYH11) and desmin (319). Formation of strictures is characterised by swelling of the intestinal wall and narrowing of the gut lumen with SMCs contributing to this process through muscle hyperplasia and hypertrophy. This was investigated in a study by Chen (2017) using resected gut strictures from CD patients and assessment of histological changes, indicating muscle hyperplasia and hypertrophy across the different layers of the gut, facilitating narrowing of the bowel lumen (380). The concept of muscle hyperplasia and hypertrophy was further shown by Flynn (2010) where primary smooth muscle cells isolated from the muscularis propria layer showed increased proliferation in the presence of insulin growth factor-I and alphaVbeta3 integrin (381). In a TNBS mouse model of colitis, data obtained increased proliferation of SMCs in the presence of PDGF and IGF-1 (319,382). These findings indicate the key role SMCs play in formation of strictures in intestinal fibrosis.

1.5.2.4 Pericytes

Pericytes are cells of mesenchymal stromal cell origin wrapped around endothelial cells; they are located in the linings of tissue micro-vessels such as capillaries. In the intestine, they are predominantly localised within the endothelium in proximity with vascular smooth muscle cells in the muscularis propria. Pericytes express markers such as α-SMA, desmin^{low}, vimentin, PDGFRβ, Neural/glial antigen 2 (NG2), melanoma cell adhesion molecule (MCAM or CD146), angiotensin I/ II, CD80, CD86 and collagen I (279,383). The mechanisms through which pericytes contribute to intestinal fibrosis are not clear, however studies suggest that chronic tissue inflammation results in proliferation, migration from vessels and differentiation of pericytes into a more contractile and collagen secreting phenotype (332) . Studies performed in col1a1-GFP transgenic mice of obstructive kidney fibrosis showed that pericytes differentiated and proliferated into myofibroblasts with increased production of collagen

(Humphreys, 2012; S. L. Lin et al., 2008). Single cell sequencing of kidney tissue showed pericytes as a major source of ECM producing myofibroblasts initiated by TGF-β and PDGF signalling pathways during wound healing (386,387).

1.5.2.5 Fibrocytes

Fibrocytes are bone marrow derived circulating cells of mesenchymal origin with features similar to monocytes. They are identified by their expression of markers such as CD34, CD45, CD14, CD11a, CD68, MHC class II and collagens I/III (388). Upon tissue injury, fibrocytes migrate to the inflamed site in response to inflammatory signals such as CCR2, CCR3, CCR4 and CCR7. They lose their stemness and gain a myofibroblast like phenotype with increased expression of markers such as α-SMA and collagen (389). A study was performed by Kuroda (2019) in a DSS induced mouse model of colitis comparing CCR2 deficient mice and wild type controls. Following the onset of inflammation, CCR2+ bone marrow derived fibrocyte infiltrated the colon with increased expression of TIMP1 in the wild type. Conversely, infiltration of fibrocytes was not observed for the CCR2 deficient mice and resulted in a reduced fibrotic response in these animals. These data suggest that the increased expression of TIMP1 inhibits degradation of ECM proteins hence driving the fibrotic phenotype in the wild type mice (391). In addition, fibrocytes have been implicated in the pathogenesis of intestinal fibrosis and IPF with increase in circulating fibrocytes in resected tissue collected from patients with both diseases. The presence of pro fibrotic factors such as TGF-β, CTGF and IL-13 were shown to modulate change in fibrocytes to a more pro-fibrotic phenotype (388,392). The data strongly suggests that fibrocytes play an important role in the pathogenesis of intestinal fibrosis and further understanding of these cells could present opportunities for a cellular target for anti-fibrotic therapy.

In general, multiple cells in a chronic inflamed microenvironment including fibroblasts, fibrocytes, pericytes, smooth muscle cells, epithelial and endothelial cells are activated in response to various pro-inflammatory and pro-fibrotic mediators, these cells trans-differentiate and acquire myofibroblast-like features with increased ECM production which in turn exacerbates fibrosis (sources of myofibroblast like phenotype in table 1.1). Most of the studies on the specific mechanisms through which these processes occur have been carried out in animal models, data for the progression of human disease remains poor. There is an urgent need to understand the mechanisms through which these heterogenous cell populations mediate inflammation and fibrosis to identify novel targets for intestinal fibrosis which is still a major unmet need.

Fibroblasts, myofibroblasts, pericytes, fibrocytes and smooth muscle cells (SMCs) are interrelated mesenchymal stromal cells; the different cell phenotypes exhibit plasticity as they tend to differentiate to any phenotype within mesenchymal stromal cells (MSCs) depending on tissue demands. In UC, the SMCs contribute to the thickening of the muscularis mucosa while in CD, these cells contribute to the thickening of the bowel wall followed by stricture formation and bowel obstruction. MSCs produce ECM proteins in response to different inflammatory mediators such as TGF-β and IL-1β, thus exacerbating the inflammatory process which eventually results in stricture formation. MSCs are currently being investigated as potential anti-fibrotic therapies in animal models of IPF, these could potentially be translated as cellular therapies in intestinal fibrosis (393,394). Other cellular sources of myofibroblasts are discussed in the next sections.

1.5.2.6 Intestinal epithelial cells

The intestinal epithelial layer plays a key immune regulatory role in maintaining the integrity of the epithelial barrier. Epithelial cells are located in the mucosal layer as part of the epithelium

which is joined by protein complexes referred to as tight junctions that regulate permeability into the intercellular space (395,396). Epithelial cells are identified by their expression of markers such as epithelial cellular adhesion molecule (EpCAM), E-cadherin and cytokeratins CK 18, 19 and 20. EpCAM has also been reported as an important trans-membrane molecule in relevant biological processes such as cell adhesion, cell migration and proliferation (397,398). Intestinal epithelial cells are a source of myofibroblasts during chronic inflammation and therefore implicated in the pathogenesis of intestinal fibrosis (336). In a TNBS induced mouse model of colitis (CD1 mice) exhibiting chronic inflammation and fibrosis, epithelial cells showed increased expression of FSP1+ fibroblast marker compared to control tissue as well as colocalization of E-cadherin and FSP1; the co-localisation indicates EMT. The group that was treated with TNBS and recombinant BMP-7 (rhBMP-7), which is an inhibitor of TGF-β/SMAD signalling showed a statistically significant reduction in total number of FSP1 fibroblasts and the degree of co-localisation with E-cadherin (399).

In the same study, double transgenic mice (Villin-LacZ) in which removal of the STOP cassette by Cre recombinase in epithelial cells at the promoter site (villin) irreversibly activated Rosa-driven LacZ reporter gene expression in all cells of intestinal epithelium was developed; LacZ reporter gene was used to trace the lineage of epithelial cells that underwent EMT; this was to further investigate EMT; fibrosis in these mice was induced by TNBS treatment. Fibrotic colon tissue from TNBS-treated Villin-LacZ mice showed increased frequency of cells with α-SMA, co-localisation of E-cadherin and FSP1 indication EMT as compared to Villin-LacZ mice treated with vehicle alone. This study suggests that treatment with rhBMP-7 both in vivo and in vitro resulted in inhibition of TGF- β-induced EMT; BMP-7 could be a potential therapeutic target for intestinal fibrosis (399,400). In an inflammatory microenvironment, increased expression of pro-inflammatory and pro-fibrotic mediators stimulate the transition of epithelial to mesenchymal cell phenotype. It is well documented that IL-17, a pro-inflammatory cytokine

is increased in colonic mucosa of CD patients. An in vitro study where an epithelial cell line (IEC-6) was stimulated with IL-17 showed a down regulation of E-cadherin and an upregulation of mesenchymal like markers such as vimentin and α-SMA. This indicates transition of epithelial to mesenchymal phenotype with the ability to produce increased levels of α-SMA which is upregulated in fibrosis (181,401).

1.5.2.7 Intestinal endothelial cells

Endothelial cells make up the lining of blood and lymphatic vessels and prevent bacterial antigens from entering the blood stream. In the intestine, endothelial cells are located in the endothelium within the submucosa layer. Endothelial cells are identified by their expression of markers such as CD31, CD34, melanoma cell adhesion molecule (MCAM/CD146), Intercellular Adhesion Molecule 1 (ICAM-1/CD54) and mucosal addressin cell adhesion molecule-1 (MAdCAM-1) (402,403). Endothelial cells facilitate the recruitment of gut homing of leukocytes from peripheral blood circulation into tissue through expression of receptors such as MAdCAM-1 and corresponding integrins such as α4β7 expressed on T cells (404).

During chronic inflammation in IBD, there is increased mucosal ulceration and damage to blood and lymphatic vessels. In response to cytokine and growth factors released into the inflamed microenvironment, endothelial cells trans-differentiate to mesenchymal like cells through a process known as endothelial to mesenchymal transition (EndoMT) (405,406). Studies performed on human endothelial cells cultured and stimulated with IL-1β, TNF-α and TGF-β resulted in change of phenotype to spindle shaped cells with increased expression of α-SMA and FSP-1 proteins. In addition, there was increased production of collagen and fibronectin 9 days post stimulation; this shows EndoMT as an important mechanism in the pathogenesis of intestinal fibrosis (405). Other molecular factors expressed on endothelial cells and have been implicated in tissue fibrosis include transcription factors Snail, Slug, Twist,

Zeb1/2; studies in cardiac fibrosis showed low oxygen levels and shear stress can lead to changes in the microenvironment and activation of SNAI1, SNAI2, TWIST1, ZEB1/2. Activation of these transcription factors led to over expression of endothelial markers such as VEGFR2, PECAM1 and increased expression of mesenchymal cell markers such as α-SMA and Collagens I and II (407,408).

Table 1.2: Cells involved in intestinal fibrosis and their markers of identification

Cell type	Positive markers	Activators	Markers upon activation	References
Fibroblasts	CD34, Vimentin, FSP-1, Collagen 1, Thy1 (CD90), Podoplanin, PDGFRβ	TGF-β, TNF-α	CTGF, α-SMA, Collagen 1A, Fibronectin	(Heinzelmann et al., 2018; Kahounová et al., 2018)
Myofibroblasts	CD34, Vimentin, FSP-1, Collagen 1, Thy1 (CD90), Podoplanin, PDGFRβ, α-SMA	TGF-β, PDGF, CTGF	CTGF, α-SMA, Collagen 1A, Fibronectin	(Eyden, 2008; Younesi et al., 2021)
Smooth muscle cells	CD34, Vimentinlow, α-SMA, Desmin, Collagen	IL-1β, TNFα, and IFN-γ	α-SMA, Collagen 1A, Fibronectin	(Latif et al., 2015; M. Liu & Gomez, 2019)
Pericytes	α-SMA, Desminlow, Collagen, RGS5, podoplanin, PDGFRβ, CD90, Vimentin	TGF-β, PDGF, Wnt	ADAMTS-1	(Baek et al., 2022; T. I. H. Park et al., 2016)
Epithelial cells	Cytokeratins, EpCAM, E-cadherin	TGF-β, Wnt, mTOR, HIF-1a	FSP-1, α-SMA	(Giotakis et al., 2021; Haunerdinger et al., 2021)
Endothelial cell	CD31, N-cadherin	IFN-γ, TNFα	VCAM-1, ICAM-1	(Goncharov et al., 2020; Grant et al., 2021)
Fibrocytes	CD45, CD34, CD11, CD14, CD80, CD86	TNFα, IL-1β, IL-6	MHC II, CCR4, CCR5	(de Oliveira & Wilson, 2020; Suga et al., 2014)

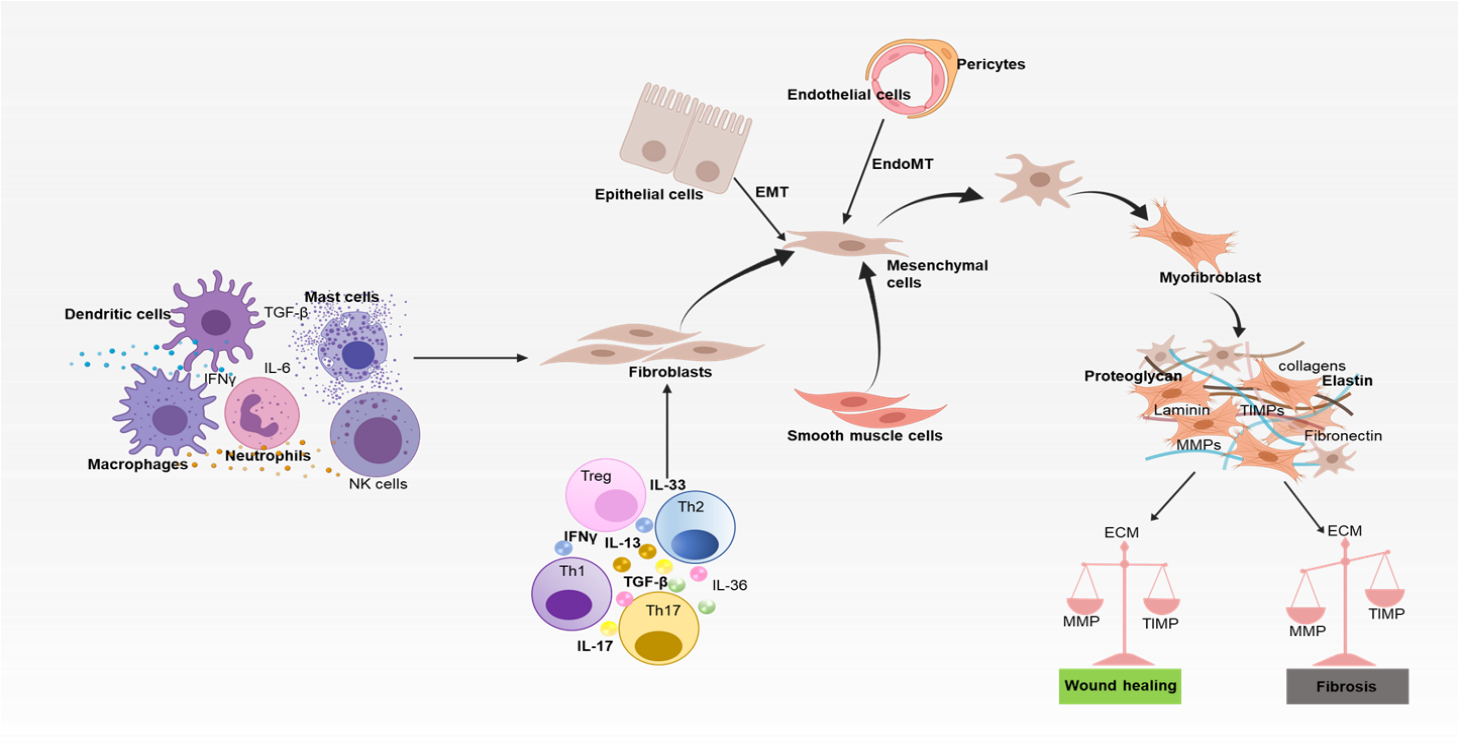


Figure 1.6: Cellular mechanisms that drive the fibrotic phenotype in intestinal fibrosis. Innate immune cells (dendritic cells, mast cells, macrophages, neutrophils and NK cells) and adaptive immune cells (Th1, Th2 and Th2) secrete cellular mediators that activate fibroblasts to differentiate into myofibroblasts that secrete pathogenic ECM. Stromal cells including smooth muscle cells and fibroblasts differentiate to myofibroblasts with increased deposition of ECM. Epithelial cells undergo Epithelial to Mesenchymal transition (EMT) while endothelial cells undergo endothelial to mesenchymal transition (EndoMT), ECM proteins with a balanced production and degradation program result in wound healing, whereas a dysregulated program results in tissue fibrosis. (*Image was designed using graphical elements from BioRender*)

1.6 Molecular mediators that drive stricture formation in intestinal fibrosis

Wound healing is a physiological process that the immune system initiates to resolve ulcerations caused due to chronic inflammation, the process results in restoration of tissue integrity and structure. Various molecular mechanisms are involved in regulating the well-balanced deposition of ECM and degradation by MMPs whose activity is inhibited by TIMPs (304,409). In CD, the wound healing process is dysregulated resulting in activation of stromal cells that deposit excess ECM proteins with an impaired degradation program; the highest producers of ECM proteins are the myofibroblasts. The ECM proteins such as collagens I, III, IV, V, and VI, glycoproteins and proteoglycans such as fibronectin, laminin, and tenascin have an altered stiffness that results in scarring and formation of strictures (275,364). Various profibrotic mediators activate signaling pathways that are key in activation and differentiation of these stromal cells, particularly fibroblasts to myofibroblasts transition (FMT). These molecular mediators also play a key role in proliferation and recruitment of both immune and stromal cells to the site of injury, hence exacerbating the pro-inflammatory and profibrotic microenvironment (410).

Studies have shown efficiency of the current anti-inflammatory therapies in resolving inflammation; however, it does not stop progression of fibro-stenosing disease; rather it becomes an inflammation independent driven process (411). Some of the key molecular mediators include cytokines such as IL-1, IL-4, IL-6, IL-13, IL-17, IL-12, IL-23, IL-33, TNF-α, MCP-1 and IFN-γ (412,413) . The mechanisms through which cytokines regulate stricture formation are discussed in section 1.5.1. Other mediators such as TGF-β1, CTGF, IGF, PDGF, ET, ROS, Transglutaminase 2 (TG2), MMPs and TIMPs drive the fibrotic response leading to a highly cross-linked extracellular network filled with growth factors, damage associated molecular pattern (DAMPs) proteins and other drivers of local immune responses (414). One of the key mediators of fibrosis is TGF-β1 (TGF-β), it is referred to as the master regulator of

tissue fibrosis across multiple organs (408). Increased levels of TGF-β and its cognate receptors I and II have been reported in fibrostenotic CD (415). TGF-β in its active form binds TGF-β I and II receptors, downward signaling results in upregulation of profibrotic genes such as PAI-1, COL1A1, FN1, FAP and PPAR-γ. The non-canonical signaling pathway involves Erk, SAPK/JNK and MAPK pathways while the canonical pathway involving SMAD proteins 2/4 and co-factor SMAD4 as shown in figure 1.7.

Phosphorylated SMAD2/3/4 complex relocate to the nucleus where they initiate transcription of target genes that are profibrotic; these include upregulation of genes that encode for profibrotic mediators such as CTGF, collagen, fibronectin, α-SMA and a down regulation of genes that encode for proteins important in regulating excess ECM deposition such as MMPs. In the mucosa of CD patients, increased levels of TGF-β, phosphorylated SMAD2/3 and increased collagen production in the stricture regions have been observed compared to the non-stricture regions (416). The TGF-β signaling pathway is regulated by SMAD7, findings from various studies in the intestine, skin, kidney, lung and liver fibrosis have shown that increased expression of inhibitory SMAD7 significantly reduced fibrosis (417,418).

A study carried out in a TNBS fibrosis induced mouse model in which SMAD7 antisense oligo-nucleotides that knocked down SMAD7 expression showed reduced collagen production and fibrosis. These findings contradict other studies that have shown TGF-β as a master regulator of fibrosis. The study indicates that SMAD7 inhibition of TGF-β signaling diminishes its anti-inflammatory mechanisms and therefore knocking down SMAD7 increases TGF-β signaling and its anti-inflammatory response (419). On the contrary, studies performed in a rat model of renal fibrosis in which SMAD7 gene insertion was carried out using the adenoviral system prevented development of unilateral ureteral obstruction whereas controls developed renal

fibrosis, suggesting SMAD7 could be used as a potential therapeutic target that can antagonize the TGF-β signaling pathway (420).

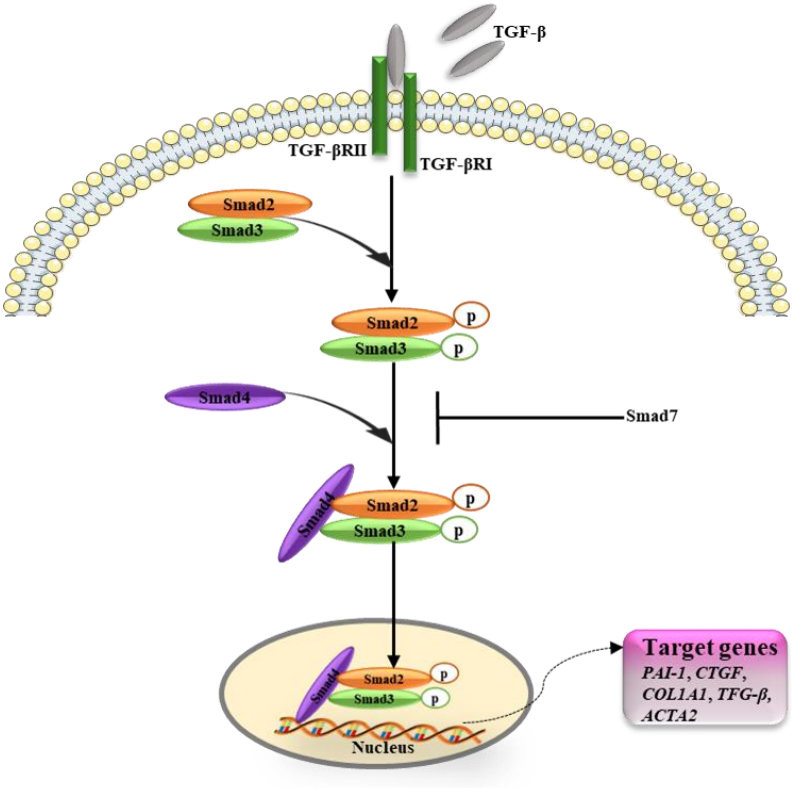


Figure 1.7: Canonical TGF-β/SMAD signaling pathway. TGF-β binds its cognate receptors I and II to activate phosphorylation of receptor-regulated Smads 2/3 (R-SMADs), pSMAD 2/3 interact with co-factor SMAD (SMAD4) to form the R-SMAD/SMAD4 complex, which associates with specific transcription factors and cofactors to modulate the expression of its target genes such as *SERPINE1*, *CTGF*, *COL1A1* and *ACTA2*. The pathway is inhibited by activation of inhibitor (I-SMAD) SMAD7 (*Image drawn from elements collected from www.servier.com*)

Another key mediator of fibrosis is CTGF, its expression is increased in the presence of TGF-β. CTGF upregulated in CD patients specifically in fibroblasts within the submucosal layer and in areas of intense damage as compared to UC patients (417,421). Studies show profibrotic mechanisms of CTGF through modulating proliferation and activation of fibroblasts to produce ECM proteins (422). Other growth factors that are important in driving the profibrotic phenotype include PDGF, bFGF (FGF2) and EGF. PDGF (PDGF-A or PDGF-B) signals

through its cognate receptors PDGFRα and PDGFRβ respectively (423,424). The presence of PDGF in the intestine has been reported to cause activation, proliferation and migration of cells such as fibroblasts, myofibroblasts and increased deposition of ECM proteins (425). The fibrotic microenvironment is complex, cytokines and growth factors initiate signaling and activation of pathways that maintain the fibrotic phenotype. Some of the key pathways that have been implicated in intestinal fibrosis include Wnt/ β catenin signalling pathway and lysyl oxidase.

The β-catenin Wnt signalling pathway controls intestinal epithelial stem cell proliferation and differentiation capacity. In the intestine, the Wnt pathway is involved in homeostasis of the gut epithelium and cell regeneration; in cases where the Wnt signalling is impaired damage of the intestinal crypts has been reported (426,427). Lysyl oxidases (LOX) are copper dependent amine oxidase enzymes involved in the cross linking and manufacture of the extracellular matrix proteins collagen, elastin and fibronectin. Of the five LOX proteins, LOXL 2 has been studied extensively and targeted as a potential antifibrotic in liver fibrosis (428). In IBD, a study was carried out that showed that myofibroblasts isolated from stenotic small bowel expressed higher levels of LOX enzymes when compared to normal gut tissue derived from the same patient (298). There is need to perform more mechanistic studies to ascertain the role of LOX proteins on collagen crosslinking and their effect on fibroblasts in intestinal fibrosis.

1.7 Gaps in current knowledge of intestinal fibrosis and the unmet need

Although there has been progress in development of anti-inflammatory therapies which have been effective in the management of CD, evidence has shown that it does not stop the natural progression of fibrostenosing disease (429). The mechanisms involved are complex and involve interactions between cells, extracellular matrix proteins, growth factors and cytokines. There is evidence that shows reversibility of fibrosis, studies have focused on inhibition of

TGF-β signalling and its downstream effects of activation of fibroblasts to myofibroblasts. A study was performed in a skin fibrosis bleomycin induced mouse model, in which the treatment group of animals were treated with a pyridine derivative compound (HPH-15) that inhibited phosphorylation of SMAD3. The treated mice experience reduction in immune cell infiltration and reduced frequencies of profibrotic macrophages, suggesting that targeting the SMAD pathway may be beneficial in this setting (430).

Another study in a unilateral ureteral obstruction (UUO) mouse model of kidney fibrosis was performed. In this study, SMAD7 overexpression plasmids were delivered into the kidney. The data showed a reduction in expression of profibrotic markers in the mice that received SMAD7 gene therapy as compared to the controls. The data gives an indication that inhibition of TGF-β signalling reduces attenuates fibrosis in these mice (431). Despite in vivo and in vitro studies that have been performed in understanding the various cellular and molecular mechanisms that drive the fibrotic phenotype, there is still no approved anti-fibrotic therapy for intestinal fibrosis in humans. Anti-fibrotic drugs used to treat fibrosis in other diseases such as pirfenidone for the management of IPF have been studied for efficacy in fibrostenosing CD; the data showed anti-proliferative effects and a reduction in collagen production by primary intestinal fibroblasts from patients with fibrostenosing CD. Pirfenidone could be repurposed for the management of fibrostenosing CD (432–434).

The challenges being faced in identification of novel targets for anti-fibrotic therapy are mostly because of the heterogenous nature of the inflammatory and fibrotic microenvironment in CD. There is a knowledge gap in understanding the specific mechanisms that drive intestinal fibrogenesis. In addition, as intestinal fibrosis is a slow progressing disease it has been difficult to detect early fibrosis with intestinal strictures being diagnosed at later stages where it has already progressed to loss of tissue function. Although there are various animal models used in

investigation of disease mechanisms (435), the challenges of in vivo studies in animals are a limitation and access to patient resection tissue is limited, hence the need to establish in vitro models that can be used to study the mechanisms in intestinal fibrosis. There is need to further understand the specific cellular and molecular mechanisms that are key in driving the fibrotic phenotype in intestinal fibrosis. For example, TGF-β signalling is a known pathway that drives fibrosis, despite the many attempts to disrupt signalling of this pathway, there has been limited success. Are there molecular mediators that can be targeted to indirectly to disrupt TGF-β signalling? Are there *in vitro* models that can be used to study mechanisms that drive intestinal fibrosis? Understanding these key questions present opportunities for novel target discovery and potential anti-fibrotic therapy in stricturing CD.

The overall aim of this project was to understand the cellular and molecular mechanisms that drive the fibrotic phenotype in CD strictures. The availability of new technologies such as single cell sequencing has made it possible to interrogate cell populations in more depth. In addition our lab is able to access fresh human resection tissue with consent from CD patients undergoing resection surgery at the Queen Elizabeth hospital, Birmingham. This presents a unique opportunity to perform studies on fresh human samples collected from patients, the results obtained would be directly applicable to patients and directly translatable to potential targets for anti-fibrotic therapy. In this study, we hypothesise that the stricture segment has a unique profibrotic molecular and cellular signature as compared to the non-stricture segments. In depth understanding of the stricture associated phenotype will provide insight into potential targets for anti-fibrotic therapy in CD. In addition, the study focuses on fibroblasts as studies have shown fibroblasts as the main cell type that produce and maintain ECM hence driving the fibrotic phenotype in intestinal fibrosis (363,378,436). Below are the specific aims of the project;

1.8 Aims for the study;

1. To study the transcriptome and proteome profiles of matched tissue from the stricture, proximal and distal segments in CD patients
2. To study fibroblast expression patterns relative to identified markers FGF2, HDAC1 and Gremlin1 in the stricture, proximal and distal segments in CD patients
3. To study the role of fibroblast derived gremlin 1 in the formation of strictures in intestinal fibrosis,

CHAPTER 2

MATERIALS AND METHODS

2.1 Sample collection, transportation and processing

2.1.1 Ethical approval for use of resected tissue samples

Ethical approval to obtain gut resection samples, biopsies and blood from IBD patients from the Queen Elizabeth Hospital was granted under the tissue remodelling and fibrostenosing in Crohn's disease (TRAFIC) study (Approval reference number RRK6508). Samples were collected and stored in accordance with the University of Birmingham's code of conduct and Human Tissue Authority (HTA) guidelines. All staff involved in the TRAFIC study completed training on details pertaining to patient eligibility, sample collection and data handling prior to commencing work on the TRAFIC study.

2.1.2 Sample collection and transportation to the lab

Samples collected included whole blood and gut resection tissue from IBD patients with fibrostenotic disease who provided consent to the TRAFIC study. Whole blood, matched to resected tissue, was collected before or on the day of the surgery depending on what was most convenient for the clinical staff and the patient. Blood was collected in vacutainer tubes containing EDTA anticoagulant. On the day of the surgery, gut resection samples were collected and immediately transported to the clinical pathology department by clinical staff. The pathologist then examined the resected tissue and decided if it was suitable to be sampled for the study. From each patient, the aim was to collect tissue segments that were distal and proximal to the stricture in addition to the stricture itself as illustrated in figure 2.1B. The size of tissue sampled ranged between 0.668-3.5 grams depending on the size of the total resection received at pathology. Samples were placed into MACS tissue storage solution (Miltenyi Biotec, Germany), and transported immediately to the lab for further processing.

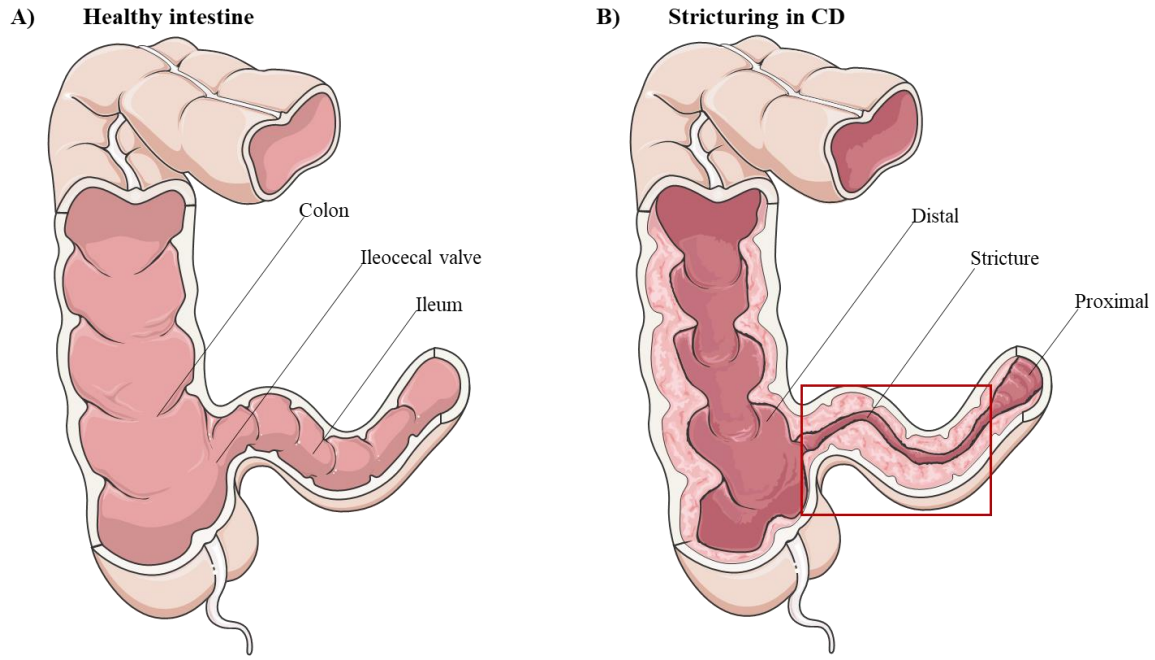


Figure 2.1: Representative images showing location of sampled segments. A) Healthy intestine showing the different segments prior to stricture formation. The segments include the large intestine (colon), the ileocecal valve, which is a sphincter muscle that separates the colon and ileum and the small intestine (ileum). **B)** Image shows regions from which the samples were collected from: the distal segment is mainly taken from the colon, a region distal to the stricture, the stricture segment was taken from the fibrotic region which was common around the ileocecal region in the ileum, and the proximal segment was taken from a region proximal to the stricture, generally located in the ileum- Image adapted from *https://smart.servier.com*

2.1.3 Sample receipt and processing in the lab

Gut resection and blood samples were pseudo-anonymised on arrival at the lab by assigning each a unique access number. Blood was processed to collect peripheral blood mononuclear cells (PBMCs) immediately for use in further experiments using the methodology described below in section 2.2.1. Resection samples were weighed and depending on how much tissue was available, were further cut for various procedures. Figure 2.2 shows the different procedures that were carried out using the tissue collected, tissue was placed in RNAlater for RNA extractions and part of it placed in formalin to generate FFPE sections. Tissue was also used for isolation of fibroblast. The number of procedures carried out varied depending on the amount of tissue supplied to the lab.

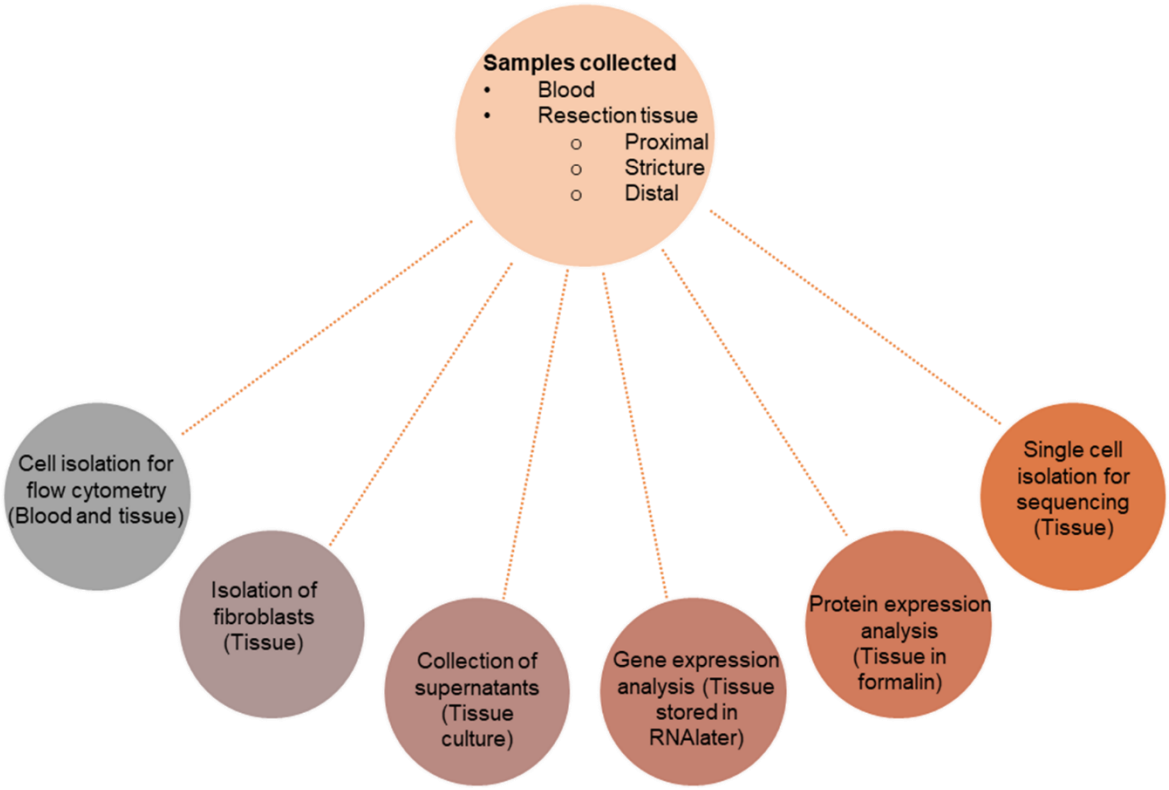


Figure 2.2: Distribution of gut resected tissue (proximal, stricture and distal) and blood for various experimental procedures. Cells were isolated from whole blood and from tissue samples for cellular characterisation using flow cytometry. Tissue samples were also used for isolation of fibroblasts. Tissue was left in culture and supernatants were collected to examine secretion of various mediators. Tissue was stored in RNA later and formalin for gene and protein expression analysis studies, respectively, and cells were isolated from tissue and used for single cell sequencing.

2.1.4 Resected gut tissue samples received and analysed in the whole study

A total of 25 samples from 25 individuals were collected for the study, but that the samples that were able to be taken varied between individuals (see Table 2.1). For example, we received blood for 5 patients which did not have matching tissue samples, but for 12 patients we were able to collect both blood and tissue from all three sites.

Table 2.1: Summary of gut resected samples delivered to the lab for in vitro experiments (Number of samples: Sample type e.g blood, proximal, stricture and distal)

Total number of samples	Blood	Proximal	Stricture	Distal
5	X			
2		X	X	X
2		X	X	
3	X	X	X	
1	X	X		X
12	X	X	X	X

2.2 Primary cell isolation

2.2.1 Isolation of Peripheral Blood Mononuclear Cells (PBMCs) from blood

Whole blood collected in EDTA tubes was used for the isolation of cells to study the immune profile of CD patients undergoing resection surgery. Whole blood was layered onto Ficoll-Paque (Sigma-Aldrich, Gillingham, UK) at a ratio of 2 parts of blood to 1 part of Ficoll-Paque. The layered mixture underwent centrifugation at 800 x g for 20 minutes with an acceleration rate of 5 and a deceleration rate of 2. The buffy coat containing PBMCs that formed between the Ficoll-Paque and plasma layers was carefully aspirated, cells were washed twice in sterile phosphate buffered saline (PBS) and then counted using a Neubauer haemocytometer. For counting, cells were mixed at a ratio of 1:10 with Trypan-blue (Sigma-Aldrich, Gillingham, UK) to ensure only viable cells were counted. Cells were then used for planned experimental procedures.

2.2.2 Isolation of primary fibroblasts

The study was focused on primary fibroblasts isolated from resection samples collected from CD patients undergoing resection surgery. However, access to resected tissue samples was limited; the centre for liver and gastroenterology research (CLGR), Birmingham had a more reliable access to explant liver samples. Therefore, explant liver was used for initial optimisation.

2.2.2.1 Isolation of primary liver fibroblasts

Approximately 20-30 g of diseased explanted liver tissue was used for isolation of liver fibroblasts following an established protocol in the lab (370). Specifically, tissue was collected and sliced into approximately 3-5 mm pieces using a scalpel and the sliced tissue was washed in ice cold PBS (Sigma-Aldrich, Gillingham, UK) to remove any residuals of blood. Tissue was transferred into a beaker and digested using an enzyme cocktail of collagenase IV (4,000 U/mL) and collagenase 1A (100 µg/mL) (Sigma-Aldrich, Gillingham, UK) in DMEM media (ThermoFisher Scientific, UK). Tissue was placed in a shaker at 370C for 20-30 minutes depending on the stiffness which was determined during slicing of the tissue.

The tissue digest was then filtered through a 70 µm strainer (Miltenyi Biotec, Germany) and cells layered onto a 77% and 33% Percoll gradient before being centrifuged to sediment debris and erythrocytes at 800 x g for 20 minutes with an acceleration rate of 5 and a deceleration rate of 2. The buffy coat containing leukocytes and stromal cells was collected, washed and resuspended in 5 mL of complete fibroblast media made up of DMEM supplemented with 100U/mL Penicillin (Gibco, Invitrogen, United Kingdom), 100 µg/mL streptomycin (Gibco, ThermoFisher, United Kingdom) and 2 mM of L-glutamine (Gibco, ThermoFisher, United Kingdom) and 16% heat-inactivated fetal bovine serum (FBS) (Sigma-Aldrich, Gillingham, UK). The cells were then transferred to a T25 flask (Corning, New York, USA) at 2X106 cells/mL. Cells were left in culture at 370C, 5% CO2 while changing media every 2-3 days, fibroblast growth was monitored over 10-14 days.

2.2.3 Isolation of primary gut fibroblasts

The gastroenterology group was being set up just as I was starting my PhD, thus there were no established protocols on handling gut samples. For isolation of human primary gut fibroblasts,

I had to develop a protocol which involved several steps of optimisation, detailed below and also discussed in chapter V of this thesis.

2.2.3.1 Protocol A: Isolation of gut fibroblasts using the 'explant liver' protocol

Similar to the 'explant liver' protocol described above, resected gut tissue was collected and sliced into small pieces approximately 3-5 mm using a scalpel. The sliced tissue was washed in ice cold PBS (Sigma-Aldrich, Gillingham, UK) to remove any residuals of blood. Tissue was transferred into a 50 mL falcon tube (ThermoFisher Scientific, UK) for digestion in an enzyme cocktail of collagenases IV (4,000 U/mL) and 1A (100 µg/mL) (Sigma-Aldrich, Gillingham, UK) in DMEM media (ThermoFisher Scientific, UK). Tissue was placed in a shaker at 370C for 20-30 minutes depending on the stiffness which was determined during slicing of the tissue. The tissue digest was then filtered through a 70 µm strainer (Miltenyi Biotec, Germany) and cells layered onto a 77% and 33% Percoll gradient before being centrifuged to sediment debris and erythrocytes at 800 x g for 20 minutes with an acceleration rate of 5 and a deceleration rate of 2. The buffy coat containing leukocytes and stromal cells was collected, washed and resuspended in 5 mL of complete fibroblast media supplemented with 100U/mL Penicillin (Gibco, Invitrogen, United Kingdom), 100 µg/mL streptomycin (Gibco, Invitrogen, United Kingdom) and 2 mM of glutamine (Gibco, Invitrogen, United Kingdom) and 16% heat-inactivated fetal bovine serum (FBS) (Sigma-Aldrich, Gillingham, UK). The cells were then transferred to a 6 well tissue culture plate (Corning, New York, USA) at 2X106 cells/mL. Cells were left in culture at 370C while changing media every 2-3 days, fibroblast growth was monitored over 10-14 days.

2.2.3.2 Protocol B: Isolation of gut fibroblasts using the organ culture protocol

A different approach was tested to isolate primary gut fibroblasts as previously described by (Sawant et al., 2016). In this case gut resected tissue from the proximal, distal and stricture

segments was transferred into a 6 well plate (Corning, New York, USA) and washed with ice cold PBS (Sigma-Aldrich, Gillingham, UK) to get rid of the excess blood residue, then cut into approximately 3-5 mm pieces using a scalpel, after which pieces of tissue were transferred to another 6 well plate and cultured in 2 mL of fibroblast media which was made up of DMEM supplemented with antibiotics including 100U/mL Penicillin (Gibco, Invitrogen, United Kingdom), 100 µg/mL streptomycin (Gibco, Invitrogen, United Kingdom), 50 µg/mL of gentamicin (ThermoFisher Scientific, UK) and 0.25 µg/mL amphotericin B (Sigma-Aldrich, Gillingham, UK). In addition, the media contained 2mM of L-glutamine (Gibco, Invitrogen, United Kingdom) and 16% heat-inactivated fetal bovine serum (FBS) (Sigma-Aldrich, Gillingham, UK). Culture conditions were at 370C, 5% CO2. Fibroblast media was replaced every 3-4 days while monitoring cell growth from the tissue.

2.2.3.3 Protocol C: Isolation of gut fibroblasts using an enzymatic digestion protocol

Another method for fibroblast isolation was performed based on enzymatic digestion using Liberase. For this method, resected gut tissue was washed twice in epithelial layer stripping buffer made up of 1X HBSS (Sigma-Aldrich, Gillingham, UK), 10 mM of HEPES (Sigma-Aldrich, Gillingham, UK), 5% FBS, 100 U/mL penicillin, 100 µg/mL streptomycin and 2 mM of EDTA. Stripping of the epithelial layer was carried out to enrich the stromal cell component of the tissue while reducing the epithelial cell component. Resected gut tissue was then placed in a 6 well culture plate and cut into approximately 3-5mm2 pieces using a scalpel. Tissue was rinsed with cold PBS (Sigma-Aldrich, Gillingham, UK) and placed in a 50 mL Falcon tube containing the epithelial layer stripping buffer. Tissue was left to incubate at 370C for 15 mins while shaking at 250rpm, this was performed twice. The resulting tissue was pelleted by centrifugation at 200 x g for 5 minutes. The pelleted tissue was further cut into smaller pieces of approximately 3-5 mm in a 6 well culture plate and transferred to gentleMACS™ C tubes containing enzymatic digestion media made up of DMEM, 2.5% FBS (Sigma-Aldrich,

Gillingham, UK), 100U/mL Penicillin (Gibco, Invitrogen, United Kingdom), 100 µg/mL streptomycin (Gibco, Invitrogen, United Kingdom), 50 µg/mL of gentamicin (ThermoFisher Scientific, UK) and 0.25 µg/mL amphotericin B (Sigma-Aldrich, Gillingham, UK). In addition, the media contained 100 µg/mL of DNAse (ThermoFisher Scientific, UK) and 100µg/mL of Liberase Thermolysin Low enzyme (Roche, Basel, Switzerland). The tissue was incubated at 370C for 30 mins while shaking. Enzymatic activity was quenched with 10% FBS in DMEM and a quick cycle of further tissue digestion was performed on the gentleMACS™ Octo dissociator with heaters (Miltenyi Biotec, Germany) for 1 minute (C tube protocol). The resultant tissue digest was filtered through a 70 µm filter into a 50mL falcon tube. The single cell suspension was centrifuged at 200 x g for 5 minutes and cells counted in trypan blue (1:10) dilution. Cells were left in culture at a dilution of 2x106 cells/mL of fibroblast media which was changed every 3-4 days while monitoring for fibroblast growth.

2.2.3.4 Protocol D: Isolation of gut fibroblasts using a hybrid (mechanical and enzymatic digestion) protocol

A fourth protocol was tested for fibroblast isolation combining enzymatic and mechanical digestion. In this protocol, resected gut tissue was washed with cold PBS (Sigma-Aldrich, Gillingham, UK), tissue was dissected into 3-5 mm pieces using a scalpel. A wash was performed using cold PBS to get rid of the excess blood residue and then the tissue was transferred to a 50 mL falcon tube containing enzymatic digestion media made up of DMEM, 2.5% FBS (Sigma-Aldrich, Gillingham, UK), 100U/mL Penicillin (Gibco, Invitrogen, United Kingdom), 100 µg/mL streptomycin (Gibco, Invitrogen, United Kingdom), 50 µg/mL of gentamicin (ThermoFisher Scientific, UK) and 0.25 µg/mL amphotericin B (Sigma-Aldrich, Gillingham, UK). In addition, the media contained 100 µg/mL of DNAse (ThermoFisher Scientific, UK) and 100µg/mL of Liberase Thermolysin Low enzyme (Roche, Basel, Switzerland. Tissue was digested at 370C while constantly dissociating on the gentle MACS™

Octo dissociator with heaters (Miltenyi Biotec, Germany) for 30 mins on a pre-set program. Digested tissue was filtered through a 70 µm strainer (Miltenyi Biotec, Germany), and the resulting single cell suspension was centrifuged at 200 x g for 5 mins at acceleration and deceleration speed of 9. The cells were washed once and re-suspended in DMEM (Sigma-Aldrich, Gillingham, UK) supplemented with 16% heat inactivated FBS in a 6 well plate at a concentration of 2.5x106 cells/mL and incubated at 370C and 5% CO2.

Table 2.2: Summary of steps used during optimisation protocol for isolation of gut derived primary fibroblasts

Steps in the fibroblast isolation process	Protocol A: "Explant liver" protocol	Protocol B: Organ culture protocol	Protocol C: Enzymatic digestion	Protocol D: Hybrid protocol
Wash of tissue	Yes	Yes	Yes	Yes
Dissection of tissue	Yes	Yes	Yes	Yes
Digestion	Enzymatic	No	Enzymatic	Enzymatic and mechanical
Wash of cells	Yes	Yes	Yes	Yes

Table 2.3: Culture media and reagents used in cell isolation and culture

Media and reagents	Company
Roswell park memorial institute 1640 medium [+] L-glutamine	Sigma-Aldrich
Dulbecco's Modified Eagle Medium (DMEM)	Thermo Fisher
Penicillin streptomycin solution	Sigma-Aldrich
Gentamicin	Sigma-Aldrich
Amphotericin	Sigma-Aldrich
Normocin	Invivogen
Dispase	Stemcell Technologies
L-glutamine	Sigma-Aldrich
Phosphate-buffered saline [+] $MgCl_2$ (1000mg/L), $CaCl_2$ (100 mg/L)	Sigma-Aldrich
Ethylenediaminetetraacetic acid solution (EDTA)	Sigma-Aldrich
Heat inactivated fetal bovine serum (FBS)	Sigma-Aldrich
Phosphate buffered saline	Sigma-Aldrich
Bovine serum albumin (BSA)	Thermo Fisher
Fibroblast growth medium	Lonza
Hank's Balanced Salt Solution (HBSS) without Ca2+ and Mg2+	Sigma-Aldrich
(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) (HEPES)	Sigma-Aldrich
DNAse	Thermo Fisher
Collagenase 1A	Sigma-Aldrich
Collagenase IV	Sigma-Aldrich
Liberase TL	Roche
Ficoll-Paque	Sigma-Aldrich

2.3 Flow cytometry

2.3.1 Surface staining

Freshly isolated single cells in suspension from resected gut tissue and peripheral blood were distributed into flow cytometry (FACS) tubes (Fisher Scientific, UK) at $1x10^6$cells/mL. Cells were initially labelled with 100 µl of the Zombie Aqua viability dye (Biolegend, San Diego, USA) diluted at 1:500 at room temperature for 20 minutes in the dark. Cells were washed in 2% FBS diluted in PBS (Sigma-Aldrich, Gillingham, UK) and further labelled with surface antibodies against cell surface antigens for Immune and stromal cells (T cells, B cells, fibroblasts, epithelial and endothelial cells). All antibodies used are listed in Table 2.4. Controls included cells only (cells without antibodies), isotype controls and all fluorochromes minus one (FMO) control, especially for markers of interest with low antigen expression. These were used to set gates above background levels during data analysis. Cells were incubated at 4oC for 30 minutes, then washed with 2% FBS + PBS. Cells were fixed with 100µl 4% Paraformaldehyde (PFA) for 20 minutes at room temperature in the dark. A final wash was performed with 1mL of 2% FBS + PBS, and the cells were then re-suspended in 300µl of 2% FBS + PBS and analysed using the BD LSRFortessa™ (BD Bioscience, New Jersey, USA) for flow cytometry, data was analysed using FlowJo version 10.8.1.

2.3.2 Intracellular staining

To study cytokine expression by tissue infiltrating and peripheral blood mononuclear cells, intracellular staining and subsequent flow cytometry analysis was performed. Freshly isolated cells from resected gut tissue and peripheral blood were distributed in 24 well plates at 2 x 106 cells/mL and left overnight in the incubator at 37°C and 5% CO2 to recover and settle. On the day of staining, cells were stimulated with 30nM PMA (Sigma-Aldrich, Gillingham, UK), 200nM ionomycin (Sigma-Aldrich, Gillingham, UK) and 4µg/mL Brefeldin A (Sigma-Aldrich, Gillingham, UK) for 4 hours. Cells were transferred to FACS tubes and washed with

1 mL of 2% FBS diluted in PBS and labelled with 100 µl of the Zombie Aqua viability dye (Biolegend, San Diego, USA) diluted at 1:500 at room temperature for 20 minutes in the dark. The viability dye was washed off using 2% FBS diluted in PBS. Cells were then fixed with 500µl of fix/perm buffer (BD Bioscience, New Jersey, USA) for 50 minutes at 40C in the dark. A wash step was performed using 2% FBS diluted in PBS, a second wash step was performed using 1X BD permeabilization buffer (BD Bioscience, New Jersey, USA) diluted at (1:10 10x Perm buffer: dH2O), and a centrifugation step was performed for 5 minutes at 500 x g with both acceleration and deceleration set to 9. The cells were stained with intracellular antibodies targeting IL-17, TNF-α, IFN-γ, IL-10 and IL-2 or isotype control for 30 minutes at 40C in the dark (details listed in table 2.4). Antibodies were prepared in the 1X BD Perm/Wash buffer. The stained cells were then washed with perm buffer followed by 2% FBS diluted in PBS; and re-suspended in 300 µl of 2% FBS diluted in PBS ready for flow cytometry using the BD LSRFortessa™ (BD Bioscience, New Jersey, USA).

Table 2.4: List of antibodies used for flow cytometry staining for immune cell, cytokine and stromal cell panels *(Abbreviations: APC: Allophycocyanin, PE: phycoerythrin, FITC: Fluorescein isothiocyanate, CY: Cyanine, BV: Brilliant violet, Rec: Recombinant, IgG: Immunoglobulin G, CD: Cluster of differentiation, k:kappa, Ms:Mouse) Note: Initial and final concentrations were not shown as some were missing on the product website*

Immune cell panel					
Target	**Fluorochrome**	**Clone**	**Dilution**	**Isotype control**	**Company**
CD45	APC CY 7	2D1	1:50	Ms IgG1, κ	Biolegend
CD19	BV650	HIB19	1:50	Ms IgG1, κ	Biolegend
CD8	BV605	SK1	1:100	Ms IgG1, κ	BD Biosciences
CD3	BV421	UCHT1	1:100	Ms IgG1, κ	BD Biosciences
CD4	APC	REA623	1:50	Rec. human IgG1	Miltenyi Biotec
CD25	PECF594	M-A251	1:50	Ms IgG1, κ	BD Biosciences
CD127	PECY7	eBioRDR5	1:50	Ms IgG1, κ	Invitrogen
CCR6	BV711	G034E3	1:100	Ms IgG2b, κ	Biolegend
CXCR3	PerCP CY5.5	11A9	1:50	Ms IgG1, κ	BD Biosciences

Cytokine panel					
IL-10	PE	JES3-9D7	1:50	Rat IgG1, κ	Biolegend
TNF-α	BV450	MAb11	1:40	Ms IgG1, κ	Invitrogen
IFNγ	BV650	4S.B3	1:100	Ms IgG1, κ	BD Biosciences
IL-17	BV785	BL168	1:50	Ms IgG1, κ	Biolegend
IL-2	APC	MQ1-17H12	1:100	Rat IgG2a, κ	eBioscience
Stromal cell panel					
CD45	APC CY7	2D1	1:50	Ms IgG1, κ	Biolegend
HLA-DR	BV450	LN3	1:50	Ms IgG2b, k	Invitrogen
Podoplanin	PerCP CY5.5	NZ-1.3	1:50	Rat IgG2a	Invitrogen
CD90	PECY7	5E10	1:50	Ms IgG1, k	Invitrogen
EpCAM	PECF594	CO17-1A	1:100	Ms IgG2a, κ	Biolegend
CD34	FITC	561	1:50	Ms IgG2a, κ	Biolegend
CD31	APC	REA730	1:50	Rec. human IgG1	Miltenyi Biotec
PDGFRβ	PE	PR7212	1:50	Ms IgG1	Bio-techne

2.4 In vitro cell culture assays

2.4.1 Cell lines used in this study

The normal human colon derived fibroblast cell line CCD-18Co (cat. No. CRL-1459™) and the human colon derived fibroblast cell line from an Adenocarcinoma donor Hs 255T (CRL-7213™) were obtained from the American Type Culture Collection (ATCC) (Manassas, VA, USA).

2.4.1.1 Cell line culture

The CCD-18Co cell line was cultured at 37ºC and 5% CO2 in complete media (Eagle's Minimum Essential Medium (EMEM) (ATCC, Manassas, VA, USA) supplemented with 10% FBS (Sigma-Aldrich, Gillingham, UK), 100U/mL Penicillin (Gibco, Invitrogen, United Kingdom) and 100 µg/mL streptomycin (Gibco, Invitrogen, United Kingdom). Cells were sub-cultured at 80% confluence by dissociation from the flask using TrypLE (ThermoFisher Scientific, UK). A confluent T75 flask (Corning, USA) was split at a ratio of 1 in 3. The Hs 255T cell line was cultured at 37ºC and 5% CO2 in complete media (Dulbecco's Modified Eagle's Medium (DMEM) (ATCC, Manassas, VA, USA) supplemented with 10% FBS

(Sigma-Aldrich, Gillingham, UK), 100U/mL Penicillin (Gibco, Invitrogen, United Kingdom) and 100 µg/mL streptomycin (Gibco, Invitrogen, United Kingdom). Cells were sub-cultured at 80% confluence by dissociation from the flask using TrypLE Express (ThermoFisher Scientific, UK). A confluent T75 flask (Corning, USA) was split at a ratio of 1 in 3.

2.4.1.2 Cell banking for long term storage

Cell lines at 80-90 % confluence in a T75 flask were washed with PBS, and 3 mL of TrypLE (ThermoFisher Scientific, UK) was added to the flask and cells left in the incubator at 37°C for 3 minutes. Once the cells were detached from the flask, 10 mL of complete media was added into the flask to quench the enzymatic reaction. Cells were then transferred to a 15 mL falcon tube (ThermoFisher Scientific, UK) and centrifuged at 500 x g for 5 minutes at an acceleration and deceleration speed of 9. The supernatant was removed, cells were re-suspended in 1 mL of CryoStor® cell freezing media (ThermoFisher Scientific, UK) and transferred into two 1.5mL Cryogenic vials and stored in Mr. Frosty freezing container that provides 1oC/min cooling rate. Cells were stored at -800C for 24 hrs then transferred to liquid nitrogen for long term storage.

2.4.1.3 Thawing cells

To recover frozen cells, defrosting was performed in a water bath for less than 2 minutes and transferred to a 15 mL falcon tube containing 10 mL of CCD-18Co media made up of EMEM (ATCC, Manassas, VA, USA) + 10% FBS (Sigma-Aldrich, Gillingham, UK), 100U/mL Penicillin (Gibco, Invitrogen, United Kingdom) and 100 µg/mL streptomycin (Gibco, Invitrogen, United Kingdom). Cells were centrifuged at 500 x g for 5 minutes at an acceleration and deceleration speed of 9, followed by a second wash with media and repeat of the centrifugation step. Cells were then re-suspended in 10 mL of CCD-18Co media and transferred to a T75 flask and left in the incubator at 37°C and 5% CO2. Cells were left to adhere to the flask for approximately 24 hours, and then media was replaced every 2-3 days

until cells were confluent enough for further experiments. Cell lines were routinely tested to monitor mycoplasma contamination of cell cultures.

2.4.1.4 In vitro stimulation of CCD-18Co cell line

CCD-18Co were counted and diluted at 5x104 cells/mL with complete media, and 500µl was transferred into wells of a 24 well plate. The cells were left in culture at 370C and 5% CO2 in an incubator for 12 hours, after which media was removed replaced with serum free EMEM with 100U/mL Penicillin (Gibco, Invitrogen, United Kingdom) and 100 µg/mL streptomycin (Gibco, Invitrogen, United Kingdom) and 0.2% BSA (Sigma-Aldrich, Gillingham, UK). The cells were incubated as before for an additional 12 hours. Serum free media was removed and CCD-18Co media containing treatments of TGFβ1 (Miltenyi Biotec, Germany), IL-13 (Peprotech, UK) and IL-36 (Peprotech, UK) at 1 ng/mL, 10 ng/mL and 100 ng/mL respectively. The cells were left in culture at 370C, 5% CO2 for 24 and 48 hours. Wells containing unstimulated cells were included as a baseline sample (control). At the end of each time point, media was removed, and cells were washed with 1X PBS (Sigma-Aldrich, Gillingham, UK) before being lysed in 350 µL of lysis buffer (RLT buffer (Qiagen, UK) containing β-mercaptoethanol (Sigma-Aldrich, Gillingham, UK) at a dilution ratio of 10 µL β-mercaptoethanol for each 1 mL RLT lysis buffer). The lysed cells were transferred to 1.5 mL Eppendorf tubes ready for RNA extraction. RNA was either extracted immediately or lysed cells were stored at -800C ready for RNA extraction.

2.4.2 Gene expression assays

Gene expression analysis was performed on tissue samples from the proximal, stricture and distal segments of resected gut as well as cells collected from in vitro cell culture assays. Fresh tissue for gene expression studies was transported to the lab in tissue storage media (Miltenyi Biotec, Germany) and transferred immediately to RNAlater at 4^0C (Sigma-Aldrich,

Gillingham, UK), this was stored further in 40C for various times until it was ready for further experimental work.

2.4.2.1 RNA extraction from cells

All reagents were provided in the kit except for ethanol solutions prepared in the lab)

Total RNA was extracted from lysed cells using the RNEasy micro kit (Qiagen, UK) according to manufacturer's instructions. The RNEasy micro kit is designed to extract RNA of up to 45 µg or in samples with cells <500,000, therefore the kit was fit for purpose based on the seeding density of cells used in the in vitro assays. Lysed cells were homogenised and 70% of ethanol (Fisher Scientific, UK) was added to precipitate the RNA from the cell lysates which was then collected within the membrane of the RNeasy MinElute spin column through centrifugation. Wash steps were performed followed by elimination of genomic gDNA using Deoxyribonuclease I (DNase I). Further wash steps were performed to purify RNA, with RW1, RPE buffers and 80% ethanol (Fisher Scientific, UK) used for these wash steps. These solutions are eliminate carbohydrates, proteins, fatty acids and residual salts retained in the column. The membrane was dried by centrifugation at maximum speed (12000 x g for 5 minutes). RNA was then eluted in 14 µL RNase-free water by centrifugation at maximum speed for 1 minute. The concentration of eluted RNA in ng/µL was measured using a Nanodrop Spectrophotometer (ThermoFisher Scientific, UK). The absorbance ratios at 260 nm/280 nm and 230 nm/260 nm were also recorded to check the purity of RNA and contaminants that may have resulted from reagents used. For an RNA sample a 260/280 ratio of >1.8 is considered pure RNA and a 260/230 ratio of >1.5 is considered clean of organic solvents.

2.4.2.2 Complementary DNA (cDNA) synthesis

Complementary DNA (cDNA) was synthesised using iScript cDNA synthesis kit (BioRad, Hercules, CA, USA) according to manufacturer's instructions. Extracted RNA (up to 1 µg) was

used as the starting RNA template. This was transcribed into cDNA by a mixture of iScript reverse transcriptase, iScript reaction mix which contained oligo(dT) and random hexamer primers and nuclease free water to a total volume of 20 µL, as detailed in the Table 2.5;

Table 2.5: Reagents used for cDNA synthesis

Reagents	Final Assay Concentration	Company
5X *i*Script reaction mix	4 µL	Biorad, UK
*i*Script reverse transcriptase	1 µL	Biorad, UK
Nuclease free water		
RNA template (up to 1 µg total RNA)	10 µg/mL	From cell lines and tissue

2.4.2.3 Real time-PCR (qPCR)

RT-PCR was used to measure gene expression in cells obtained from in vitro cell culture experiments (list of genes shown in table 2.6). Taqman assays which contain primers and Taqman probes labelled with a 5' reporter dye (FAM) and a 3' quencher dye (probe assay IDs in table 2.6) were used. Serine and arginine Rich Splicing Factor 4 (SRSF4) was used as an endogenous control. The assay was run on the ABI Prism 7900HT sequence detection system (ThermoFisher Scientific, UK) at the University of Birmingham. Differential gene expression levels were calculated using the delta Ct method by (Livak and Schmittgen, 2001) as shown below;

$$\text{Delta Ct } (\Delta Ct) = \text{Av. Ct. of gene of interest} - \text{Av Ct. of endogenous gene})$$

$$\text{Delta delta Ct } (\Delta Ct) = \Delta Ct \text{ of measured condition (Stimulated)} - \Delta Ct \text{ control (unstimulated)}$$

$$\text{Relative gene expression} = (\text{Power } 2^{-\Delta\Delta Ct})$$

Table 2.6: Reagents used for qRT-PCR

Reagents	Final Assay Volume (Total 10 µl)	Company
Taqman 20X primers	0.25 µL	Life Technologies
PCR 2X master mix	5 µL	Life Technologies
RNAse free water	3.75 µL	Life Technologies
cDNA	1 µL	

Table 2.7: Protocol for qRT-PCR

Step No.	Process	Treatment temperature	Duration	Number of cycles
1	**Pre-incubation**	95^0C	10 minutes	1
2	**Denaturation**	95^0C	10 seconds	40
3	**Annealing**	60^0C	30 seconds	
4	**Extension**	72^0C	30 seconds	

Table 2.8: List of genes and primer sequences for targets used in qPCR

Gene	Name	Assay ID	Dye
SRSF4	Serine And Arginine Rich Splicing Factor 4	Hs00900675_m1	FAM
COL1A1	Collagen Type I Alpha 1 Chain	Hs00164004_m1	FAM
ACTA2	Actin Alpha 2, Smooth Muscle	Hs00426835_g1	FAM
FN1	Fibronectin 1	Hs01549976_m1	FAM
TIMP1	TIMP Metallopeptidase Inhibitor 1	Hs01092512_g1	FAM
GREM1	Gremlin 1	Hs01879841_s1	FAM
CTGF (CCN2)	Cellular Communication Network Factor 2	Hs00170014_m1	FAM
BMP2	Bone Morphogenetic Protein 2	Hs00154192_m1	FAM
BMP4	Bone Morphogenetic Protein 4	Hs00370078_m1	FAM
BMP7	Bone Morphogenetic Protein 7	Hs00233476_m1	FAM
TGFB1	Transforming Growth Factor Beta 1	Hs00998133_m1	FAM
BMPRIA	Bone Morphogenetic Protein Receptor 1A	Hs04980288_g1	FAM
BMPRIB	Bone Morphogenetic Protein Receptor 1B	Hs01010965_m1	FAM
BMPRII	Bone Morphogenetic Protein Receptor II	Hs00176148_m1	FAM

2.5 Bulk RNA sequencing

2.5.1 RNA extraction

Total RNA was extracted from 9 matched gut resected tissue samples (proximal, stricture and distal segments) that had been stored in RNAlater (Sigma-Aldrich, Gillingham, UK) using the RNEasy mini kit (Qiagen, UK) according to manufacturer's instructions. The RNEasy mini kit is designed to extract RNA of up to 100 µg from cells or tissue. Tissue was transferred from RNAlater solution to a gentleMACS™ C tube containing RLT lysis buffer containing β-mercaptoethanol, and then the tissue was disrupted using a pre-set protocol "RNA_01 for tissue" on the gentleMACS™ dissociator. Further tissue disruption was performed using a 10 mL syringe, mixing tissue "up and down" to obtain a homogenous solution. Disrupted tissue was then centrifuged for 5 mins at 12000 x g for 5 minutes, after which the supernatant was transferred to a fresh tube while the pellet was discarded, and 70% of ethanol (Fisher Scientific, UK) was added to the supernatant to precipitate the RNA.

The homogenised sample including any precipitate was transferred to an RNeasy Mini spin column, and centrifugation was performed at ≥8000 x g. A wash step was then performed using the RW1 buffer followed by elimination of genomic gDNA using Deoxyribonuclease I (DNase I). Further wash steps were performed to purify RNA, with RW1 and RPE buffers (Qiagen, UK) were used for the wash steps. These solutions eliminate carbohydrates, proteins, fatty acids and residual salts stuck in the column. The membrane was dried by centrifugation at maximum speed (12000 x g for 5 minutes). RNA was then eluted in 30 µl RNase-free water by centrifugation at maximum speed for 1 minute ready for the next experimental steps.

2.5.2 RNA quality assessment

Extracted RNA was subjected to quality checks: RNA concentration and purity (A260/A280) was measured using the Nanodrop Spectrophotometer (ThermoFisher Scientific, UK) while

RNA quality and integrity was checked using the 2100 Bioanalyzer (Agilent) by measurement of the RNA integrity number (RIN). Samples with purity of >1.8 at OD260/280 and RNA integrity number (RIN) of >7 were used to prepare cDNA libraries.

2.5.3 cDNA library preparation for sequencing

Preparation of cDNA libraries was performed using the QIAseq UPX 3' (Qiagen, UK) according to manufacturer's instructions. In summary, 0.5ng of RNA was reverse transcribed to cDNA; during this step each sample was tagged with a unique molecular index (UMI). All samples with tagged UMIs were then pooled into a single tube and subjected to further amplification through PCR using the Qiagen HiFi PCR master mix (Qiagen, UK). The amplified cDNA sequence was fragmented into uniform short reads to make sequencing possible; the fragmented ends were repaired and "A-tailed" by addition of an adenine base. The library was completed by ligation of adaptors to the cDNA fragments, that serve to identify the library and attach it to the flow cell during the sequencing process. Finally, the cDNA fragments each tagged with their UMI and library adaptors were further amplified by PCR to enrich the amplicon and increase sensitivity of the assay. The amplified fragments in the cDNA library were quantified using the QIAseq Library Quant Assay Kit (Qiagen, UK) and sequenced using the Nextseq Illumina platform. Nextseq Illumina sequencing was performed by the genomics department within the Institute of Biomedical Research, University of Birmingham. Raw data was generated as FastQ files ready for analysis.

2.5.4 Analysis of bulk RNA-seq data

Bulk RNA-seq data analysis was performed by Dr Animesh Acharjee, a senior research fellow with specialty in 'omics' data analysis from the Institute of Cancer and Genomic Sciences, University of Birmingham. The FastQ files obtained as raw data were de-multiplexed, aligned to the genome, quantified and normalised using the CLC Genomics workbench (Qiagen, UK).

The normalised values of reads reported in transcripts per million (TPM) were quality controlled by excluding genes with transcript expression of mean normalised counts that were less than a set threshold (< 1) and a false discovery rate corrected p value of <0.05 was considered and filtered for further downstream data analysis. Partial least squares discriminant analysis (PLS-DA) modelling was performed on these genes and a Variable Importance in Projection (VIP) score was used as a filter of genes of importance as per the model, genes with VIP Score of ≥1 were considered to be important in the PLS-DA model. DEGs were identified using the Linear models for Microarray data analysis (limma) package. Gene expression was compared between proximal vs stricture (P vs S), distal vs stricture (D vs S) and proximal vs distal (P vs D). Pathway enrichment was done using the databases Gene ontology biological pathways, Wiki pathways and Kyoto Encyclopaedia of Genes and Genomes (KEGG).

2.6 Single cell RNA sequencing

2.6.1 Isolation of single cells from resected gut samples for RNA sequencing

Single cell RNA sequencing was performed on cells collected from the stricture (n=3) and proximal (n=3) segments. The decision to sequence tissue from the stricture and proximal was taken because both segments are located in the small bowel, and strictures in CD commonly occur in the terminal ileum located within the small bowel. These made the two segments more comparable in terms of gene expression patterns, except for the changes that may arise due to disease phenotype in each segment. Tissue was collected from pathology in MACS tissue storage solution (Miltenyi Biotec, Germany). Tissue was weighed and transferred into a petri dish (ThermoFisher Scientific, UK), sliced into 3-5 mm sections using a scalpel, washed in cold PBS and transferred into gentleMACS™ C tubes (Miltenyi Biotec, Germany) which contained 2.5 mL of digestion media. The digestion media contained DMEM (Sigma-Aldrich, Gillingham, UK), 100 µg/mL Liberase enzyme (Roche, Basel, Switzerland), 100U/mL Penicillin (Gibco, Invitrogen, United Kingdom), 100 µg/mL streptomycin (Gibco, Invitrogen,

United Kingdom), 50 µg/mL gentamicin (ThermoFisher Scientific, UK) and 0.25 µg/mL amphotericin B (Sigma-Aldrich, Gillingham, UK). Tissue digestion was performed on a gentleMACS™ Octo dissociator with heaters (Miltenyi Biotec, Germany). A protocol was set up on the Octo dissociator to mechanically digest tissue while maintaining a temperature of 370C for 20-30 minutes depending on tissue stiffness. The tissue digest was then filtered in a 70µm cell strainer (Miltenyi Biotec, Germany). Culture media containing DMEM (Sigma-Aldrich, Gillingham, UK), 16% heat-inactivated FBS (Sigma-Aldrich, Gillingham, UK), 100U/mL Penicillin (Gibco, Invitrogen, United Kingdom), 100 µg/mL streptomycin (Gibco, Invitrogen, United Kingdom), 50 µg/mL of gentamicin (ThermoFisher Scientific, UK) and 0.25 µg/mL amphotericin B (Sigma-Aldrich, Gillingham, UK) was added to the flow through to stop the enzymatic reaction. The flow through containing cells was subjected to centrifugation at 500 x g for 5 minutes, the supernatant was discarded, and the pellet was resuspended in sterile PBS and another wash step was performed. Cells were then counted ready for the next experimental steps.

2.6.2 Enrichment of viable cells by Fluorescence Activated Cell Sorting (FACS)

To enrich viable cells for sequencing, freshly isolated single cells from the proximal and stricture segments of resected gut were stained with propidium iodide (PI) dye ready for sorting on the FACSAria machine (BD Biosciences). Viable cells were collected, and a viability test was performed using trypan blue on the Biorad TC20 automated cell counter (Biorad, UK). Samples containing >80% viable cells were used for single cell sequencing. A total of three matched samples from the stricture (n=3) and proximal (n=3) were used for sequencing.

2.6.3 10X Genomics Chromium single cell sequencing

The 10X Genomics Chromium system is a platform used for single cell sequencing, and the approach is based on rapid droplet-based encapsulation of single cells using a gel bead in

emulsion (GEM). Each gel bead is labelled with oligonucleotides that consist of a unique barcode, a 10 bp UMI, sequencing adapters/primers, and an anchored 30 bp oligo-dT (See et al., 2019). Complementary DNA synthesis and library preparation was performed using Chromium Single Cell 3' Library & Gel Bead Kits (10X genomics, USA) by the genomics facility team at the University of Birmingham, UK. The cDNA was quantified, and quality controlled by tape station analysis before the libraries were prepared and sequenced using Illumina NextSeq500/550High-Outputv2.5Kit (150 cycles) sequencing reagents to obtain 5000 cells per sample.

2.6.4 Analysis of single cell sequencing data

Single cell RNA sequencing data analysis was performed by Dr Csilla Varnai, a Research Fellow in Bioinformatics from the Institute of Cancer and Genomic Sciences.

The FastQ files obtained as raw data were de-multiplexed, and quality control of the data was performed. Cell filtration criteria was based on quality control metrics in Seurat data analysis software. The exclusion criteria included: cells in which <200 genes were detected as low quality cells have few genes; cells in which >2,500 genes were detected as cell doublets or multiplets may exhibit high gene count; and cells with a mitochondrial content >5% as low quality cells or dying cells tend to exhibit a high contamination of mitochondria. Filtered cells were then merged and subjected to cluster analysis for cell type identification using the R package Seurat. Based on gene expression patterns, cells were grouped under the following major cell categories: epithelial, stromal/mesenchymal, myeloid, and lymphoid cells. For similarity-based cell network analysis and visualization, dimensionality reduction was performed based on gene expression profiles to identify cells that clustered together. Uniform Manifold Approximation and Projection (UMAP) figures were generated for the different clusters. Reference markers (gene expression) were then used to identify cell type in each cluster. For cell types with low expression profiles, sub-clustering was performed to identify

these cell types. Further downstream analysis was performed to identify expression profiles of the genes of interest comparing expression in the proximal vs the stricture. (Adams et al., 2020).

2.7 CCD-18Co transfection with small interfering RNA (siRNA)

To study the downstream signalling effect of the identified gene on interest (GREM1), small interfering RNA (siRNA) was used to silence GREM1 expression in CCD-18Co. The protocol involved the use of the FlexiTube kit (Qiagen, USA) that contained 2 types of reagent: FlexiTube gene solution small interfering RNA (siRNA) and HiPerFect transfection reagent. Four siRNAs (Qiagen, USA) designed to target different regions of human gremlin 1 (Refseq ID: NM_001191322) were used. CCD-18Co cells were plated in a 24 well plate at 5x104 cells/mL in culture media made containing EMEM (ATCC, Manassas, VA, USA) supplemented with 10% FBS (Sigma-Aldrich, Gillingham, UK), 100U/mL Penicillin (Gibco, ThermoFisher, United Kingdom) and 100 µg/mL streptomycin (Gibco, ThermoFisher, United Kingdom). The cells were left in culture at 370C and 5% CO2 for 24 hrs to adhere to the plate. Cells were then transfected with GREM1 siRNA (Qiagen, UK) and negative control siRNA (Qiagen, UK) at 60nM in a final volume of 600 µl as described below.

The respective siRNAs were made up in 100 µl of serum free media and mixed by vortexing for 10 seconds. In addition to the siRNA mix, 4.5 µl of HiPerFect transfection reagent was added and mixed by vortexing for 10 seconds. The mixture was incubated for 10 minutes at room temperature to allow formation of transfection complexes. Culture media was removed from cells and replaced with 500µl of culture media. 100µl of psiRNA complexes containing HiPerFect transfection reagent and siRNA (GREM1 and negative control, prepared as described above) were added onto the respective wells while gently swirling the plate to ensure uniform distribution of the transfection complex mixture across the well; the final concentration of siRNA in each well was 60nM. Cells were then incubated at 370C and 5%

CO2 for 72 hrs. To measure the efficiency of siRNA transfection in CCD-18Co, cells were lysed after 72 hrs of transfection and GREM1 gene expression was measured in the cell lysates using qPCR as described in sections 2.5.1, 2.5.2 and 2.5.3. Transfected cells were also used for further downstream in vitro experiments to further understand effects of GREM1 knockdown in CCD-18Co. The nucleotide sequences used for GREM1 siRNA optimisation are shown in table 2.9.

Table 2.9: Sequences of the four synthetic siRNAs targeting *GREM1*

siRNA no.	Target position	Target sequence (5'-3')	RNA oligo sequences (5'-3')
siRNA1	477-499	TCGTTGCATATCCATCGATTTGG	AAAUCGAUGGAUAUGCAACGAGUUGCAUAUCCAUCGAUUUGG
siRNA2	559-581	GACCTAAAACAACCAGATTCTTA	AGAAUCUGGUUGUUUUAGGUCCCUAAAACAACCAGAUUCUUA
siRNA3	989-1011	TCCATCTCTTCTTAAGTTGATAG	AUCAACUUAAGAAGAGAUGGACAUCUCUUCUUAAGUUGAUAG
siRNA4	1218-1240	AAGCTTGAAAGGCCAATACCAGA	UGGUAUUGGCCUUUCAAGCUUGCUUGAAAGGCCAAUACCAGA

2.8 CCD-18Co lentiviral transduction

Lentiviruses have become a stable and fast means of inserting transgenes into animals or cell lines to address outstanding scientific questions. To study the biological function of abundant GREM1 expression and its effect on downstream signalling, CCD-18Co were transduced with lentiviral particles coding for GREM1 and/or GFP. This was performed to integrate GREM1 into the cell genome resulting in gene overexpression, and transduction efficiency was determined by examining cells for GFP expression from a separate locus using a fluorescence microscope. The lentiviral particles used for GREM1 transduction were generated by members of the Weston lab (Centre for liver and gastrointestinal research, Birmingham, UK). CCD-18Co were seeded in 6-well plates in culture media (10% FBS+ EMEM+ 100U/mL penicillin, 100 µg/mL streptomycin) at a concentration of 2.5x105 cells per well. CCD-18Co were either co-transduced with both GFP- and GREM1-encoding particles or GFP-encoding particles alone for the control cells, , by adding 5 µl of lenti-GREM1 and/or lenti-GFP particles to the respective wells. For this experiment, multiplicity of infection (MOI) was not pre-determined. Cells were cultured at 370C, with 5% CO2 for 48 hrs. At this point, cells emitting the GFP green fluorochrome could be visualised under a fluorescence microscope; this was an indicator of a successful transduction process. Culture media containing any remaining viral particles was gently removed and the cells were washed with PBS (X2). 2 mL of fresh culture media was added onto the cells and left in culture. The cells were passaged to expand for further experiments.

2.9 Semi- quantitative PCR

To check for transduction efficiency, GREM1 transduced cells were lysed, RNA was extracted (section 2.5.1) and cDNA obtained through reverse transcription of RNA using iScript kit (BioRad, Hercules, CA, USA) (section 2.5.2). For the amplification of cDNA, a master mix was prepared as shown in table 2.10, the forward primer sequence was

GCCACCATGAGCCGCACAGC while the reverse primer sequence was TCACTATTAATCCAAATCGATGGATATGCAAC. The PCR reaction was amplified for 20 and 30 cycles respectively. The PCR product was diluted in 1X DNA gel loading dye (ThermoFisher Scientific, UK) and loaded onto a 2% (w/v) agarose gel (Sigma-Aldrich, Gillingham, UK) containing 0.5% ethidium bromide (Sigma-Aldrich, Gillingham, UK) in Tris-borate-EDTA (TBE) buffer (ThermoFisher Scientific, UK). The ethidium bromide binds to the DNA and can be visualised under ultraviolet light. A 100 base pair DNA Ladder (ThermoFisher Scientific, UK) was loaded in the first lane of the gel to enable the size of the target gene to be determined. The gel was run for 60 minutes at 100 V after which GREM1 DNA fragments with an expected size of 567bp were visualised by imaging under the Chemidoc imaging system (Biorad, UK).

Table 2.10: Reagents used for qRT-PCR

Reagents	Final Assay Volume (Total 10 µl)	Company
2X MM Q5 Hot Start Polymerase	0.25 µl	New England Biolabs Ltd (NEB)
10 µM reverse primer	1.25 µl	New England Biolabs Ltd (NEB)
10 µM forward primer	1.25 µl	New England Biolabs Ltd (NEB)
RNAse/DNAse free water	6.25	Qiagen, UK
cDNA (1ng/mL)	1 µl	

Table 2.11: Protocol for semi-quantitative PCR

	Treatment temperature	Duration	Number of cycles
Pre-incubation	98°C	30 s	1
Denaturation	98°C	10 s	20 or 30
Annealing	67°C	30 s	
Extension	72°C	2 min	
Cooling	4°C	∞	1

2.10 CCD-18Co proliferation using the MTS assay

CCD-18Co proliferative activity was measured with the MTS reduction assay (#ab197010; Abcam, UK). The assay is based on MTS tetrazolium reduction by NAD(P)H-dependent dehydrogenase enzymes in viable cells, which generates a coloured formazan dye. The greater the production of the formazan dye, the higher the number of proliferating cells. CCD-18Co

were plated into a 96-well plate at densities of 6,250 and 25,000 cells per well and either left unstimulated or stimulated with TGF-β (100 ng/mL; Miltenyi Biotec, Germany), or IL-1β (100 ng/mL; Miltenyi Biotec, Germany) or PDGF-BB (100 ng/mL; Peprotech, UK). Cells were incubated at 370C and 5% CO2 and left in culture for 20 hours, 36 hours and 48 hours. At each time point, 20µl of MTS reagent were added and cells were incubated at 370C, with 5% CO2 for 4 hours. Quantification of the formazan dye reaction product was performed by measuring the absorbance of the culture medium at 490nm using the ELISA plate reader (BioTek Synergy HT).

2.11 Protein expression assays

2.11.1 ELISA

An ELISA kit to measure human Gremlin 1 (#EKH4238; Lot No: H2119F033) was purchased from Nordic BioSite. CCD-18Co cells were stimulated with pro-fibrotic factors such TGF-β (100 ng/mL; Miltenyi Biotec, Germany), IL-1β (100 ng/mL; Miltenyi Biotec, Germany) PDGF-BB (100 ng/mL; Peprotech, UK), PAI-1 (100 ng/mL; Peprotech, UK) or FGF2 (100 ng/mL; Peprotech, UK); supernatant was collected. Gremlin 1 proteins levels were determined using ELISA performed according to manufacturer's instructions.

2.11.2 Haematoxylin and Eosin staining (H & E)

The H and E stain involves haematoxylin which is a basic dye that stains acidic structures including the nucleus purple/blue, and eosin which is an acidic dye that stains basic structures such as the cytoplasm pink. Paraffin sections of 5 µm thickness were dewaxed by incubating in 3 sequential basins containing xylene (Pioneer research chemicals, UK) for 2 minutes in each, to enable the aqueous labelling solutions to penetrate the fixed tissue. The tissue sections were then passed through 3 basins of ethanol (Sigma-Aldrich, Gillingham, UK) in descending order from absolute ethanol, 96% ethanol and then 70% ethanol for 2 minutes in each solution.

Tissue sections were then placed in 2 basins of distilled water (2 minutes in each) to rehydrate. The sections were transferred to a rack containing Heamatoxylin Harris (Sigma-Aldrich, Gillingham, UK) solution for 4 minutes and the excess stain washed off by placing the sections in water for 2 minutes. The sections were transferred to acid alcohol (ThermoFisher Scientific, UK) for 30 seconds to remove excess staining, and then washed again with water for 2 minutes. Next, the sections were transferred to Scott's tap water substitute (Cell path, UK) for 30 seconds. This solution helped in bluing the sections to improve quality of the H and E stain. Sections were then dipped in distilled water (dH2O) for 2 minutes before transferring to the eosin (Sigma-Aldrich, Gillingham, UK) stain for 1 minute. Sections were dehydrated by passing sequentially through 2 basins of distilled water, 3 basins of alcohol in ascending order at concentrations of 70%, 96% and absolute ethanol for 2min in each solution. The sections were then passed through 3 basins of xylene (2 minutes in each). Finally, the sections were mounted in distyrene, plasticizer and xylene (DPX) (Cell path, UK) with a cover slip, left to dry and imaged using the Zeiss AxioScan Z.1 slide scanner (Zeiss, Munich, Germany)

2.11.3 Immunohistochemistry

Immunohistochemistry has long been used for identification of protein expression based on the principle of antibodies binding to the antigen of interest. The target antigens in this case were FGF2, Gremlin 1 and HDAC1. Formalin fixed paraffin embedded (FFPE) tissue blocks collected from the proximal, stricture and distal segments of resected gut were used with sections cut into 5 µm. Sections were dewaxed and rehydrated by placing in xylene (x3) (Pioneer research chemicals, UK) for 3 minutes in each basin, then transferred to descending concentrations of alcohol (Sigma-Aldrich, Gillingham, UK) with concentrations including absolute ethanol, 96% of ethanol and 70% ethanol for 3 minutes in each solution, then in 2 basins of distilled water. The next step involved unmasking epitopes of the target protein to improve detection and antibody binding. This was performed using an antigen retrieval solution

(pH 9) which was diluted in distilled water (10 mL of pH high solution (Vector Labs, UK) in 1L of distilled water). The solution was heated in the microwave at high power for 10 minutes. The tissue sections were transferred into the heated antigen retrieval solution and placed in the microwave to heat on high power for further 30 minutes. The solution was allowed to stand for 10 minutes and then brought to room temperature by running cold water. Tissue sections on each slide were circled using a wax pen (ImmEdge Pen, Vector Laboratories, UK) to prevent leakage of any staining compound being added onto the sections. The sections were then placed in Tris-Buffered Saline containing 0.1% Tween (TBS-T) for 5 minutes and then blocked with 2 drops of BLOXALL reagent (Vector Labs, UK) per tissue section for 10 minutes; BLOXALL reagent is a dual enzyme blocking reagent that inactivates endogenous peroxidase and alkaline phosphatase. The tissue was then washed in TBS-T for 5 minutes and blocked in Casein reagent (Vector labs, UK) for a further 10 minutes to reduce unspecific binding. A volume (100µl) of the respective primary antibody solutions prepared as listed in table 2.11 was added onto the sections and left at room temperature with gentle agitation (platform shaker) for one hour.

The tissue sections were then washed with TBS-T twice for 5 minutes. For antigen detection, 100µl of secondary antibody raised against the isotype of the primary antibody and linked to the horseradish peroxidase (ImmPRESS HRP Anti-IgG) (Vector Labs, UK) was added into each section at room temperature with gentle agitation (platform shaker) for 30 minutes. Tissue sections were then washed twice in TBS-T for 5 minutes. Meanwhile, 3,3' Diaminobenzidine (DAB) solution was prepared by adding 1 drop of chromogen into 1 mL of diluent as per manufacturer instructions (Vector Labs, UK). 5-6 drops of DAB solution were added on the tissue sections. The reaction was stopped by washing with tap water after 5 minutes. The tissue sections were then counter stained with 3 drops per section of haemotoxylin (Pioneer research chemicals, UK) for 45 seconds and the reaction stopped by washing with tap water. Hot tap water was then run on the tissue sections for 2 minutes gently to avoid washing the tissue off,

excess water was drained off, and sections were then transferred to 3 basins of alcohol in ascending order at concentrations of 70%, 96% and absolute ethanol for 3x3minutes, then into xylene for 3x3minutes.The sections were drained of excess xylene and immediately mounted in distyrene, plasticizer and xylene (DPX) (Cell path, UK), covered with a cover slip, left to dry and then imaged using the Zeiss AxioScan Z.1 slide scanner (Zeiss, Munich, Germany).

2.11.4 Immunofluorescence

Immunofluorescence (IF) staining technique allows for detection of multiple protein expression on the same section if the antibodies are raised in different species or are of differing isotypes. In addition, co-expression of proteins and how they interact within the same section can be visualized. This was especially important because certain cells of interest in this study such as fibroblasts are detected by a combination of different markers. IF was performed on 5µm FFPE sections from the proximal, stricture and distal segments of resected gut tissue. Sections were labelled, dewaxed and rehydrated by placing into xylene (x3) (Pioneer research chemicals, UK) for 3 minutes in each basin, then transferred to descending concentrations of alcohol (Sigma-Aldrich, Gillingham, UK) with concentrations including absolute ethanol, 96% of ethanol and 70% for 3 minutes in each solution, then in 2 basins of distilled water.

Heat induced antigen retrieval at a basic pH was performed (pH 9) for 30 minutes. Tissue sections on each slide were both circled using a wax pen (ImmEdge Pen, Vector Laboratories, UK) to prevent leakage of any staining compound being added onto the sections. The sections were then placed in TBS-T)for 5 minutes, and unspecific antigen binding was minimised by incubation with Casein blocking solution (Vector Labs, UK) for 30 minutes. Primary antibodies were prepared in 10% goat serum (Sigma-Aldrich, Gillingham, UK) and PBS; 100µl of the primary antibody solution and the corresponding isotype controls as listed in table 3.3 were added onto the sections and left at room temperature with gentle agitation (platform

shaker) for one hour. The tissue sections were then washed twice with TBS-T for 5 minutes. For antigen detection, slides were stained with secondary antibodies conjugated with fluorochrome raised against the species of the primary antibody. Secondary antibodies were prepared in PBS containing Hoechst dye (Sigma-Aldrich, Gillingham, UK) at a dilution of 1:1000 µl in PBS. Slides were stained for 30 minutes and then washed twice with TBS-T for 5 minutes. Slides were then mounted on to glass microscope slides using prolong gold anti fade mounting reagent. Imaging was performed using the confocal microscope (Zeiss, Munich, Germany).

Table 2.12: List of antibodies used for IHC, IF and mIF
(Abbreviations: IgG: Immunoglobulin G, CD: Cluster of differentiation, k:kappa, Ms:Mouse, AF: AlexaFluor)

Unconjugated primary antibodies					
Target	**Clone**	**Stock concentration**	**Dilution**	**Isotype**	**Company**
Mouse anti-human desmin	D33	111 mg/L	1:100	Ms IgG1, k	DAKO
Mouse anti-human vimentin	V9	156 mg/L	1:100	Ms IgG1, k	DAKO
Mouse anti-human α-SMA	1A4	71 mg/L	1:200	Ms IgG2a, k	DAKO
Rat anti-human podoplanin	NZ-1.3	0.5 mg/mL	1:50	Rat / IgG2a	Invitrogen
Rabbit anti-human CD90	EPR3133	0.08 mg/mL	1:100	Rabbit IgG	Abcam
Rabbit anti-human PDGFRβ	42G12	0.162 mg/mL	1:100	Rabbit IgG	Abcam
Mouse anti-human HDAC1	10E2	200 µg/mL	1:100	Ms IgG1, k	SantaCruz
Mouse anti-human FGF2	C-2	200 µg/mL	1:100	Ms IgG2a, k	SantaCruz
Goat anti-human Gremlin 1	Polyclonal	0.2 mg/mL	1:50	Goat IgG	R and D
Secondary antibodies					
Immunogen	**Conjugate**	**Stock concentration**	**Dilution**		**Company**
Rat IgG (H+L)	AF488	2 mg/mL	1:1000		Invitrogen
Rabbit IgG (H+L)	AF488	2 mg/mL	1:1000		Invitrogen
Rabbit IgG (H+L)	AF555	2 mg/mL	1:1000		Invitrogen
Mouse IgG1	AF647	2 mg/mL	1:1000		Invitrogen
Mouse IgG2a	AF555	2 mg/mL	1:1000		Invitrogen
Rat IgG (H+L)	AF647	2 mg/mL	1:1000		Invitrogen
Mouse IgG2a	AF488	2 mg/mL	1:1000		Invitrogen
Rat IgG (H+L)	AF488	2 mg/mL	1:1000		Invitrogen
Mouse IgG2a	AF546	2 mg/mL	1:1000		Invitrogen
Mouse IgG1	AF488	2 mg/mL	1:1000		Invitrogen
Mouse IgG1	AF594	2 mg/mL	1:1000		Invitrogen
Mouse IgG1	AF546	2 mg/mL	1:1000		Invitrogen
Rabbit IgG (H+L)	AF647	2 mg/mL	1:1000		Invitrogen

2.12 Statistical analysis

2.12.1 Graphpad prism

Data analysis was performed using Graphpad PRISM 9 (Graphpad Software Inc, USA). Data sets were tested for normality - a parametric test was used to analyse normally distributed data while non-parametric test was used for data that was not normally distributed. To compare two data sets, a student's t test was used. One way ANOVA or two-way ANOVA was used in cases where comparison was being made in means of more than two groups. A value of $P \leq 0.05$ was considered statistically significant. Statistical analysis was performed for both the bulk RNA-seq and single cell seq data, P values were generated to check for statistical significance in gene expression profiles between proximal, stricture and distal segments

CHAPTER 3

TRANSCRIPTOMIC AND PROTEOMIC PROFILING OF STRICTURE, PROXIMAL AND DISTAL SEGMENTS TO IDENTIFY MEDIATORS AND PATHWAYS OF INTESTINAL FIBROSIS IN CROHNS DISEASE

Overview of this chapter

In this chapter I studied: a) the transcriptomic profile of the proximal, stricture and distal segments from resected gut tissues from patients with Crohn's disease using bulk RNA-sequencing and single cell RNA-sequencing and b) determined the frequencies and localization of immune and stromal cells in the three segments. The data from this chapter will give an insight into the molecular and cellular mediators as well as the key pathways involved in initiation and progression of inflammation and fibrosis in CD.

3.1 Introduction

The most common complication of CD is development of strictures with approximately 40-70% of CD patients requiring surgical intervention within 10 years of diagnosis (199,429). Deep data analysis from GWAS in CD patients have identified single nucleotide polymorphisms (SNPs) within genes such as *NOD2*, *ATG16L1*, *IRGM* and *IL-23R*. Patients diagnosed with CD and predisposed to these polymorphisms have an increased risk of developing severe disease and non-responsiveness to anti-inflammatory treatment specifically therapies that target IL-12/IL-23. Single nucleotide polymorphisms were also identified on *TGFB1*, *MMP3*, *JAK2* and *IL12B*; CD patients with these mutations have an increased risk of developing ileal fibrostenosing CD and undergoing multiple resection surgeries (285,437,438).These findings suggest a strong contribution of genetic factors in immune dysregulation and development of fibrostenosing CD. In agreement with this, multiple studies suggest that if the inflammatory response is not resolved it becomes a precursor to development of strictures in CD (72,285,439,440).

Genetic factors account for only 13.6% in disease variation; other factors such as epigenetic modifications are thought to influence disease outcome (441,442). Epigenetic changes such as DNA acetylation, deacetylation or methylation regulate gene transcription and can influence disease behaviour (443,444). For example, post-translational modification of histones by histone acetylases and deacetylases in mouse models of colitis have been shown to regulate transcriptional activity (445), hence affirming a key role of epigenetics in the pathogenesis of CD (441,446–448). Cells of the innate and adaptive immune system have also been reported to infiltrate inflamed mucosa in CD, these include T and B cells, neutrophils, monocytes, macrophages and NK cells among others. Cellular activation and differentiation results in increased levels of pro-inflammatory mediators such as IFN-γ, TNF-α, IL-6, IL-1β, TGF-β, IL-22 and IL-17. These cytokines promote proliferation and activation of MSCs including

fibroblasts, smooth muscle cells, endothelial and epithelial cells to produce extracellular matrix (ECM) proteins (297,310,336). Cytokines such as IL-33, IL-36 and IL-6 also act on MSCs to prevent production of MMPs and increase secretion of TIMPs. The imbalance between MMPs and TIMPs drives an increase in the deposition of ECM proteins (328,449). Although inflammation is required to initiate fibrosis, the fibrotic process can be maintained even after inflammation has been resolved, and this is mainly attributed to profibrotic pathways driven by activation of MSCs (450).

Progress has been made in the development of effective anti-inflammatory therapies; however, there is still a huge unmet need in the lack of anti-fibrotic therapies for the management of intestinal fibrosis (451). Understanding molecular and cellular mediators of key pathways that drive the fibrotic process and identifying those unique to intestinal fibrosis is key in the development of therapies for the management of fibrostenosing CD (452,453). Studies described in this chapter aimed to identify novel biomarkers and associated pathways that drive intestinal mucosal inflammation and fibrosis in the stricture vs non-stricture segments (proximal and distal) in resected gut of CD patients. This was performed with the aim of generating a data set that informs the knowledge gap on mechanisms that drive intestinal fibrosis in CD.

3.2 Aims of this chapter:

The specific aims of this chapter were to:

1. To study the transcriptomic profile of stricture, proximal and distal resection segments using bulk (bulk-seq) and single-cell RNA sequencing (scRNA-seq).
2. To study the proteome profile in expression and localisation of key markers by immunofluorescence.

3.3 Results

3.3.1 Patient demographic information and characteristics

Dr Uday Shivaji, a clinical research fellow and a member of our research group identified CD patients that were scheduled to undergo resection surgery at the QE hospital, Birmingham. A total of 25 patients were recruited into the tissue remodelling and fibrostenosing in Crohn's disease (TRAFIC) study (approval reference number RRK6508) also referred to as the fibrostenosing Crohn's disease (FCD) cohort. Of the 25 patients, 20 were Caucasian, 2 Asian and 3 Afro-Caribbean in ethnicity. Consent was obtained from both male and female patients with a median age of 37 years (range 19-65 years) and median disease duration of 14 years (range 0.3 – 38 years). Resected samples with consent from the patients were collected as per approved ethical guidelines. Details of patient demographics, disease classification and medications are summarised in table 3.1;

Table 3.1: Patient demographics and characteristics
(Medication: 5-ASA (1), Thiopurines (2), Methotrexate (3), Anti-TNF (4), Vedolizumab (5), Ustekinumab (6), Other immune-suppressive (7) Steroids (8))

Study code/ID	Age (years)	Sex	Montreal classification	Duration of disease (years)	Medications within 1 year of surgery	Medication at the time of surgery
FCD001	57	Male	B3 penetrating p perianal disease modifier	5 to 9 years	2;4	1; 2; 4
FCD002	65	Male	B1 non-stricturing, non-penetrating	5 to 9 years	No data	No data
FCD003	54	Male	B2 stricturing	>20 years	1; 2; 7	1; 2
FCD004	63	Female	B2 stricturing	>20 years	4	6
FCD005	24	Male	B2 stricturing p perianal disease modifier	10 to 20 years	7	8
FCD006	28	Male	B1 non-stricturing, non-penetrating	5 to 9 years	7	No data
FCD007	43	Female	B1 non-stricturing, non-penetrating p perianal disease modifier	5 to 9 years	No data	No data
FCD008	58	Female	B2 stricturing	>20 years	2; 4	2; 4
FCD009	19	Male	B3 penetrating	< 6 months		
FCD010	35	Female	B3 penetrating	5 to 9 years	2; 4	8
FCD011	55	Male	B3 penetrating p perianal disease modifier	>20 years	No data	No data
FCD012	55	Female	B2 stricturing	>20 years	4	4
FCD013	52	Male	B3 penetrating	>20 years	No data	No data
FCD014	33	Male	No data	No data	No data	No data
FCD015	28	Female	B2 stricturing	10 to 20 years	6	6; 8
FCD016	37	Male	B3 penetrating	6 to 12 months	4; 7	4; 8
FCD017	39	Male	B2 stricturing	10 to 20 years	2; 4; 6	2; 6
FCD018	22	Male	B2 stricturing	5 to 9 years	5	5
FCD019	26	Male	B3 penetrating	10 to 20 years	2;4	2; 4
FCD020	24	Male	B3 penetrating p perianal disease modifier	10 to 20 years	2;4	4
FCD021	55	Female	B3 penetrating	10 to 20 years	4	4
FCD022	42	Female	B2 stricturing	5 to 9 years	1; 2	4
FCD023	25	Female	B3 penetrating	No data	3; 6	8
FCD024	48	Female	B2 stricturing	5 to 9 years	No data	4
FCD025	56	Female	B2 stricturing	10 to 20 years	6	6

3.3.2 Sequencing as a tool to study transcriptome profiles "the old versus the new"

Transcriptional changes that occur in normal tissue versus diseased conditions have provided insight into variation in gene expression of different disease conditions. Such data has informed decisions on potential therapeutic targets that have been translated to life changing therapies (412). Bulk RNA-sequencing is used as a tool to study average gene expression profiles in each given sample, the method is robust, sensitive and provides information on variation in disease versus normal conditions. ScRNA-seq is a more recent technique that was introduced by Tang (2009); It is a more advanced and informative technique that permits the identification of individual cell specific gene expression profiles. The technique has been used to identify novel cell sub types, monitoring cell fate and plasticity in different states of disease (455). Our study incorporated both bulk- and scRNA-seq methods for an in depth understanding of gene expression profiles in the tissue as a whole and at a cellular level in stricture and non-stricture segments. Figure 3.1 provides an overview of the processes involved in the two techniques.

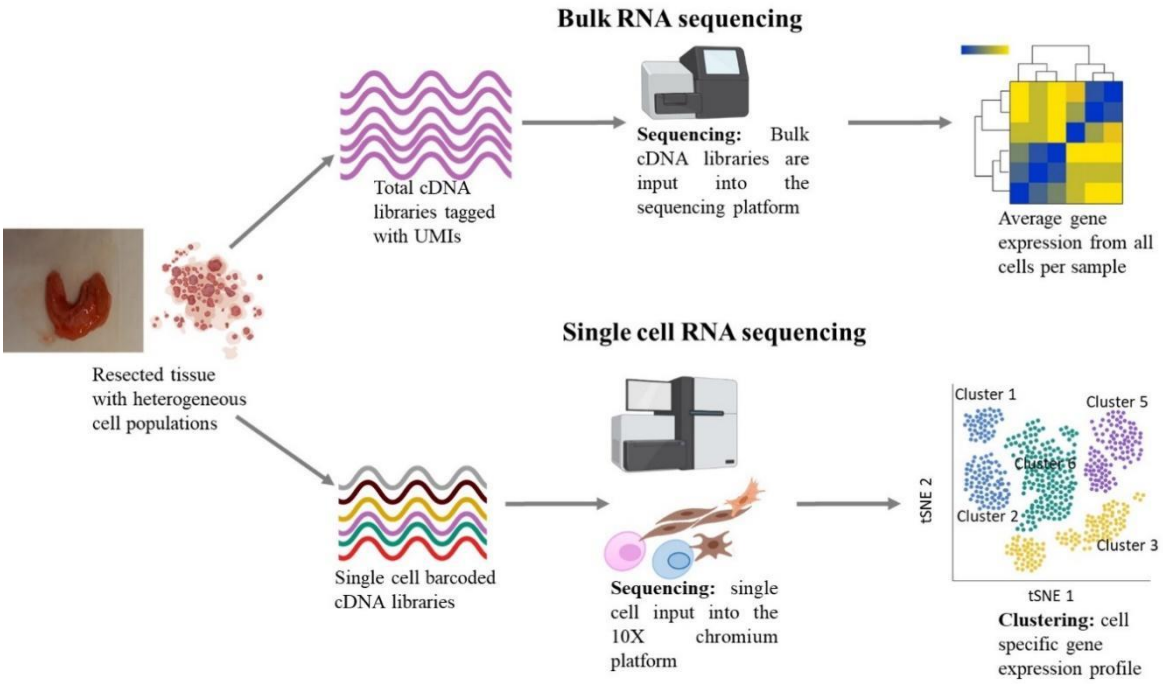


Figure 3.1: Schematic diagram that shows workflow involved in bulk RNAseq and scRNAseq processes. Bulk RNAseq involves isolation of total RNA from tissue, each sample (total RNA) is labelled with a unique identifier, cDNA libraries are prepared ready for sequencing. scRNA-seq involves isolation of single cells from tissue, each cell is then labelled with a unique identifier, cDNA libraries are prepared ready for sequencing.

3.3.2.1 Bulk RNA-sequencing

Bulk RNA-seq was performed to understand transcriptome profiles and variation in gene expression in the proximal (P), stricture (S) and distal (D) segments from resected gut of CD patients. Bulk RNA-seq provides an insight into the transcriptome by measuring average gene expression based on RNA from a heterogeneous population of cells in each sample; in so doing, it is possible to study transcriptional changes in stricture compared to the non-stricture segments. Of the 9 samples selected for bulk RNA-seq, 7 included matched tissue from the stricture (fibrotic) region and those regions proximal and distal to it. Isolated RNA was checked for quality and integrity using a Nanodrop spectrophotometer, Qubit fluorometer (Life Technologies) and 2100 Bioanalyzer (Agilent Technologies). Extracted RNA was measured to

establish concentration, purity (A260/A280) and RNA integrity number (RIN) as seen in table 3.2; the RIN scale ranges from 0 to 10, with 10 indicating maximum RNA integrity. It is important to check RNA quality and integrity prior to sequencing because RNA can easily be digested in the presence of ubiquitous RNase enzymes. In general, RNA with an absorbance ratio (A260/A280) close to 2.0 and RIN above 8 is considered as good enough quality for sequencing (456). The QIAseq™ UPX 3' transcriptomic profiling kit used for this study enabled sequencing of ultralow amounts of cells or RNA from each sample (1 to 100 cells or 10 pg to 1 ng isolated RNA per samples), which remarkably increases transcriptome profile accuracy, specificity and sensitivity.

The 27 samples included in this study had a median concentration of 253.7ng/mL (range: 37.2-1675.5), a median A260/ A280 absorbance ratio of 2.02 (range: 1.89-2.07) which suggested that the RNA was suitably free of impurities and contaminants and a median RIN value of 7.6 (range: 5.8-8.8), which indicated if the RNA integrity was acceptable for sequencing which was fair quality of RNA. Samples FCD012-P (Proximal) was not analysed for RIN due to technical mistake; however, the sample was sequenced based on a high A260/ A280 absorbance ratio of 2.03. The variable range in RNA concentrations was an indicator of the heterogeneity of the samples. cDNA libraries were prepared and quantified ready for sequencing using the Nextseq Illumina platform.

Table 3.2: Showing concentration, absorbance and RIN values for RNA samples
(P= Proximal; S= Stricture and D=Distal)

Sample ID	Concentration (ng/µl)	Absorbance A260/A280	RIN
FCD003-P	426.4	2.03	7.1
FCD003-S	241.6	2.04	8.2
FCD003-D	75.1	2.02	7.8
FCD004-P	1675.5	2.06	7.7
FCD004-S	77.9	2.04	7.6
FCD004-D	399.3	2.02	7.6
FCD005-P	158.6	2.03	7.6
FCD005-S	107	2.03	7.6
FCD005-D	1137.6	2.09	8.0
FCD006-P	228	2.02	8.2
FCD006-D	85.8	2.02	6.7
FCD009-P	253.7	2.02	7.2
FCD009-S	253.7	2.02	6.1
FCD009-D	78.4	1.94	7.1
FCD010-P	120.3	2.01	7.6
FCD010-S	163.7	1.99	7.6
FCD010-D	45.1	1.90	6.5
FCD012-P	309.5	2.03	-
FCD012-S	590.1	2.10	8.0
FCD012-D	519.8	1.99	8.1
FCD015-P	1004.5	2.07	5.8
FCD015-D	37.2	1.89	7.0
FCD016-P	316.3	2.03	8.3
FCD016-S	265.7	2.02	8.1
FCD016-D	1305.4	2.07	8.1
FCD017-P	116.8	2.00	8.4
FCD017-S	251.7	2.02	6.8
FCD017-D	451.3	2.03	6.2
FCD019-P	282.5	2.05	8.8
FCD019-S	212	2.01	8.0
FCD019-D	259.2	2.04	8.7

A summary of the analysis process is highlighted in figure 3.2, fastQ sequencing files were obtained through BaseSpace (Illumina) and reads de-multiplexed, cleaned, aligned, quantified and normalised to human genome using the CLC Genomics Workbench (Qiagen). Dr Animesh Acharjee, a senior research fellow from the institute of cancer and genomics performed the data analysis. Quality control (QC) was performed on the data by filtering out lowly expressed genes (≤ 1 transcript), and all samples passed the QC process; filtered data was then subjected

to linear models for microarray data (limma). Limma is a bioconductor package written in R and used for identification of differentially expressed genes (DEGs) to investigate transcriptional changes and differences in gene expression profiles in the stricture, proximal and distal segments.

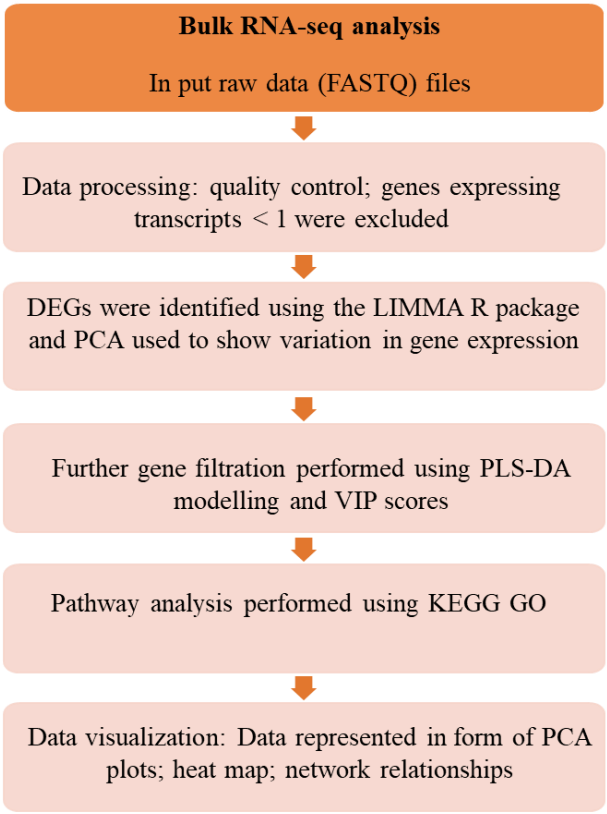


Figure 3.2: An outline showing steps involved in bulk RNA-seq data analysis. Raw data in the form of FASTQ files is subjected to quality control and cell filtration, DEGs were identified using the limma- R package and visualised variation in gene expression patterns determined using PCA and PLS-DA. Pathway enrichment analysis was performed based on KEGG and GO reference data bases. Data was visualised in PCA plots, heat map and bar plots

3.3.2.2 Principal component analysis of gene expression data revealed differences in gene expression profiles in the stricture, proximal and distal segments

To study variation in gene expression of the 3 segments, a total of 501 DEGs were analysed by principal component analysis (PCA). PCA is a univariate regression analysis model that is used to show trends and similarities in the data set based on gene expression patterns; clusters are formed based on similarities in gene expression profiles. PCA reduces dimensionality of the data into principal components and is visualised as a 2-dimensional (2D) graph displayed as principal component 1 (PC1) vs principal component 2 (PC2); PCA is used to cluster samples into different groups based on their location in the PC1 vs PC2 plot, each PC is independent of

the other. Principal components are linear combinations of the data's original variables; PC1 is the matrix projection containing features with the largest variance along the first axis. For examples, PC1 contains a combination of genes that contribute to the highest variation within samples along the PC1 axis. Principal component 2 (PC2) contains a combination of genes that result in the second largest sample variation within the data set. Therefore, PC1 vs PC2 plots show the highest spread of the data resulting in better recognition of similarity or dissimilarity in gene expression profiles across samples (457). Figure 3.3A shows unsupervised PCA plots in which the training data set for the analysis algorithm was not provided with specific information (labels) or variables about the data. The algorithm in unsupervised PCA analysis uses machine learning to detect hidden patterns in the data to show similarities or differences without information on the different variables. The PCA analysis scores in figure 3.3A show a 47.2% and 10.3% variance as per PC1 vs PC2 respectively. Samples from the stricture segment clustered together and apart from the non-stricture (proximal and distal), the non-stricture segments formed a discrete cluster which suggests similarity in gene expression profiles in the proximal and distal segments as compared to the stricture.

Supervised PCA analysis was performed where labels and specific information about the data set such as the group to which the samples belong (stricture, proximal and distal) is input during training of the algorithm. PCA plots (figure 3.3B) showed a 47.1% and 5.8% variance for PC1 and PC2 respectively. Supervised PCA plot showed tighter clustering of samples from the proximal and distal segments as compared to stricture samples. A wide range of variability in gene expression within the stricture samples was also observed, this could be attributed to samples collected at different stages of fibrosis and hence undergoing different biological processes and mechanisms despite all the tissue being fibrotic.

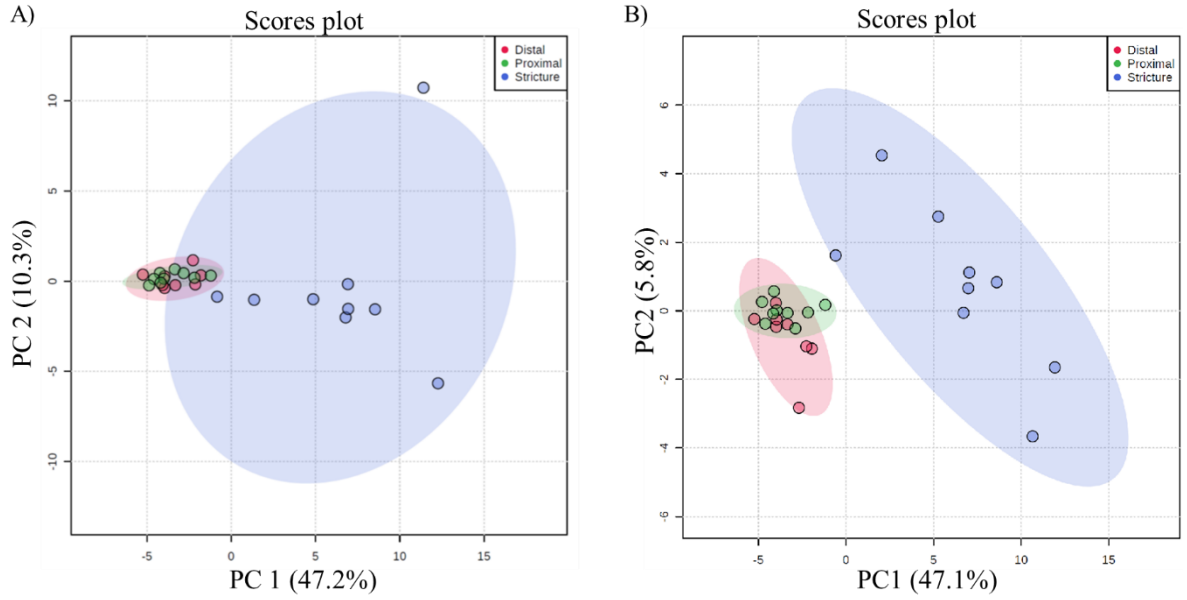


Figure 3.3: Transcriptome analysis showed differential gene expression profiles in the stricture, proximal and distal segments of resected gut. RNA was isolated from matched proximal, stricture and distal segments. cDNA libraries were prepared and sequenced using the 3'RNA-seq kit and Illumina NGS. **A)** Unsupervised principal component analysis (PCA) showing an overview of gene expression patterns in the proximal (green circles), stricture (blue circles) and distal segments (red circles). **B)** Supervised PCA showing differences in gene expression of the stricture (blue circles) to proximal (green circles) and distal (red circles) which appear similar in their pattern of gene expression proximal n=9, stricture n=8 and distal n=8

3.3.2.3 Variability in gene expression profiles within samples in each segment

To study gene expression patterns in samples between and within each segment, pairwise supervised partial least-square discriminant analysis (PLS-DA) was performed. PLS-DA is a multivariate regression analysis model that is suitable for discrimination of multiple variables according to their gene expression profiles. PLS-DA plots in figure 3.4A show distinct gene expression patterns between the proximal and stricture with PC1 and PC2 accounting for a 37% and 7.2% variance respectively. Although samples from each segment clustered together as seen in PCA plots in figure 3.4A-C, there is variation in gene expression within each group: proximal being much more similar which is evident through how close the samples cluster together based on gene expression profile as compared to those in the stricture (figure 3.4A).

Pairwise comparison of distal and stricture showed a similar pattern with PC1 and PC2 accounting for 38.7% and 7.5% variance respectively (figure 3.4B). Pairwise comparison of distal and proximal in figure 3.4C showed a similar pattern with PC1 and PC2 accounting for a 35.8% and 11.8% variance respectively. The distal samples seemed to spread across the cluster while the proximal samples clustered into 2 separate groups. PLS-DA analysis of the data set showed differences in gene expression patterns in each segment. The data also showed separation in gene expression profiles within samples from the same segment which demonstrates the heterogeneity and variability in the transcriptome of proximal, stricture and distal tissue.

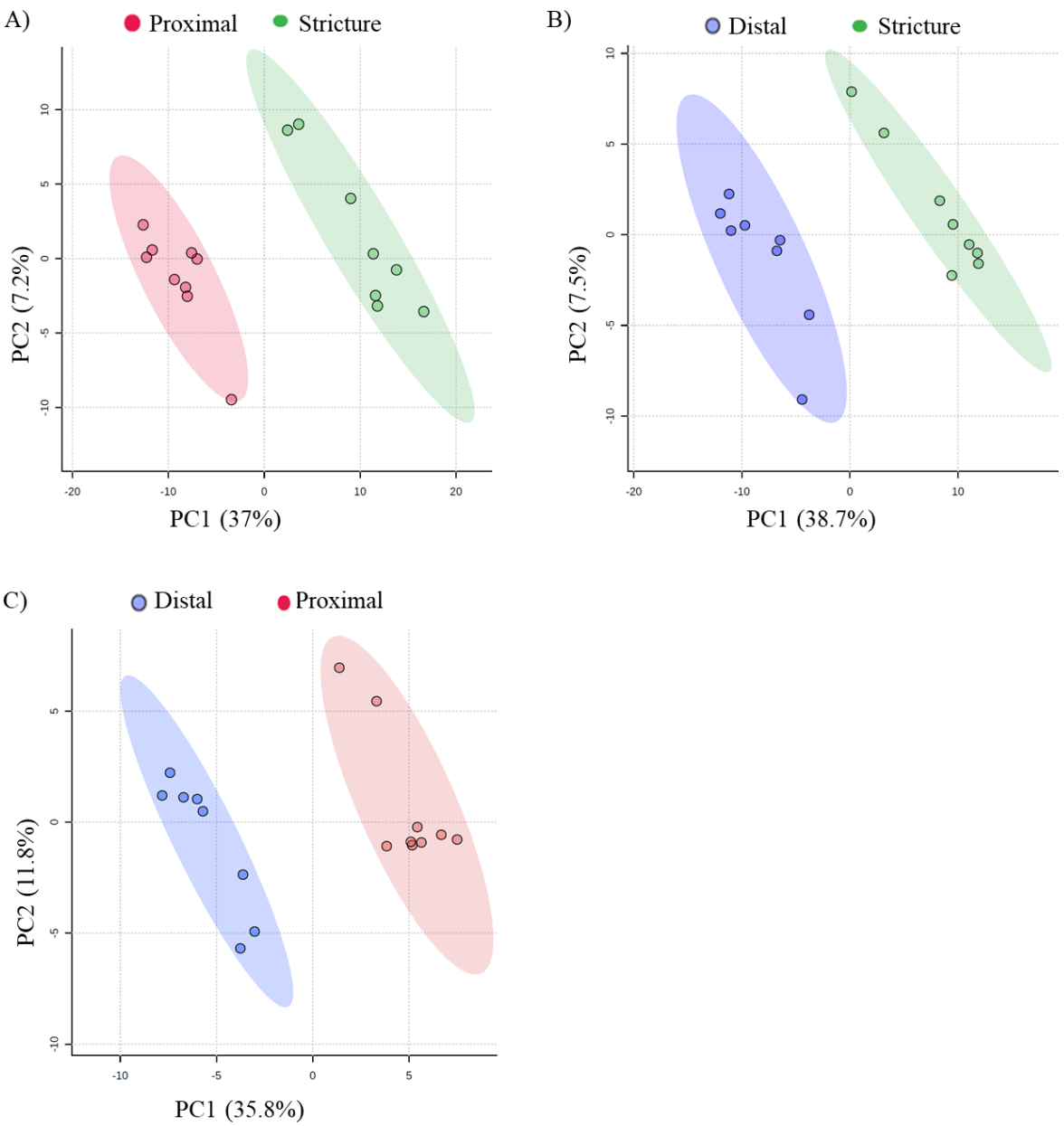


Figure 3.4: Supervised PLS-DA of DEGs showed variation in gene expression between segments and within samples from the same segment. Differentially Expressed Genes (DEGs) identified through limma-R analysis were subjected to PLS-DA to classify gene expression profiles between and within segments A) Supervised PLS-DA comparing proximal vs stricture with PC1 and PC2 variance at 37% and 7.2% respectively B) Supervised PLS-DA comparing distal vs stricture with PC1 and PC2 variance at 38.7% and 7.5% respectively C) Supervised PLS-DA comparing distal vs proximal with PC1 and PC2 variance at 35.8% and 11.8% respectively (proximal n=9; stricture n=8; distal n=8)

3.3.2.4 Assessing performance of the PLS-DA model

To evaluate performance of the trained PLS-DA predictive model, goodness of fit (R2) and predictive variation or goodness of prediction (Q2) matrices were measured for each pairwise comparison as seen in table 3.3. R2 and Q2 are usually used in combination to test the robustness of the algorithm that was developed based on a training data set and determine if it is good enough to reproduce similar results when used to analyse a test data set. Determining R2 values helps explain variation in a given variable in comparison to another while Q2 gives an insight of how well a variable can be predicted. R2 and Q2 values range from (0-100 %), the closer to 100% both values are, the more robust the model is. A good model works independently of the specific data set that was used to train the algorithm.

The PLS model used in the analysis was designed by making comparison for the 3 segments. R2 and Q2 values were determined for each comparison. R2 and Q2 values for the proximal vs distal comparison were 47% and 92.2%, these values suggest that 47% of variation in the proximal is explained by variation in the distal, the Q2 value at 92.2% shows a good predictive value in the results for the comparison. Similar R2 and Q2 values for the distal vs stricture and the proximal vs stricture comparison with R2 values of at least 50% and Q2 values of at least 85%. These 2 matrices confirm that the model is well suited to analyse the data and can be used for another test data set (458).

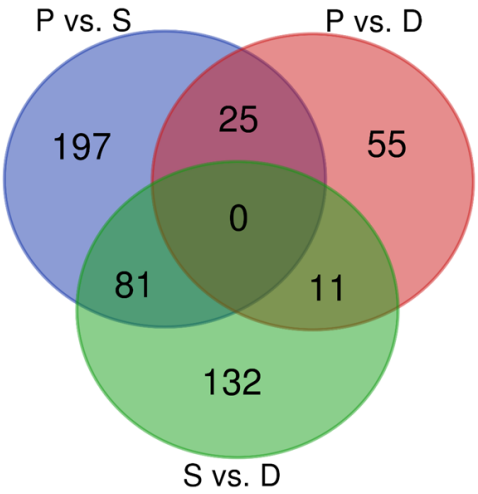
Table 3.3: PLS model validation; model variance (R2) and predictive ability (Q2) of each of the comparisons

Comparison	Model variance (R^2)	Predictive ability (Q^2)
Proximal vs. Distal	47%	92.2%
Distal vs. Stricture	51.8%	85.1%
Proximal vs. Stricture	50%	87%

3.3.2.5 Differentially expressed gene identification

A total of 501 differentially expressed genes (DEGs) were identified in the 3 segments as seen in the Venn diagram in figure 3.5A, which illustrates the overlapping associations of DEGs in a pairwise comparisons of the 3 segments. Figure 3.5B shows DEGs that were upregulated and downregulated in each pairwise comparison. The highest number of DEGs was seen in the proximal vs stricture comparison with a total of 303 DEGs, of which 126 were upregulated and 177 downregulated in proximal as compared to the stricture respectively. The lowest number of changes was seen in the proximal vs distal comparison with a total of 91 DEGs, of which 46 were upregulated and 45 downregulated in proximal as compared to the distal respectively. There was no gene expression overlap in the comparison of all the 3 segments. The highest number of DEGs in common was 81 as seen in the proximal vs stricture and stricture vs distal comparisons. The overlap in gene expression could be due to similar biological processes that might be occurring and shared across the 3 segments.

A)



B)

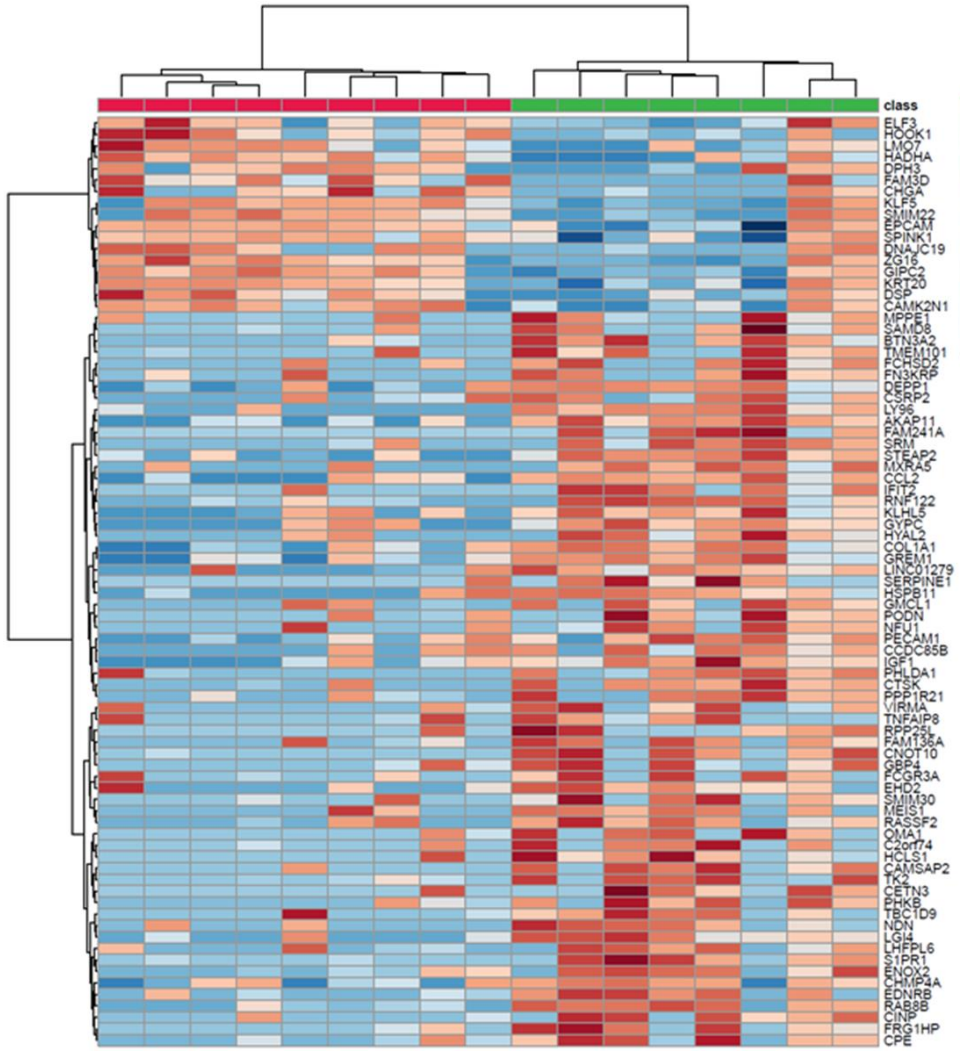
Segments	Total DEGs	Summary
P vs S	303	126 upregulated 177 downregulated
P vs D	91	46 upregulated 45 downregulated
S vs D	224	163 upregulated 61 downregulated

Figure 3.5: Transcriptome analysis showed differential gene expression profiles in the stricture, proximal and distal segments of resected gut. RNA was isolated from matched proximal, stricture and distal segments. cDNA libraries were prepared and sequenced using the 3'RNA-seq kit and Illumina NGS. **A)** Venn diagram showing differential gene expression across the three segments. The genes selected were based on an adjusted p value <0.05, blue shows the number of genes upregulated in the proximal (P) vs stricture (S), red shows the number of genes upregulated in proximal (P) vs distal (D) and green shows the number of genes upregulated in stricture (S) vs distal (D). Overlap regions show the number of genes upregulated in both comparisons **B)** Table corresponding to the Venn diagram showing the number of genes up regulated or down regulated in the 3 segments

3.3.2.6 A unique gene signature was upregulated in the stricture segment

To explore transcriptome profiles in the stricture, a group of 81 DEGs identified in the previous section which were regulated in both the proximal versus stricture and stricture versus distal comparisons were selected for further analysis. The selected 81 genes were subjected to hierarchical clustering and data visualised as a heat map as seen in figure 3.6, expression of the 81 genes was analysed in samples from the stricture and proximal respectively. Hierarchical clustering revealed a separation in gene expression profiles between stricture and proximal samples. There was a unique transcriptomic fingerprint of fibrotic related genes whose expression was upregulated in samples from the stricture segment. This suggests a distinct fibrotic molecular signature within the stricture segment. Examples of genes upregulated in the

stricture segments included *COL1A1* that encodes for the most abundant ECM protein collagen 1 and a key gene in regulating pathways that drive fibrosis such as ECM-receptor interaction and collagen formation, *CCL2* which encodes for macrophage chemoattractant protein which is important in antigen presentation through the recruitment of myeloid cells, a process that initiates and sustains chronic inflammation that eventually leads to fibrosis. *GREM1* which encodes Gremlin 1 protein known to play a key role in regulation of BMP/TGF-β1 signalling pathways was also identified, this was considered significant as the TGF-β1 signalling pathway is widely reported as a key pathway in fibrogenesis of multiple organs. Downstream signalling TGF-β1 results into activation, proliferation, and differentiation of MSCs particularly fibroblasts; the TGF-β/SMAD signalling is the main mechanism that drives fibrosis (459–461). Other genes of interest that were upregulated in the stricture were *SERPINE1*, *PODN*, *PECAM1*, *IGF1* and *SIPR1* among others. These genes have also been reported in the literature to regulate pathways that are linked to fibrosis. The heat map in figure 3.6 also showed a clustering pattern in which samples from each segment (stricture and proximal) clustering together.



Gene Symbol	Gene Name
	S<P
CHGA	chromogranin A(CHGA)
DSP	desmoplakin(DSP)
HADHA	hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctional protein), alpha subunit(HADHA)
KRT20	keratin 20(KRT20)
KLF5	Kruppel like factor 5(KLF5)
	S>P
CCL2	C-C motif chemokine ligand 2(CCL2)
CAMSAP2	calmodulin regulated spectrin associated protein family member 2(CAMSAP2)
CTSK	cathepsin K(CTSK)
CHMP4A	charged multivesicular body protein 4A(CHMP4A)
COL1A1	collagen type I alpha 1 chain(COL1A1)
CSRP2	cysteine and glycine rich protein 2(CSRP2)
EHD2	EH domain containing 2(EHD2)
EDNRB	endothelin receptor type B(EDNRB)
FCHSD2	FCH and double SH3 domains 2(FCHSD2)
GYPC	glycophorin C (Gerbich blood group)(GYPC)
GREM1	gremlin 1, DAN family BMP antagonist(GREM1)
HCLS1	hematopoietic cell-specific Lyn substrate 1(HCLS1)
HYAL2	hyaluronoglucosaminidase 2(HYAL2)
IGF1	insulin like growth factor 1(IGF1)
LY96	lymphocyte antigen 96(LY96)
MEIS1	Meis homeobox 1(MEIS1)
NDN	necdin, MAGE family member(NDN)
PECAM1	platelet and endothelial cell adhesion molecule 1(PECAM1)
PHLDA1	Pleckstrin homology like domain family A member 1(PHLDA1)
RASSF2	Ras association domain family member 2(RASSF2)
SERPINE1	serpin family E member 1(SERPINE1)
SRM	spermidine synthase(SRM)
S1PR1	sphingosine-1-phosphate receptor 1(S1PR1)
STEAP2	STEAP2 metalloreductase(STEAP2)

Key

- Epihelial barrier integrity
- Protein polymerisation
- ECM remodelling
- Cell differentiation and tissue remodelling
- Inflammation
- Stress response
- Adipogenesis
- Metabolism
- Other

Figure 3.6: Hierarchical clustering of 81 genes to classify gene expression patterns in the stricture and proximal segments. A) The 81 DEGs shown in the heat map were selected from the pairwise comparisons of the three segments (P vs S and S vs D). Gene expression profiles were analysed by hierarchical clustering in samples from the proximal (red) and stricture (green). The colour range indicates expression level - from deep red (high levels in gene expression) to deep blue (low levels in gene expression). **B)** Pathways enriched in the 81 genes and their associated genes, the pathways are coordinated to a colour code and matched with specific genes that are important in the pathway. S<P indicates reduced expression in stricture *vs* proximal, S>P indicates elevated expression in stricture *vs* proximal tissue.

3.3.2.7 Gene enrichment analysis identified important biological processes in the stricture, proximal and distal segments

Gene network interaction analysis was performed on a list of DEGs to identify variation in biologically relevant processes that are enriched in communities of genes linked to cell signalling in inflammation and fibrosis within the three segments. Network interaction analysis was performed on DEGs in proximal vs stricture, stricture vs distal and proximal vs distal segments. The Infomap algorithm has been used to detect communities and their interactions within large datasets; therefore, DEGs were analysed using Infomap with pairwise comparison in the three segments to study gene to gene interactions (figure 3.7).

Each node represents an “enriched term”, and the edge represents the connections between two enriched terms that have a gene-overlapped ratio greater than 0.5. Different modules in different colours represent nodes belonging to specific communities in the gene interaction network, which were defined using Infomap algorithm as seen in figures 3.7 A (proximal vs stricture), B (stricture vs distal) and C (proximal vs distal). The network is in a circular layout according to the gravity of the two nodes. The list of enriched terms in each module were identified based on pathway analysis using the Gene ontology biological pathways (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) reference databases (462). These

databases link DEGs to the relevant biological processes, molecular function and cellular components. A list of biological processes that are over-represented within each module were identified. The identified biological processes in each module were further represented in bar plots as seen in figures 3.8A (proximal vs stricture), B (stricture vs distal) and C (proximal vs distal) respectively. A total of nine different modules were identified for each comparison, with one non-identified (unassigned) module. In the bar plots, each enriched term (pathway) is represented by a bar where the length of the bar relates to its' enrich ratio. The enrich ratio is calculated as a ratio of the "number of observed genes compared to expected genes that regulate the enriched term (pathway). Pathways that show a high number of significant genes being expressed will have a high enrich ratio and most likely a high level of importance in regulating different biological processes.

Within each enriched term, there is a group of genes that are differentially regulated - Figure 3.8A shows a pair wise comparison of proximal and stricture indicating pathways that are differentially expressed in the two segments. An example is the HIF-1 signalling pathway that regulates hypoxia, and studies have shown that chronic hypoxia is involved in the pathogenesis of fibrosis through upregulation of ECM related genes such as *COL1A1*, *COL4A1*, *COL4A2*, *COL9A1* and *FN1* (463). Understanding the metabolic shifts and pathways that influence cellular changes in an inflammatory or fibrotic state are important in identification of potential therapeutic targets.

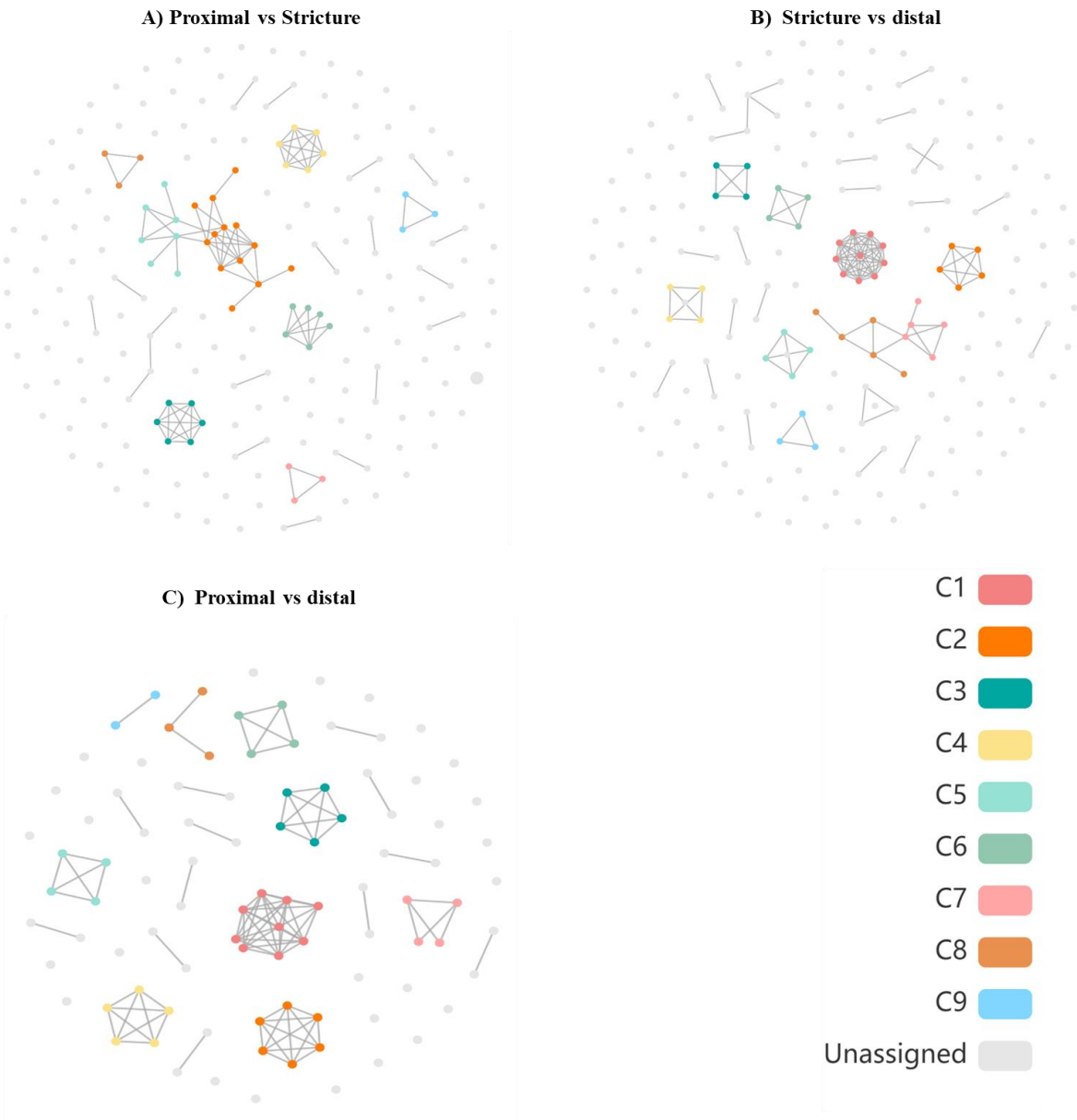


Figure 3.7: Enriched term connections visualised in circular network. Pairwise comparison between segments **A)** proximal vs stricture **B)** stricture vs distal and **C)** proximal vs distal. Each node represents an enriched term, and the edge represents the connections between two enriched terms that have a gene-overlapped ratio more than a specific cut off (>0.5). Different modules in different colours represent nodes belonging to specific topologic communities in a structure network, which were defined using Infomap algorithm. The network is in a circular layout according to the gravity of the two nodes.

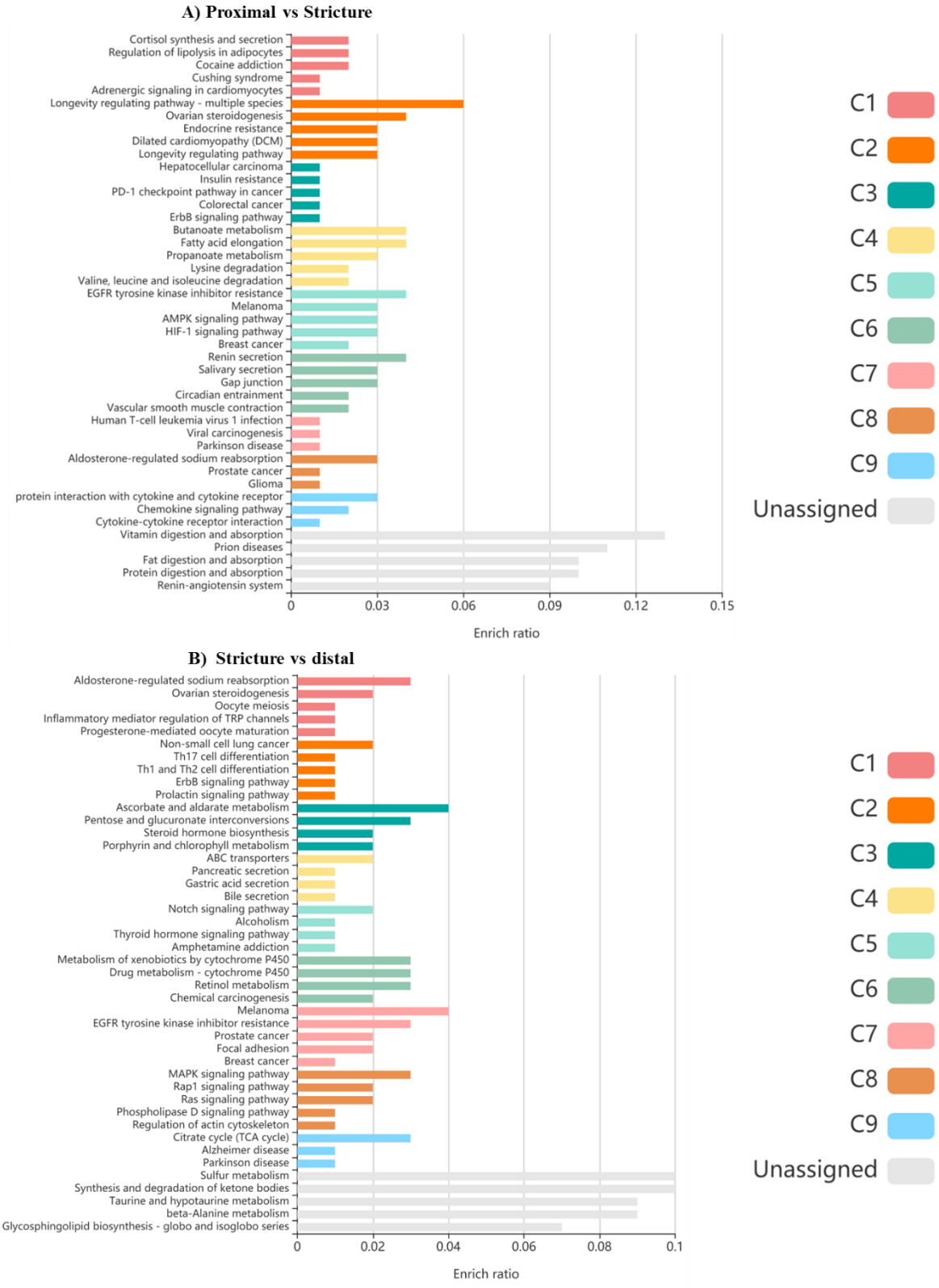


Figure 3.8: Figure legend on the next page

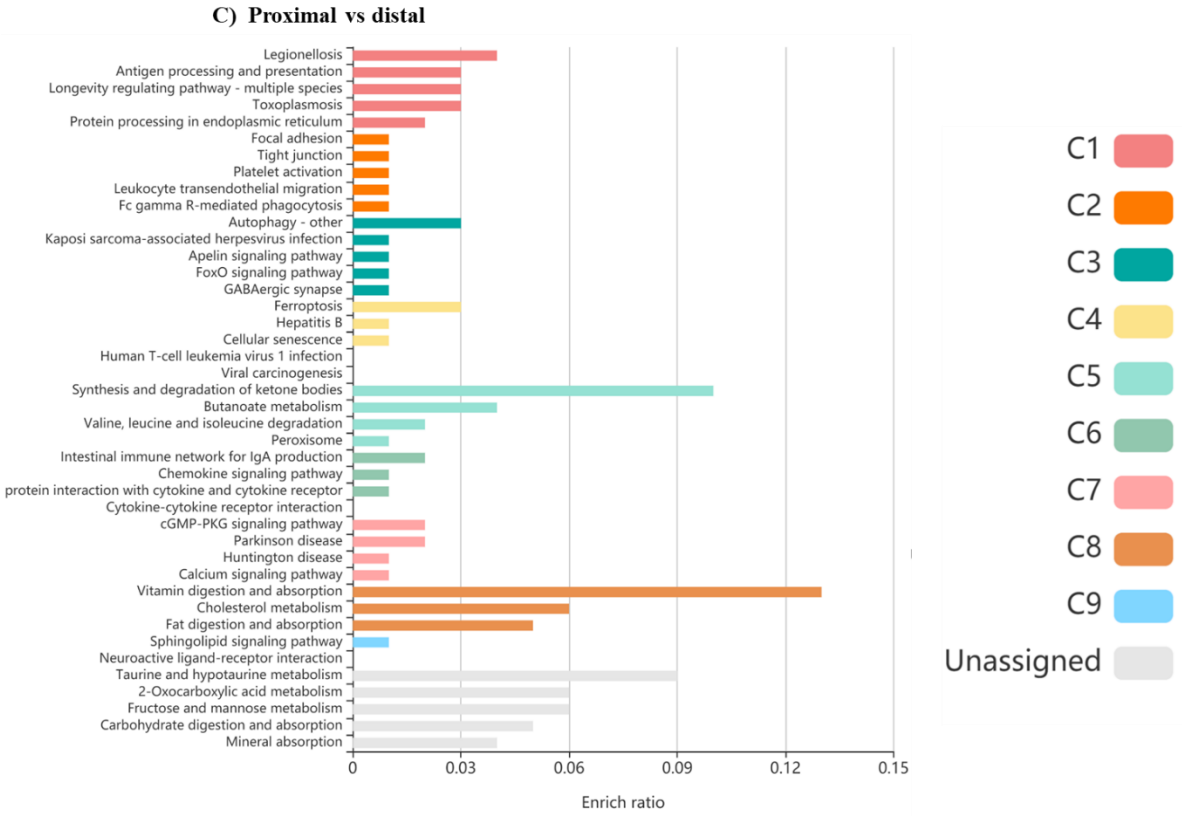


Figure 3.8: DEGs identified as significant (FDR adjusted p<0.05) for the segment pairwise comparisons. Bar plots represent classification of significant DEGs in the gene ontology (GO) database for every pairwise comparison. Each bar represents an enrich term (module/GO term) plotted to show the enrich ratio to determine overall upregulation of the module **A)** proximal vs stricture **B)** stricture vs distal and **C)** proximal vs distal.

3.3.2.8 Selection of DEGs as predictors of stricturing fibrostenosing disease

A group of genes were selected as predictors of stricturing in intestinal fibrosis, these included *LY96, GREM1*, *SERPINE1*, *HDAC1* and *FGF2*. The group was selected based on the highest ranked genes in the list of DEGs with a cut off set at a p value <0.05 and their biological association to pathways that drive a profibrotic phenotype. To test model performance for prediction of stricturing, AUC (Area Under the Curve) and confidence interval were calculated as listed in table 3.4. AUC values measure overall accuracy of a model for prediction based on sensitivity (measure of true positives) and specificity (measure of true negatives), where the higher the sensitivity and specificity, the higher the accuracy. This means that if an individual

is randomly selected from the group and subjected to gene transcriptome analysis, the 5 selected genes can potentially predict development of stricturing disease based on the AUC values. AUC values range from 0-1 with values close to 1 indicating an excellent model for prediction while those close to 0 show the model as a poor predictor of stricturing in CD. Pairwise comparisons for corresponding AUC values for each gene and a combination of all the genes were calculated. *GREM1*, *SERPINE1* and *FGF2* showed the highest potential to predict intestinal tissue stricturing with AUC of *GREM1* (0.88), *SERPINE1* (0.77), *FGF2* (0.7). *HDAC1* and *LY96* had lower AUC values of 0.65 and 0.69 respectively. These genes gave a high combined predictive AUC value of 0.81 at the 95% confidence interval, which suggests that the group of genes can be used as predictors of stricturing disease, especially *GREM1* which had the highest AUC value.

The selected genes are involved in various biological processes in inflammation and fibrosis. Current research in various fibrotic conditions such as the kidney, liver, heart and lung are focused on biomarkers encoded by these genes: Gremlin 1 encoded by *GREM1* is of interest for its role in the TGFβ/BMP signalling pathway. PAI-1 protein encoded by *SERPINE1* is of interest due to its protective role in cardiac fibrosis (464), and HDAC1 plays a major role in epigenetic modulation through mechanisms such involving histone modification. *HDAC1* has been shown to promote the differentiation of MSCs such as fibroblasts and pericytes to myofibroblasts, a more activated phenotype with increased potential to produce excess ECM proteins which promotes a profibrotic microenvironment (465,466). By contrast, FGF2 has demonstrated antagonistic effects on activation of fibroblasts into myofibroblasts; Nintedanib an anti-fibrotic small molecule inhibitor targets Fibroblast growth factor receptor (FGFR1, 2 and 3) is currently licenced for the management of IPF patients (467). Based on these reasons, there was an interest to understand the biological pathways through which these markers modulate profibrotic processes and how their dysfunction links to progression of stricturing in

CD. Understanding these regulators of fibrosis and their associated pathways provides avenues to explore potential therapeutic targets for intestinal fibrosis.

Table 3.4: Table showing AUC for individual genes and a combined score

Genes	Proximal vs. Stricture		Distal vs Stricture		Non stricture vs stricture	
	AUC	C.I	AUC	C.I	AUC	C.I
GREM1	0.83	0.68-0.94	0.92	0.82-0.99	0.88	0.77-0.95
FGF2	0.77	0.54-0.90	0.67	0.48-0.87	0.7	0.55-0.84
SERPINE 1	0.78	0.59-0.92	0.77	0.60-0.91	0.77	0.59-0.90
HDAC1	0.61	0.42-0.78	0.68	0.50-0.86	0.65	0.48-0.80
LY96	0.67	0.48-0.84	0.7	0.50-0.87	0.69	0.50-0.85
Combined	0.73	0.51-0.95	0.85	0.61-0.1	0.81	0.56-0.96

3.3.2.9 Validation of DEGs by qPCR

The bulk RNA-seq data provided insight into potential regulators of fibrosis using bulk RNA-seq data, RT-qPCR was used to validate these findings. Nine samples used in bulk-RNAseq run and an additional five samples matched as proximal, stricture and distal (a total of 14 samples) were used for the validation process in which expression of the five DEGs selected to build a predictive model were analysed by RT-Qpcr, β- actin was used as a house keeping genes. The data was presented as median with a 95% confidence interval, statistical analysis was performed using the "Kruskal-Wallis non-parametric test". *GREM1* was highly expressed in the stricture as seen in figure 3.9A, proximal vs stricture (p value=0.004) and stricture vs distal (p value <0.001). *SERPINE1* was upregulated in the stricture as seen in figure 3.9D with a statistical significance (p value=0.009). There was no statistical significance in expression of *LY96* (figure 3.9C), *FGF2* (figure 3.9B) and *HDAC1* (figure 3.7E) though *LY96* and *FGF2* showed upregulation in the stricture. *HDAC1* expression was higher in the distal segment as compared to the stricture and proximal but this was not statistically significant.

Overall, qPCR data matched and validated the expression patterns seen in the bulk RNA-seq data and confirmed the suitability of the group of genes selected as predictors of stricturing disease: *GREM1*, *LY96*, *FGF2*, *HDAC1* and *SERPINE1* as shown in figure 3.7 is reproducible in comparison to the bulk RNA-seq. Particularly, *GREM1* and *SERPINE1* were upregulated in the stricture as compared to the proximal and distal segments (data expressed as ($2^{-\Delta Ct}$).

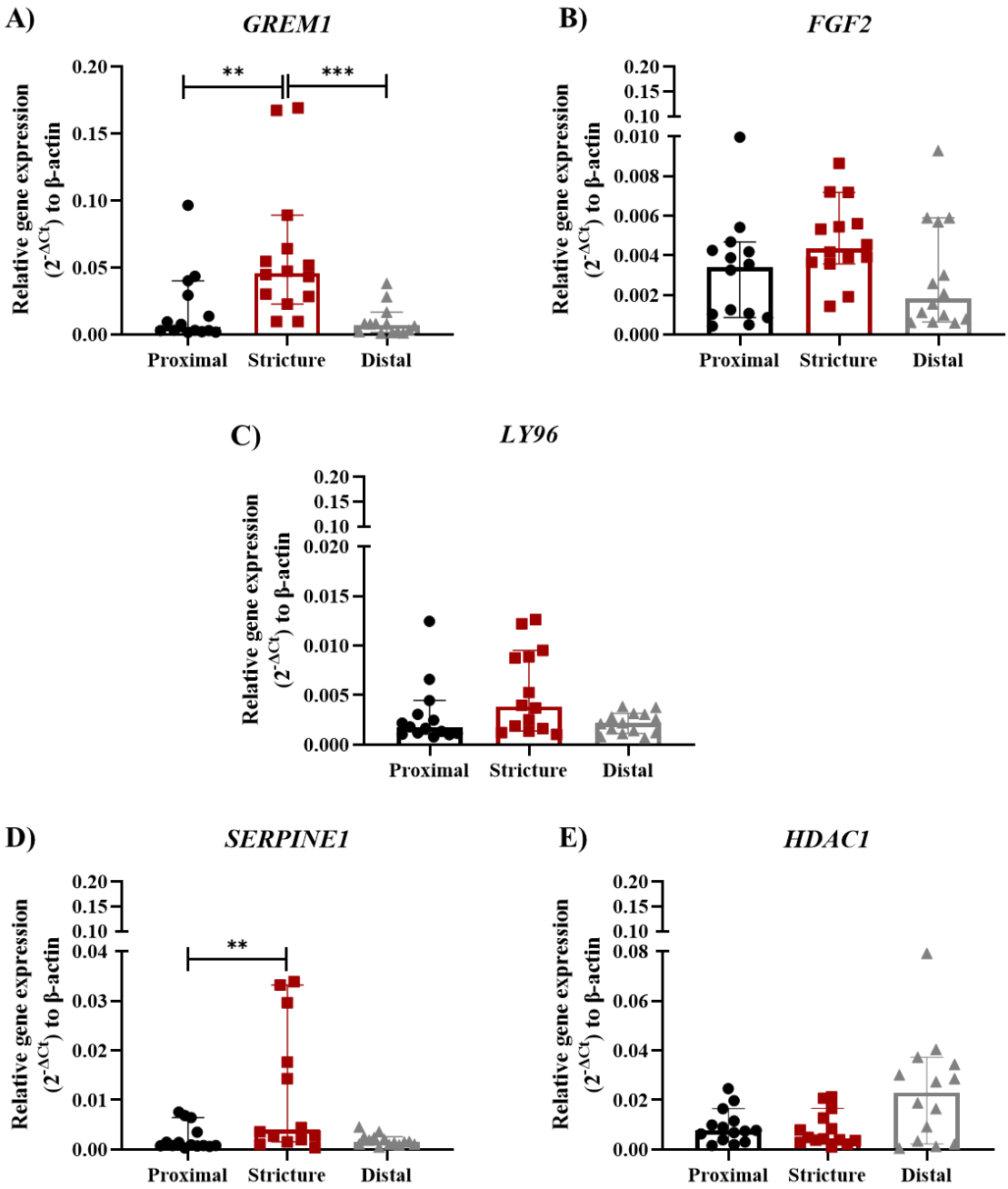


Figure 3.9: Validation of DEGs by qPCR. Gene expression profile of the selected biomarkers is similar in bulk RNA-seq and RT-qPCR data across the proximal, stricture and distal segments. The expression of (A) *GREM1*, (B) *FGF2*, (C) *LY96*, (D) *SERPINE1* and (E) *HDAC1* was studied by RT-qPCR analysis in n=14 matched samples from proximal, stricture and distal segments from the FCD cohort. Data shown as relative expression to β-actin as the endogenous control and reported as 2-ΔCt (where ΔCt = Ct of gene of interest – Ct of endogenous control). Data represented as median with 95% confidence interval. Statistical analysis was analysis was performed on PRISM using Kruskal-Wallis non-parametric test, and Dunn's multiple comparison test *p<0.05, ****p<0.0001

3.3.2.10 Determining sample size and power estimation analysis

For any study, it is critical to determine the appropriate sample size to answer a particular research question. For a gene to be considered differentially expressed, a particular quantity of transcriptome for the specific gene from a group of samples must be met. This is common particularly when a study has a small sample size which was the case for this study (468). To have confidence in the data and minimise false positive error rates, it is important to calculate the number of samples required to achieve gene specific effect, and therefore a retrospective power analysis was performed (469). A power analysis is used to estimate the smallest sample size needed for an experiment to a chieve a required level of significance and effect. It helps to determine if a result from an experiment is due to chance or genuine and significant (470). The power values calculated in figure 3.10 determine the probability that a specific gene will be differentially expressed when a specific sample size is used. According to our calculations, the probability that *GREM1* will meet the criteria turn out to be differentially expressed when 5 samples are used to determine *GREM1* gene expression by qPCR is 0.25, but when the sample size is doubled to 10 samples, the probability becomes 1.

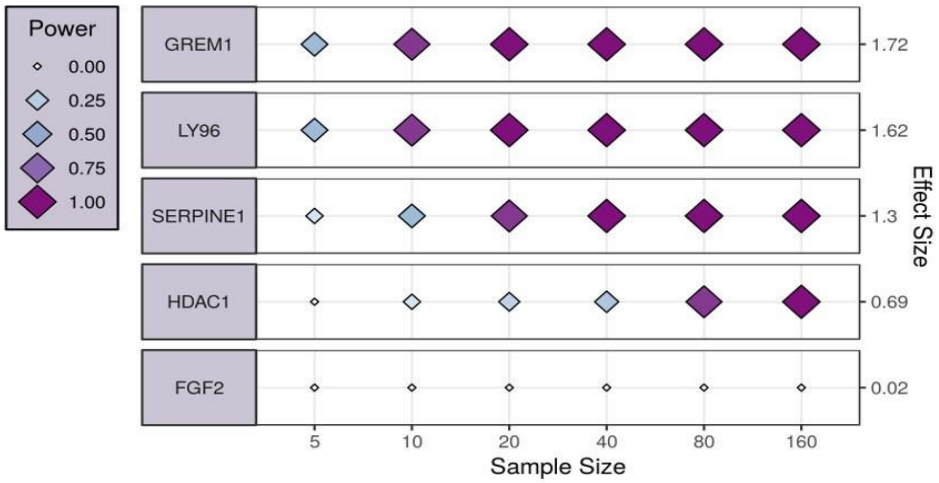


Figure 3.10: Figure shows calculations to determine sample size and power analysis. Power values for each gene are determined corresponding to the shape and colour of the diamond, Sample size is on the x-axis. The power values are the probability of using a specific number of samples to achieve a desired effect size.

3.3.3 Single cell RNA-seq to determine unique cell phenotypes in stricture and proximal tissues

Bulk RNA-seq as described above provided insight into the overall transcriptome profiles in the stricture, proximal and distal segments but due to the heterogenous nature of tissue in terms of its cellular composition, it is not possible to distinguish cell specific gene expression profiles and their contribution to disease pathogenesis. To determine cell specific changes and pathways that contribute to the fibrotic phenotype, gene expression profiles were studied in single cell populations from the proximal and stricture segments. Single cell RNA sequencing (scRNA-seq) has extensively been used to study transcriptional similarities and differences in individual cells and within a population of cells. This technique has led to identification of rare cell populations which would have been undetected by use of the traditional sequencing methods such as bulk RNA-seq. Moreover, during the process of scRNA-seq data analysis, unique molecular identifiers (UMI) are used to accurately track the mRNA molecules that have been transcribed, and barcodes are used to link the mRNA transcript to the cell of origin. Due to the efficiency in using UMIs and barcodes, it makes the technology very accurate even when very

few cells are analysed (471). Single cell RNA-seq was used to further understand transcriptome profiles at individual cellular levels in the proximal and stricture segments. In this section of the study, resected gut tissue from the stricture (n=3) and proximal (n=3) segments of CD patients were used to perform scRNA-seq. The bulk RNA-seq data suggested that the proximal and distal segments were most similar. Therefore, either could serve as a control relative to the stricture given the need to save costs. Single cells were isolated and sequenced on the 10X chromium platform, quality control was performed on the sequenced raw data to exclude cells that showed low quality transcriptomes and low viability, leaving such cells as part of the analysis can skew and create bias in the data.

3.3.3.1 Quality control and cleaning up of scRNA-seq data

A total of 31,195 cells were recovered from 6 samples after sequencing. Quality control (QC) was performed based on a set-criteria, including <5% mitochondrial gene expression and 200-2,500 cellular features which considered genes related to apoptosis, metabolism as well as ribosomal and mitochondrial genes. Based on these criteria, cells were classified as high and low quality. A total of 6,106 cells passed the set QC threshold and were included in further analysis. From the filtered cells, 8,387 unique genes were detected. Table 3.5 shows the number of cells pre and post QC for each sample and the number of genes used for further analysis. Cell numbers recovered for analysis were variable across samples. Many cells were lost during the QC process, and figure 3.11 shows the number of cells that were releasing mitochondrial content in each sample. Genes encoded in the mitochondria are commonly expressed and used in daily cellular physiological processes that are necessary in normal function of the cell. Their expression is estimated at 5-10% reads in each cell (472,473). If a cell is expressing mitochondrial genes higher than the normal range, it may overshadow the signal from low expression nuclear genes whose changes give an indication of possible biological alterations that a cell undergoes from normal to a diseased state. In addition, cells

undergoing stress and possibly cell death have shown increase in expression of mitochondrial genes. These factors constitute an exclusion-criteria for cells expressing high mitochondrial content.

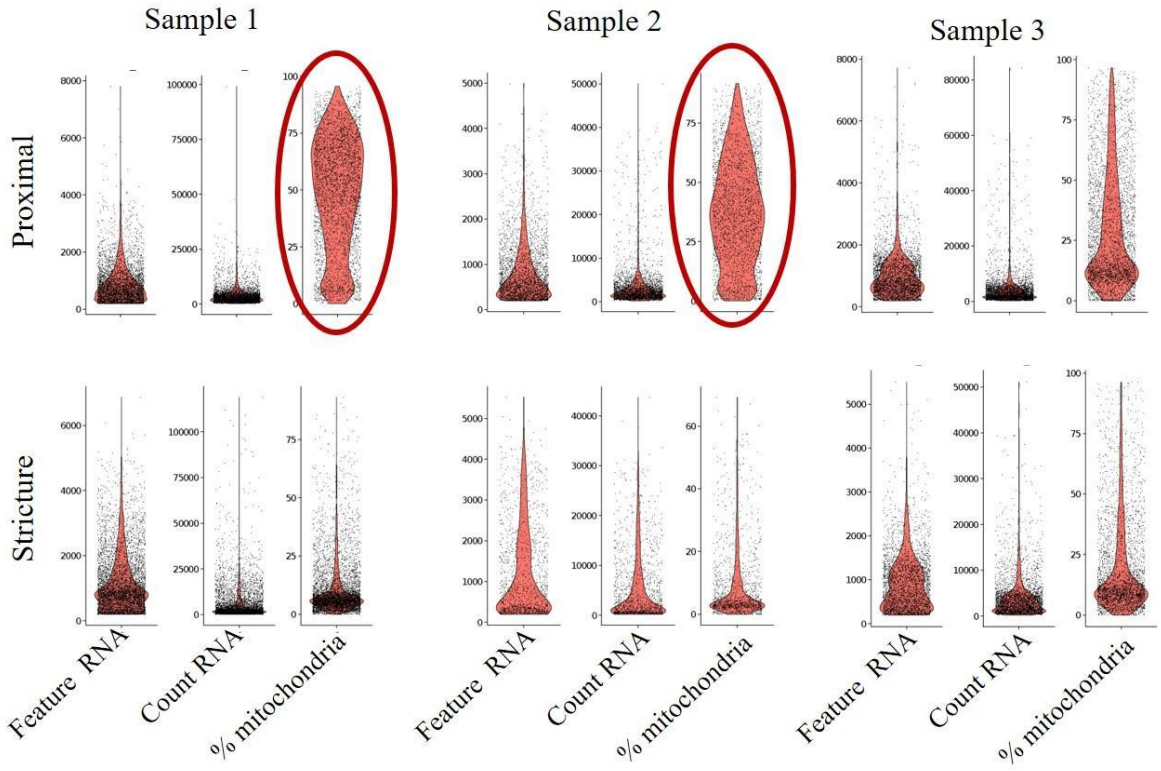


Figure 3.11: Quantity of mitochondrial content being released by cells from each segment.
Violin plot showing release of mitochondrial content.

The ideal number of cells required for scRNA-seq data analysis to acquire in depth information on the transcriptional state of cells or tissue is not clear. However, studies suggest as low as 50-10000 cells per sample can be used to perform scRNA-seq analysis but the biological question to be studied should be taken into account (474,475). Table 3.5 shows the variation in the number of filtered cells across all samples that were submitted for further analysis, with patient 3 having the lowest number of cells (320 cells obtained from the proximal region) and patient 1 having the highest number of cells (2990 cells obtained from the stricture). The number of filtered cells used for analysis were within the range previously used in several studies to analyse cell specific gene transcription profiles.

Table 3.5: Quality control was performed on the cells for data analysis

Sample	Number of cells recovered	Number of cells after filtering	Average number of genes detected per cell
Patient 1- Proximal	7808	584	1527
Patient 1- Stricture	10156	2990	1979
Patient 2- Proximal	4612	886	1494
Patient 2- Stricture	1488	813	1800
Patient 3- Proximal	3788	320	827
Patient 3- Stricture	3343	513	760

3.3.3.2 Gene expression profiling identified 14 distinct cell populations

To identify cell populations and modulatory mechanisms that are unique to inflammation and fibrosis in the stricture and proximal segments, scRNA-seq data was analysed using Seurat, an R-studio data analysis package. Due to the low cell numbers being used for analysis in each sample, cells were pooled from all six samples (Stricture n=3 and proximal n=3) for analysis, totalling 6,106 filtered cells. Seurat data analysis software uses dimensionality reduction and unsupervised uniform manifold approximation and projection (UMAP) clustering. In total, 14 different cell types consisting of both immune and stromal cell populations were identified as seen in figure 3.12. Unsupervised clustering was based on the principal that cells that clustered together have a similar gene expression profile and biological function. Identification of cell types in each cluster was based on cell defining markers that were highly expressed in respective clusters, gene lists from each cluster were compared to publicly available literature to check for phenotypic markers that are unique to a cell population (476). Table 3.6 shows reference markers used to identify cells in the different clusters.

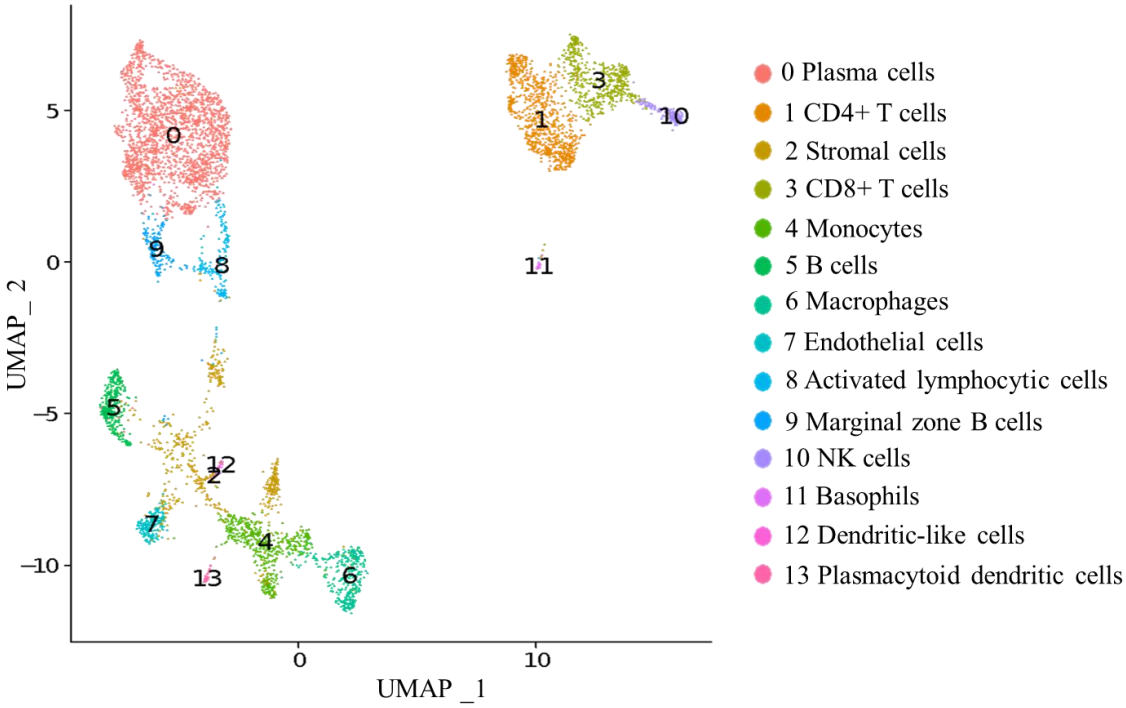


Figure 3.12: Multiple cell populations identified in both the stricture and proximal segments. scRNA-seq was performed on single cell suspensions isolated from resected stricture (n=3) and proximal (n=3) using the 10X chromium platform. Cell clustering was performed using Seurat (R-studio package) based on gene expression patterns and visualised using the Uniform Manifold Approximation (UMAP) plot. A total of 6,106 cells pooled from six samples (stricture n=3 and proximal n=3) were used to identify 14 cell clusters. Each cluster identified corresponds to a specific colour.

The data from this study showed immune cells constituting the highest frequency found in diseased tissue, in agreement with the published literature (477). Immune cell populations identified include lymphocytes (B and T cells), NK cells, monocytes and macrophages, basophils and dendritic cells. Based on gene expression, B cells were identified as cluster 5 expressing CD19 and CD20. Two subsets of B cells were identified, these include cluster 0 as plasma cells based on their expression of *CD39*, *CD79A* and several immunoglobulins as expected of plasma cells (478). Cluster 9 was identified as marginal zone B cells, and this cluster had a gene expression profile of immunoglobulin related genes and markers such as *MZB1* which are known to be expressed by marginal zone B cells. T cells were identified as

clusters 1 for CD4 T cells and cluster 3 for CD8 T cells respectively. Some of the genes that were expressed within the CD4 cluster such as *CCR6, RORA, LTB* and *FOXP3* indicated presence of Th17 and regulatory T cell phenotypes within this cluster. Cluster 8 was identified as activated lymphocytic cells due to the expression of lymphocyte markers at low expression levels and effector molecules such as *IFNG* and *GZMB*. Cluster 10 was identified as NK cells due to the expression of markers characteristic of this lineage including *KLRD1, NKG7* and *NCAM1* commonly known to encode for CD56 protein. Clusters 4, 6, 12 and 13 were identified as monocytes, macrophages, and dendritic cells respectively. Monocytes were identified based on *CD14, CXCL8* and *HLADR* expression. Macrophages were defined as *CD68* and *HLA-DRA* expressing cells while dendritic cells in cluster 12 were identified by their expression of *CD83* and *IDO1* which encodes for indoleamine 2,3-dioxygenase, an enzyme that is reported to modulate various pathophysiological processes and expressed in dendritic cells. Plasmacytoid dendritic cells were identified in cluster 13 using their expression of *IL3RA* and *IRF7*.

The identified immune cells play a key role in chronic inflammation, they activate resident stromal cells that contribute to tissue stiffness and formation of strictures. In addition to immune cells, stromal cell populations were also identified; these were classified in cluster 2 because of their expression of genes such as *COL1A1, COL1A2, VIM, EPCAM* and *PDGFRα* associated with fibroblasts, pericytes and epithelial cells. Endothelial cells were identified in cluster 7 with expression of *CD34*, a well-known mesenchymal cell marker and *PECAM1* which encodes for *CD31*. Having pooled cells together from all the 3 samples and segments, it was important to establish the origin of these cells in terms of samples 1 2 or 3 and either the proximal or stricture segment.

Table 3.6: Genes selected for cluster identification of the different cell populations

Cluster 0	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**	**Cluster 6**
Plasma cells	CD4+ T cells	Stromal cells	CD8+ T cells	Monocytes	B cells	Macrophages
CD38	*CD3*	*THY1, VIM*	*CD3D*	*CD14*	*CD79A*	*CD68*
CD79A	*CD4*	*COL1A1,COL1A2*	*CD8*	*CXCL8*	*CD74*	*HLA-DRA*
IGLV2-8	*CD2*	*COL18A1,FAP*	*CD2*	*IL-1β*	*CD19*	*IL-1β*
IGLL5	*IL-7R*	*PDGFRα, EPCAM*	*GZMB*	*HLA-DR^{low}*	*MS4A1 (CD20)*	*CD163*
IGKV4-1	*FOXP3*	*C-KIT*	*IFNγ*			
Cluster 7	**Cluster 8**	**Cluster 9**	**Cluster 10**	**Cluster 11**	**Cluster 12**	**Cluster 13**
Endothelial cells	Activated lymphocytic cells	Marginal zone B cells	NK cells	Basophils	Dendritic cells	Plasmacytoid dendritic cells
PECAM-1	*CD3,HLA-DR,*	*FCGR2B*	*KLRD1*	*TPSAB1*	*LAMP3*	*IL-3RA*
CD34	*CD79^{low},MZB1^{low}*	*MZB1*	*NKG7*	*MAS4A2*	*CD83*	*CLEC4C*
	IGHG1^{low},IGHA1^{low}	*IGHM*	*GZMB*		*IDO-1*	*NRP1, NRP2*
	CADM1,		*NCAM1*			*IRF7*
	TNFRSF17					

Figure 3.13 shows the gene expression pattern of the main markers used in the identification of specific clusters. The central figure shows 14 clusters that were identified and around it are UMAP plots that show gene expression pattens within clusters for some of the markers that were used to identify the various cell types.

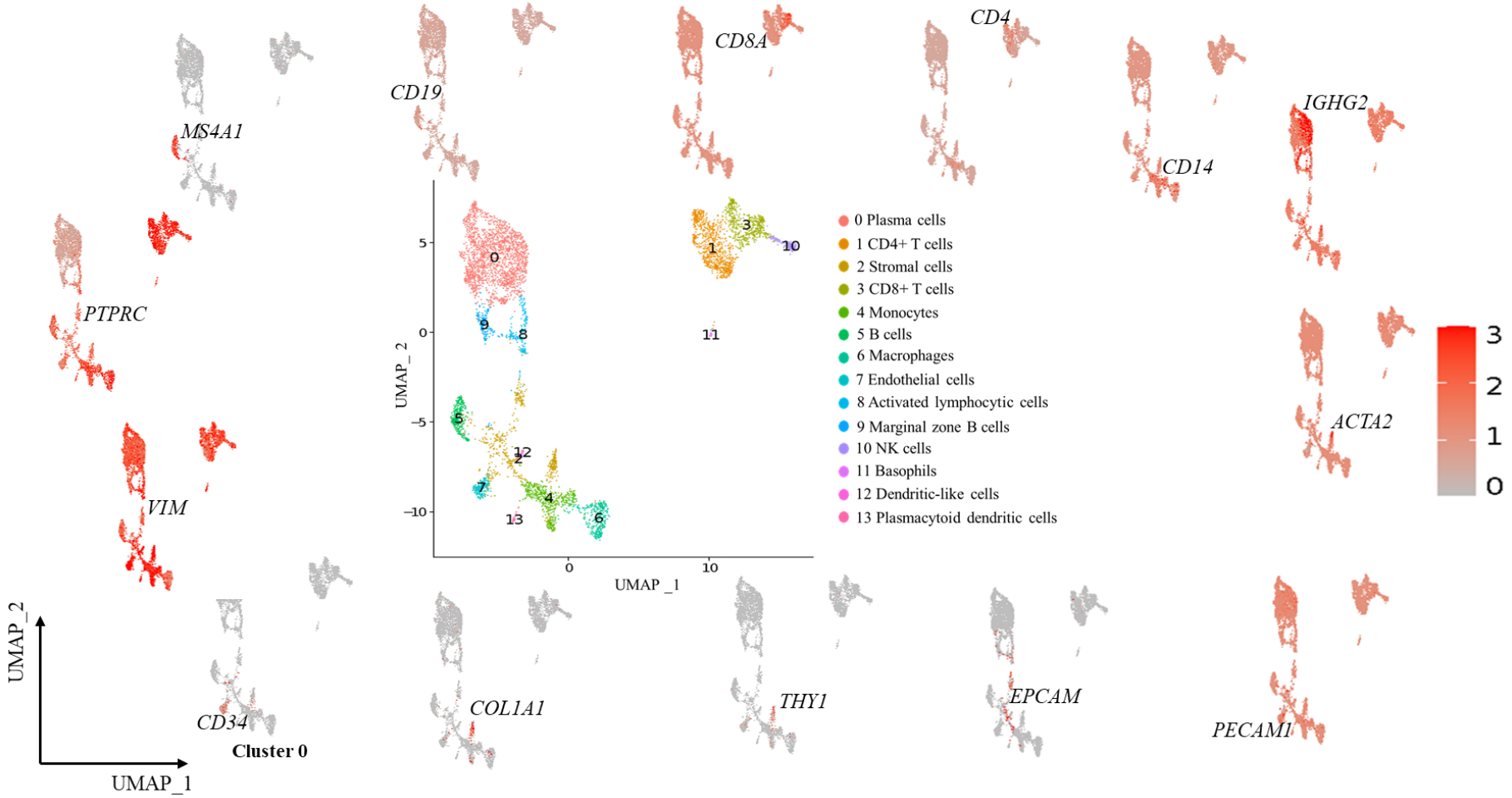

CD19
CD8A
CD4
IGHG2
MS4A1
CD14
PTPRC
ACTA2
VIM
CD34
COL1A1
THY1
EPCAM
PECAM1
UMAP_2
UMAP_1
Cluster 0
UMAP _2
UMAP _1
0 Plasma cells
1 CD4+ T cells
2 Stromal cells
3 CD8+ T cells
4 Monocytes
5 B cells
6 Macrophages
7 Endothelial cells
8 Activated lymphocytic cells
9 Marginal zone B cells
10 NK cells
11 Basophils
12 Dendritic-like cells
13 Plasmacytoid dendritic cells
3
2
1
0

Figure 3.13: **UMAP plots showing gene expression patterns of markers that were used to define 14 clusters that represent individual cell populations.** Each dot represents a single cell. The central plot shows already identified cell populations and surrounding it are UMAP plots showing transcriptional expression of some of the biomarkers used to identify specific cell populations across the 14 clusters. Genes with high transcripts being expressed corresponded to the scale (0-3) where high expression is towards the red colour (3) and no expression is towards grey (0)

3.3.3.3 Determining the origin of identified cell types corresponding to sample and segment

To study the distribution of cells across the identified 14 clusters, Seurat analysis software was used to track the origin of the cells in each cluster to the sample and segment (proximal or stricture) using UMIs and barcodes. The cell types identified from the 3 samples were distributed across clusters as seen in figure 3.14. However, samples 1 consists of the highest number of cells and cluster around cell types CD4+ T cells (cluster 1), CD8+ T cells (cluster 3) and NK cells (cluster 10) as seen in figure 3.14B. This was expected as the highest number of cells after QC was from sample 1. The lowest number of cells used for analysis was from sample 3 which is quite evident when looking at the distribution plots (figure 3.14B). Overall, cell distribution across clusters is even despite variation in cell numbers. Figure 3.14C shows cell distribution across clusters as per the segment of origin (proximal and stricture), the plot shows a similar distribution of cells from both segments across the 14 clusters. As expected, the cell frequencies from the stricture are more visible (green dots) as compared to those from the proximal (red dots). This was expected because of the 6,106 cells that were being used for analysis, 71% were from the stricture.

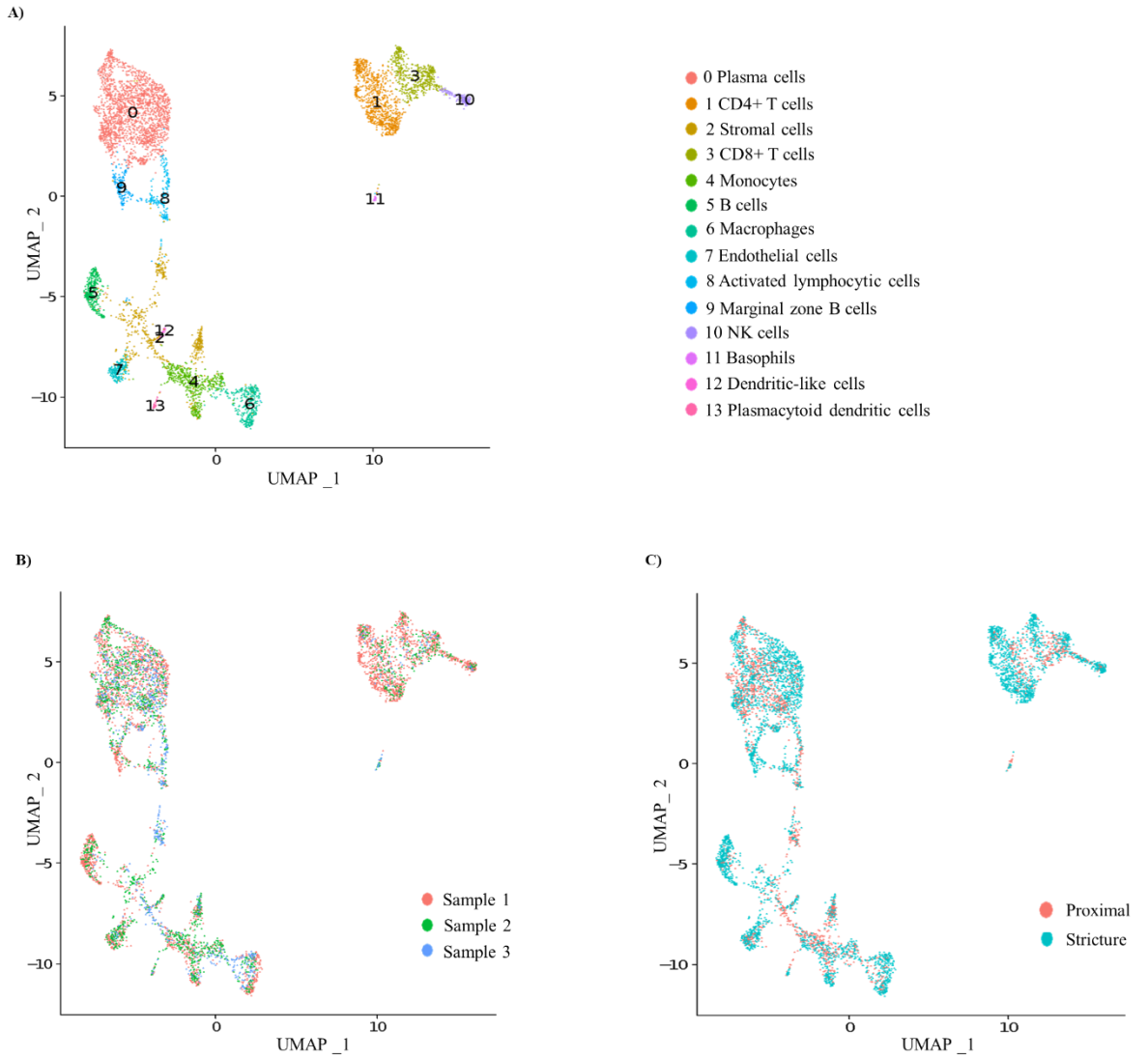


Figure 3.14: Even cell distribution across clusters. A) Plot showing identified cell types, analysis was based on cells pooled together from 3 samples, each sample containing matched tissue from the proximal and stricture **B)** With the use of Seurat and basing on unique molecular identifiers and barcodes, cells in each cluster were linked back to the sample of origin identified with a unique colour (red dots show sample 1 distribution, green dots=sample 2 and blue dots=sample 3) **C)** Plot shows distribution of cells from either the proximal or stricture segments across the different cell clusters, segment of cell origin is identified with a unique colour (red dots show cells from the proximal, green dots show cells from the stricture).

3.3.3.4 Increased expression of markers of immune cell activation across clusters

Immune cells have been shown as major drivers of chronic inflammation which is known to initiate the fibrosis process in CD (265). To understand the state of immune cell activation in the tissue microenvironment, gene expression patterns of activation markers within the 14

cluster were examined as seen in figure 3.15. Expression of activation markers such as *CD69*, *CD2* and various *HLA* molecules is an indication of an activated immune cell phenotype. Increased expression of *CD69*, which is an early marker of lymphocyte activation was seen in clusters 1, 3, 5 and 10, these clusters were identified as T cells, B cells and NK cells respectively (479).

Other molecules whose expression indicate an activated immune phenotype include *CD2* which is a costimulatory molecule for activation of T cells (480) and *IL1B* that is mostly expressed on antigen presenting cells such as monocytes and macrophages, these are important proteins in priming and activation of the adaptive immune response (481). Figure 3.15 shows expression of *CD2* was seen in clusters 1, 3, and 10 which were identified as CD4, CD8 and NK cells while expression of IL1B was seen in clusters 4 and 6 which correspond to monocytes and macrophages. Other molecules of importance especially in antigen presentation and the major histocompatibility complex (MHC) include CD74, HLA-DRB1, HLA-DRB5 and HLA-DRA, whose expression is increased in clusters 2, 4, 5, 7,12 and 13 which correspond to T and B cells, monocytes, macrophages and dendritic cells.

Figure 3.15 also shows a moderate expression of MHC Class II molecules including HLA-DRB1, HLA-DRB5 and HLA-DRA in cluster 2 which has been identified as stromal cells particularly fibroblasts, pericytes and epithelial cells. Epithelial cells are the primary barrier and first line of tissue defence, intestinal epithelial cells are important in regulation of tolerance while preserving immunity against pathogens. Studies have shown expression of class II MHC molecules on the surface of IECs which aligns with findings from our data (482,483).

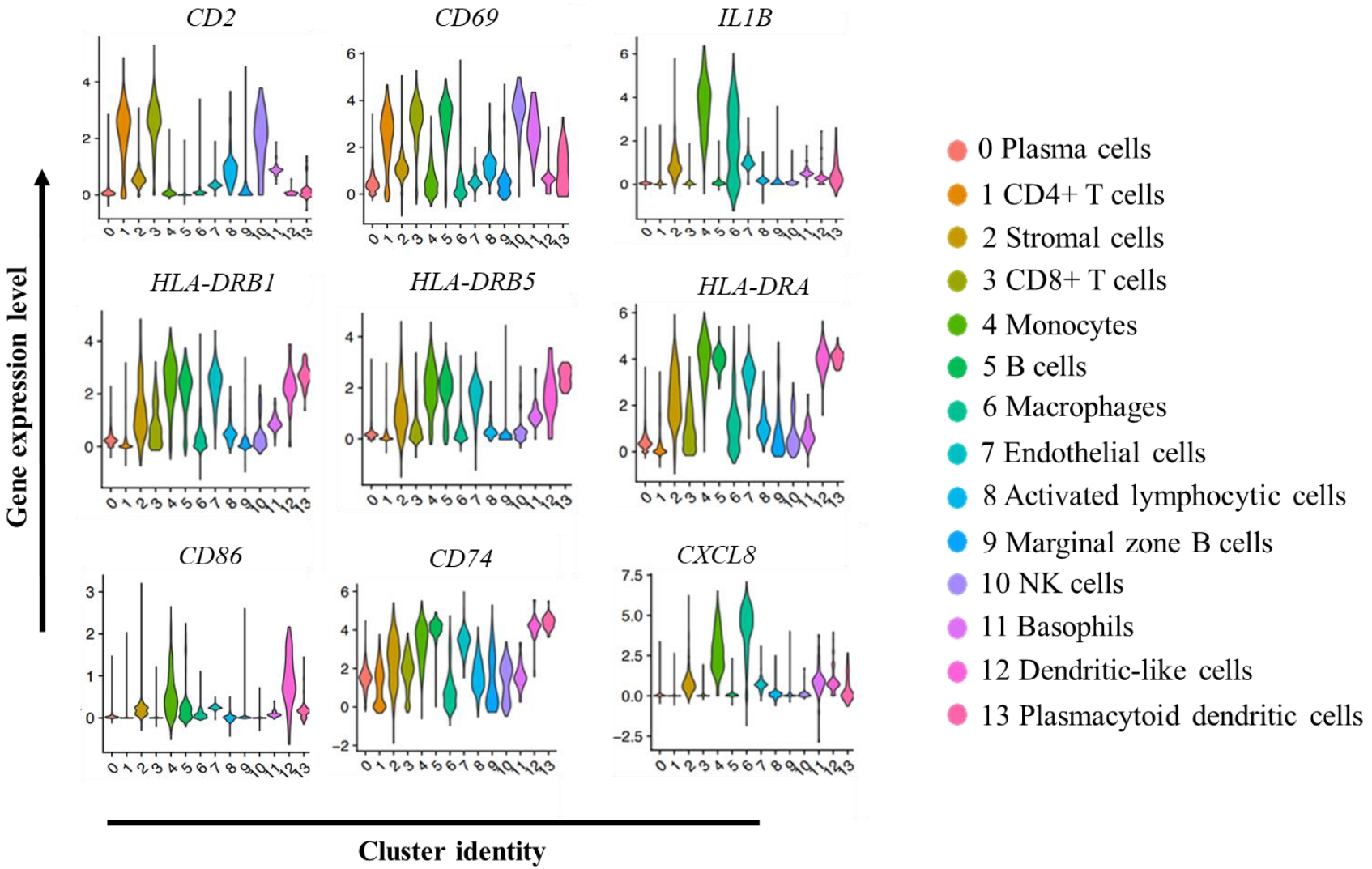

CD2
CD69
IL1B
HLA-DRB1
HLA-DRB5
HLA-DRA
CD86
CD74
CXCL8
Gene expression level
Cluster identity
0 Plasma cells
1 CD4+ T cells
2 Stromal cells
3 CD8+ T cells
4 Monocytes
5 B cells
6 Macrophages
7 Endothelial cells
8 Activated lymphocytic cells
9 Marginal zone B cells
10 NK cells
11 Basophils
12 Dendritic-like cells
13 Plasmacytoid dendritic cells

Figure 3.15: Violin plots show activation markers with increased expression across immune cell clusters. Each violin plot is representative of markers related to activation of immune cells, these include *CD2, IL1B, CD69, HLA-DRB1, HLA-DRB5, HLA-DRA, CD86, CD74* and *CXCL8*. All the 14 clusters are shown, each violin is coloured corresponding to a specific cluster. The x-axis shows the cluster identity while the y-axis shows gene expression levels which increase higher up on the axis

CXCL8 which encodes for IL-8, a chemokine that signals through receptors CXCR2 and CXCR1, was observed on clusters 4 and 6 that were identified as monocytes and macrophages, IL-8 is produced by macrophages and other cells such as epithelial and endothelial cells. IL-8 is a chemoattractant for neutrophils and is important in neutrophil recruitment.

3.3.3.5 Sub-clustering identified stromal cell populations in the data set

Immune cells are the main drivers of inflammation, but they also contribute to activation of stromal cells in sites of inflammation. The activated microenvironment leads to a chronic inflammatory response which if not resolved eventually results in formation of strictures in CD patients. Extensive research has been performed on mechanisms through which immune cells drive inflammation and fibrosis. However, less focus has been placed on the role of stromal cells in inflammation and fibrosis in CD. The scRNA-seq data was analysed to identify stromal cell populations in the proximal and stricture segments. Stromal cells are much lower in frequency in the tissue as compared to immune cells, and therefore it was challenging to detect most stromal cell markers during initial clustering because of the low cell numbers used to perform the analysis and the overall low expression of these markers.

Of the 14 clusters that were identified in the previous sections, clusters 2, 4 and 12 (figure 3.16A) exhibited expression of stromal like markers such as *CD34*, *THY1*, *COL1A1*, *COL1A2*, *PECAM1* and *EPCAM*. Through further analysis of cells from the three selected clusters, a

further 5 sub-clusters were identified as seen in figure 3.16B. Figure 3.16C below shows violin plots of representative genes expressed in each sub-cluster, with these biomarkers used for identification of the stromal cell populations represented in each sub-cluster. Genes expressed by cluster 0 were indicative of plasma cells, they include *IGHA1*, *IGHM* and *MZB1*. This may suggest an overlap in gene expression between immune and stromal cells during the initial clustering. Cluster 1 was identified as a mixture of endothelial and smooth muscle cells due to expression of *CD34*, *PECAM1*, *CAV1*, *ACTA2* and *VIM.* The cell populations in cluster 2 were identified as a possible mix of fibroblasts, myofibroblasts and pericytes due to their expression of genes such as *THY1 (CD90)*, *COL1A1*, *COL1A2*, *PDPN*, *PDGFRα* and *ACTA2*. Cluster 3 was identified as epithelial cells with expression of *EPCAM* and *VIM*. Cluster 4 showed a gene expression signature that was characteristic of dendritic like cells, these included genes such as *LAMP3*, *CD83* and *IDO1*.

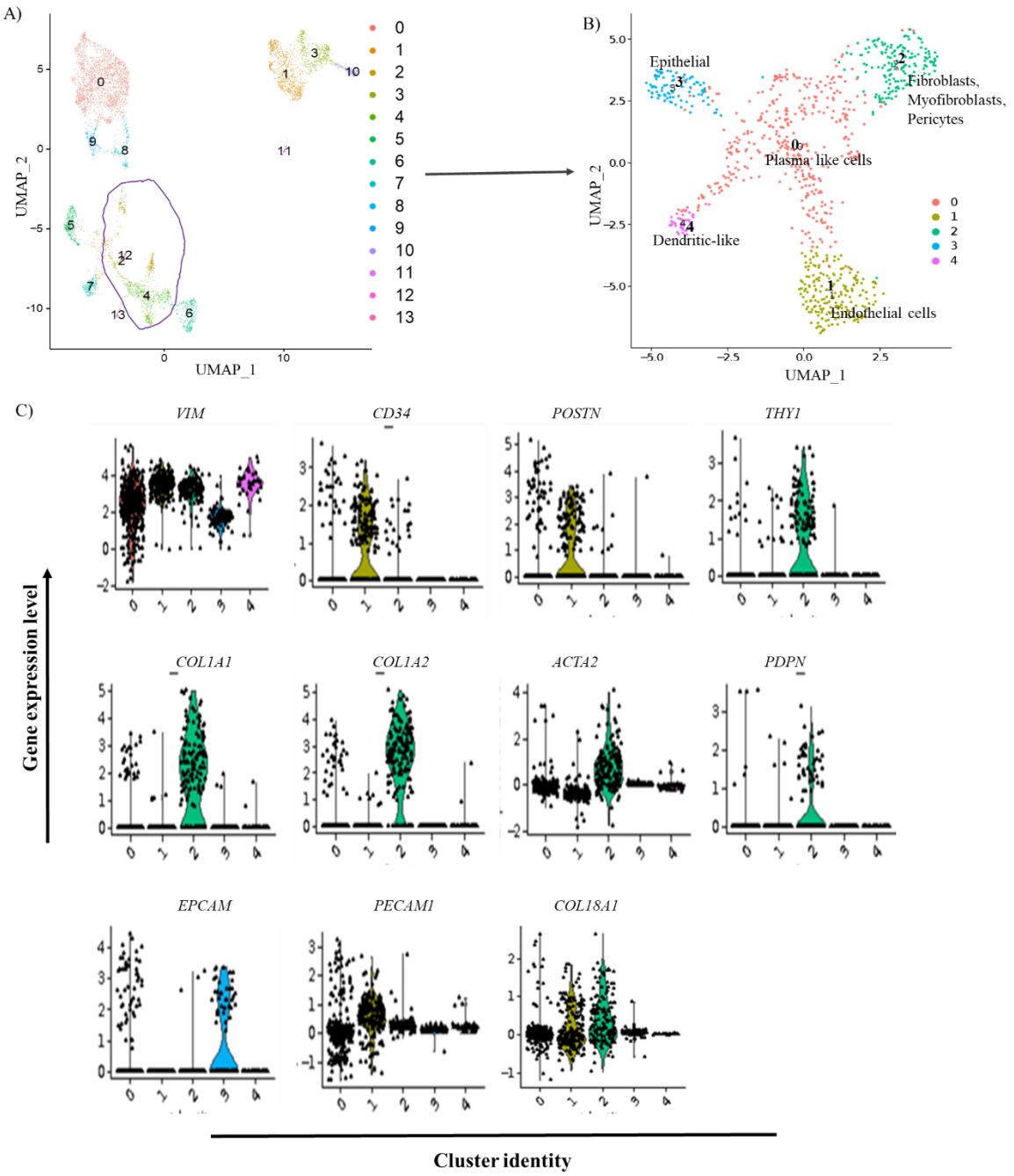


Figure 3.16: Sub clustering of cells based on gene expression to identify stromal cells with a low transcription profile. A) UMAP plot showing 14 identified clusters based on gene expression patterns, clusters 2, 4 and 6 circled in purple were selected for sub clustering, each dot represents a single cell and cells clustered together with the same colour represent a cell population. **B)** Sub-clustering identified 5 cell populations, these were identified as; **cluster 0** as Plasma cells; **cluster 1** as endothelial cells, **cluster 2** as fibroblasts, myofibroblasts and pericytes **cluster 3** as epithelial cells and **cluster 4** as dendritic cells **C)** Violin plots show pattern of select stromal markers, each dot represent a single cell plotted as per level of expression. The y-axis shows gene expression profile level with zero being the lowest expression which increases towards the top of the axis. The x-axis shows the identity of the cluster ranging from 0-4

3.3.4 Fibroblasts and biomarkers of fibrosis

Myofibroblasts are the activated phenotype of fibroblasts and play a major role in intestinal fibrosis. Myofibroblasts exhibit profibrotic features such as increased production of extracellular matrix (ECM) and contractile proteins that result in tissue stiffness due to fibrosis (336,484,485). The mediators and pathways through which fibroblasts regulate intestinal fibrosis have been studied but are not fully understood. The bulk RNA-seq data from this study revealed differences in expression of markers that are currently being studied in other fibrotic diseases.

A particular interest was on *GREM1*, *HDAC1* and *FGF2* which were shown as differentially expressed in the proximal, stricture and distal segments. Expression of these markers was further investigated in the scRNA-seq data with interest to localise their expression to a specific cellular phenotype. *HDAC1* was broadly expressed in several clusters including immune and stromal populations as seen in figure 3.17D, whilst *FGF2* and *GREM1* localised predominantly to the stromal clusters specifically cluster 2 which was identified as stromal cells (figure 3.17 B and C). In the next sections, the expression profile of *HDAC1*, *FGF2* and *GREM1* within the stromal cell clusters was explored.

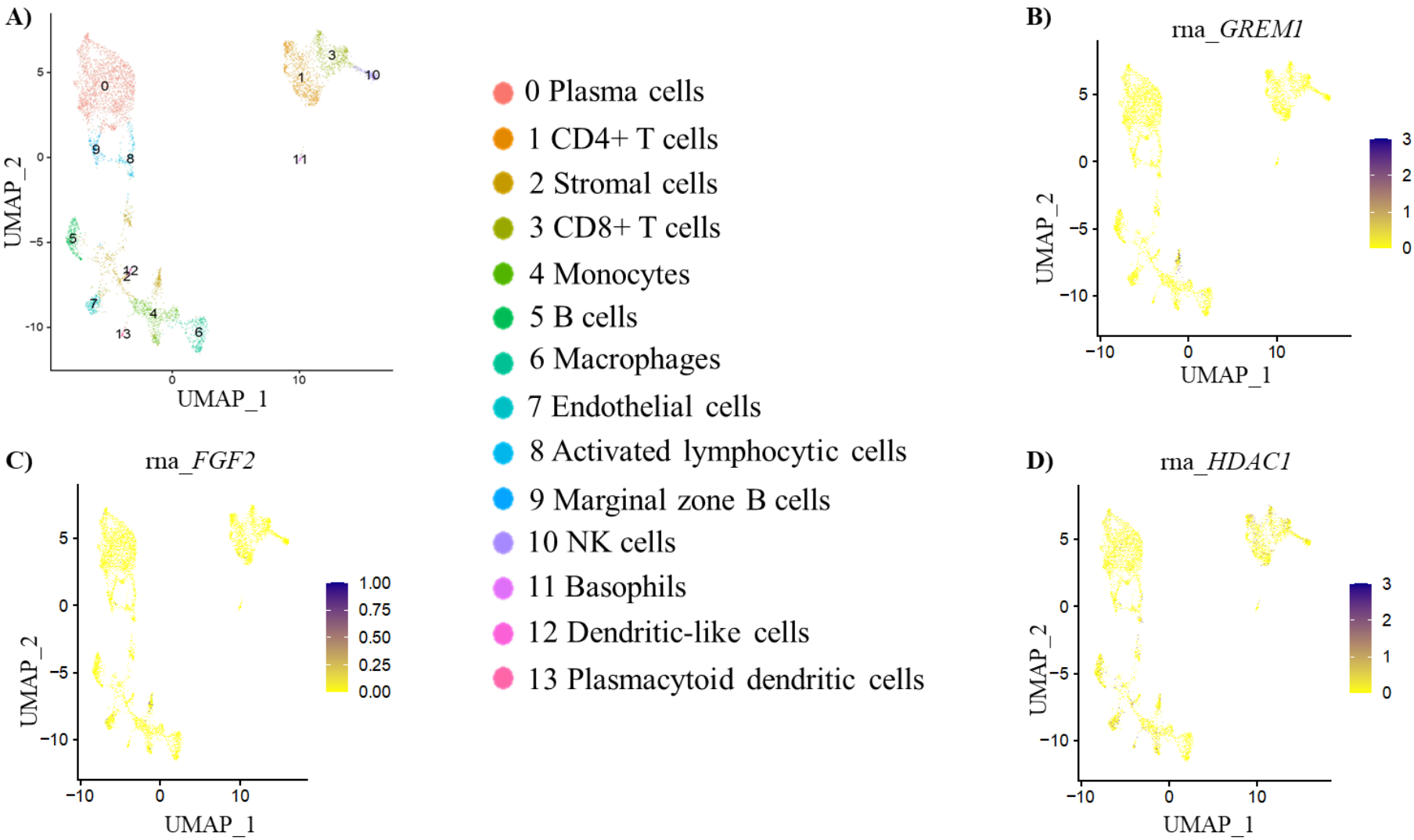

A)
UMAP_2
UMAP_1
0
1
2
3
4
5
6
7
8
9
10
11
12
13
0 Plasma cells
1 CD4+ T cells
2 Stromal cells
3 CD8+ T cells
4 Monocytes
5 B cells
6 Macrophages
7 Endothelial cells
8 Activated lymphocytic cells
9 Marginal zone B cells
10 NK cells
11 Basophils
12 Dendritic-like cells
13 Plasmacytoid dendritic cells
B)
rna_GREM1
C)
rna_FGF2
1.00
0.75
0.50
0.25
0.00
D)
rna_HDAC1

Figure 3.17: Expression of target genes in scRNA-seq data from proximal and stricture segments. 10X single cell sequencing was performed on total cell populations isolated from the stricture and proximal regions of the resected bowels of three patients. Single cells from all patients and locations were pooled and clustered using the UMAP_1 vs UMAP_2 parameters. **A)** Clusters were identified based on their expression of lineage defining marker genes **B)** UMAP_1 vs UMAP_2 expression plots for *GREM1* **C)** *FGF2* and **D)** *HDAC1*

3.3.4.1 *GREM1* mRNA expression localised to stromal cells in resected intestinal tissue

Genetic studies have suggested a central role of stromal cells in mediating fibrosis in multiple organs including the intestine (486). Therefore, a more focused analysis on expression of select biomarkers including *FGF2*, *GREM1* and *HDAC1* was performed on the five clusters that were identified during sub-clustering. The bulk RNA seq data showed upregulation of *GREM1* in the stricture segment as compared to the proximal and distal segments. Data from the scRNA-seq showed *GREM1* mRNA expression to be localised to the stromal cell cluster containing a mix of fibroblasts, myofibroblasts and pericytes as seen in figure 3.18B. However, it also shows that not all stromal cells within cluster 2 express *GREM1*. This may suggest that of the mixture of stromal cells in cluster 2, *GREM1* expression might be specific to one cell type within that group. A further analysis of *GREM1* mRNA expression between proximal and stricture segments showed a trend in increased expression of *GREM1* in the stricture as compared to the proximal as seen in figures figure 3.18 C, D and E. This however is not conclusive because the data was not statistically significant.

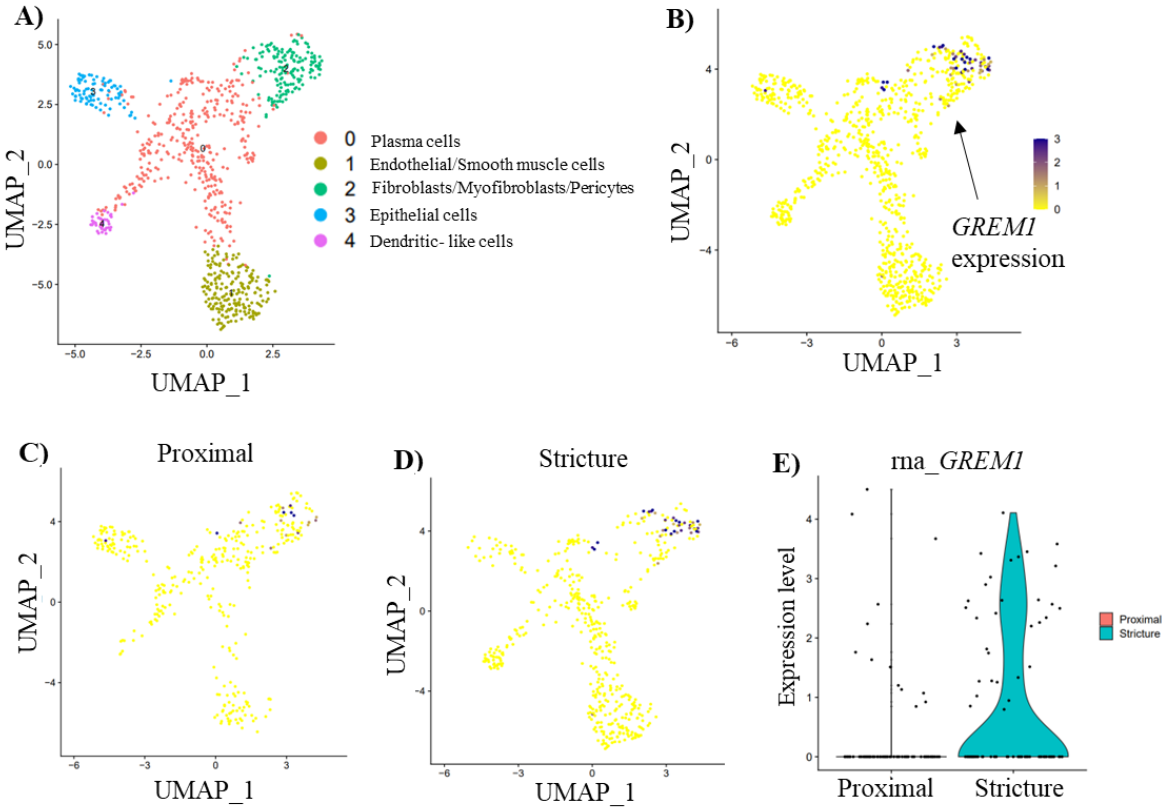


Figure 3.18: *GREM1* expression identified in intestinal stromal cells (cluster 2) A) Representative UMAP plot showing sub-clustering to identify stromal cell populations **B)** UMAP plot showing expression of *GREM1*, the arrow points to cells expressing the gene, the scale corresponds to level of gene expression, purple shows the highest expression while yellow shows no expression. **C)** UMAP plot showing expression of *GREM1* in cells from the proximal segment **D)** UMAP plot showing expression of *GREM1* in cells from the stricture segment **E)** Violin plot showing *GREM1* expression level in cells from the stricture and proximal.

3.3.4.2 *HDAC1* ubiquitously expressed in resected intestinal tissue

The second biomarker of interest was *HDAC1*, bulk RNA-seq data showed that *HDAC1* was differentially expressed across proximal, stricture and distal segments with the highest expression seen in the distal tissue (figure 3.19E). The analysis performed on scRNA-seq data showed ubiquitous expression of *HDAC1* in all the 5 sub-clusters as seen in figure 3.19B. Figure 3.19 shows a trend in increased expression of *HDAC1* within cluster 1 which was identified as constituting a combination of endothelial and smooth muscle cells. However, this is inconclusive as statistical analysis was not performed due to the very low cell numbers that

were used for analysis. Comparison in *HDAC1* expression between the proximal and stricture segments showed expression of *HDAC1* in both segments which confirms the ubiquitous nature and expression profile of *HDAC1*.

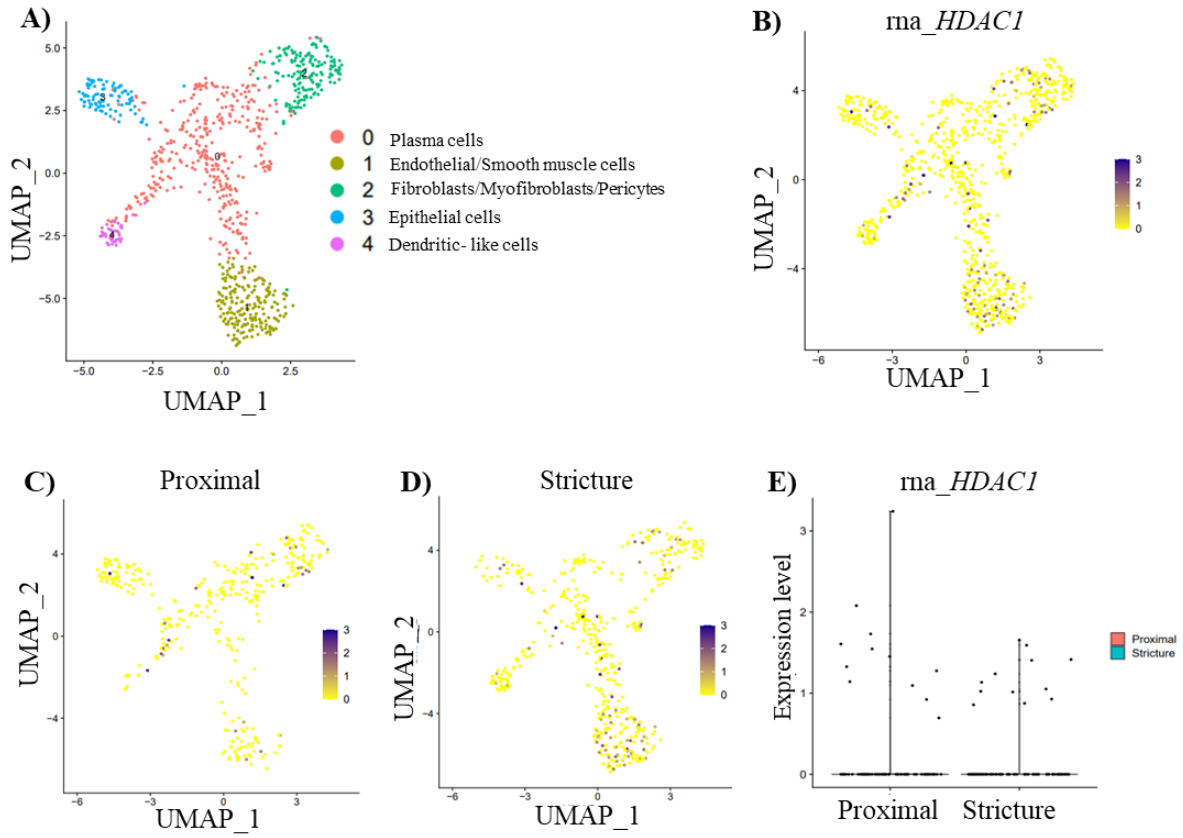


Figure 3.19: *HDAC1* expression seen across sub-clustered cell populations in both the proximal and stricture segments A) Representative UMAP plot showing sub-clustering to identify stromal cell populations **B)** UMAP plot showing expression of *HDAC1* across the cell populations identified through sub-clustering **C)** UMAP plot showing expression of *HDAC1* mRNA expression in cells from the proximal segment **D)** UMAP plot showing expression of *HDAC1* in cells from the stricture segment **E)** Violin plot showing *HDAC1* expression level in cells from the stricture and proximal tissue.

3.3.4.3 FGF2 mRNA expressed in stromal cells of resected intestinal tissue

To further assess the mRNA expression profile of *FGF2* in the proximal and stricture segments of resected gut tissue, *FGF2* gene expression was measured in cell populations identified upon sub-clustering of the scRNA-seq data. The data showed localised gene expression to cluster 1

identified as endothelial/smooth muscle cells and cluster 2 identified as a mix of fibroblasts, myofibroblasts and pericytes respectively as seen in figure 3.20B. A further comparison between the proximal and stricture segments in graphs C and D in figure 3.20 showed an interesting expression pattern in which FGF2 expression was restricted more to cluster 2 whereas in the stricture, gene expression is seen in both cluster 1 and 2 identified as endothelial/smooth muscle cells and fibroblasts, myofibroblasts and pericytes respectively.

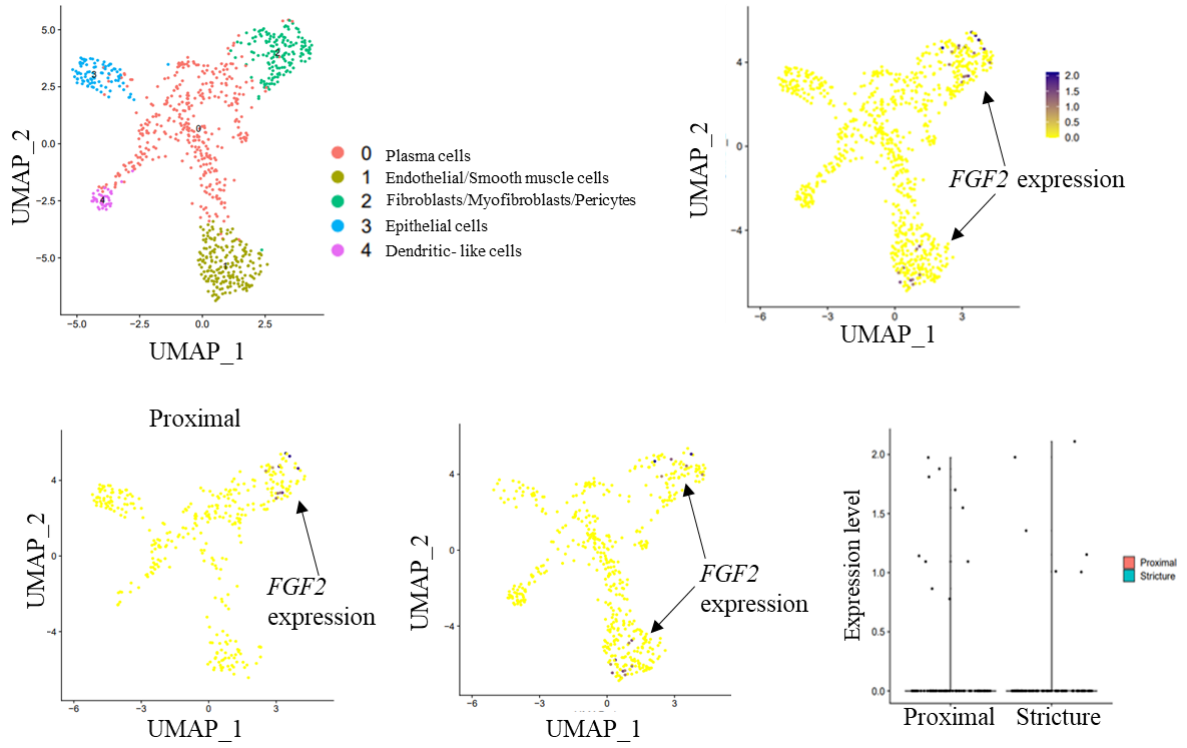


Figure 3.20: *FGF2* expression localised to clusters identified as endothelial/smooth muscle cell cluster and fibroblasts A) Representative UMAP plot showing sub-clustering to identify stromal cell populations **B)** UMAP plot showing expression of *FGF2* across the cell populations identified through sub-clustering, gene expression localised to endothelial, smooth muscle and fibroblast clusters **C)** UMAP plot showing expression of *FGF2* mRNA expression in cells from the proximal segment, gene expression localised to the stromal cell cluster **D)** UMAP plot showing expression of *FGF2* in cells from the stricture segment, gene expression in clusters identified as endothelial/smooth muscle cells and the fibroblast cluster **E)** Violin plot showing *FGF2* expression level in cells from the stricture and proximal segments.

3.4 Discussion

Studies to understand transcriptional changes that occur in normal versus diseased conditions have provided insight into variation in gene expression patterns across disease states. The data acquired has informed decisions on potential therapeutic targets that have been translated to life changing therapies (487). Bulk RNA-seq platforms have been used to study average gene expression profiles in each given sample, and the method is robust, sensitive and provides information on variation in disease versus normal conditions. Despite the vast amount of

informative data that has been generated using bulk-RNA-seq, the limiting factor with the technique is that it is not possible to analyse cell specific gene expression profile. ScRNA-seq platform is the more recent technique that was introduced by Tang (2009), and provides the means to identify individual cell specific gene expression profiles. This approach has been used to identify rare cellular phenotypes, identify unique molecular signature in a specific disease state and link to a particular cell phenotype. scRNA-seq data has also been used to monitor cell fate and plasticity in response to disease (455).

As discussed in the previous chapters, there is still an unmet need in understanding mechanisms that drive fibrostenosing disease, with approximately 50% of patients with CD still developing fibrosis (488). This has been attributed to the heterogenous nature of the disease. This study aimed to complement multiple studies that have already been carried out in identification of disease-classifying molecular and cell type profiles associated with fibrostenosing CD (279). The uniqueness to this study was the access to matched tissue from the fibrotic region (stricture) and segments that were proximal and distal to the stricture from the same patient. The study cohort constituted of 25 samples in total, and it would have been ideal to perform bulk RNA-seq on all the 25 samples for a higher statistical power. However, only 9 samples out of the 25 were sequenced due to cost limitations while additional samples served as a validation tool for follow up RT-qPCR analysis.

Bulk RNA-seq data provided an overview of the transcriptome profile of the stricture as compared to the proximal and distal segments. There was a clear alteration in gene expression signature in the stricture as compared to the proximal and distal (figure 3.3B). The data showed a strong genetic profile that was linked to pathways relevant in fibrosis as seen in hierarchical clustering in figure 3.6. Some of the genes upregulated in the stricture segment such as *CTSK*, *COL1A1*, *GREM1*, *PECAM1* and *SERPINE1* have been identified as those having a key role

in ECM remodelling, cell differentiation and tissue remodelling. Upregulation of profibrotic genes such as CCL2 was also seen in the stricture in comparison to the non-stricture segments; CCL2 is important in the recruitment of infiltrating CCR2+ monocytes and fibrocytes. Both cell types signal to produce TIMP-1 which inhibits degradation of collagen. CCL2 is also important in recruitment of fibroblasts to the site of injury hence driving the inflammatory and fibrotic phenotype (390).

The findings from the bulk RNA-seq data were interesting however not specific to the cell type in which gene expression alterations occurred. The study was complemented using single cell sequencing approach comparing tissue from the proximal (n=3) and stricture (n=3) segments. The single cells were pooled from all 3 samples and segments, quality control was performed and eventually the resulting cells used for analysis were a total of 6,106 cells. Analysis resulted in identification of 14 clusters in total constituting of immune and stromal cells as seen in figure 3.12. A further analysis showed that the cells were distributed in a similar pattern between the proximal and distal; and across the 3 samples. This is indicative of how heterogenous the intestinal microenvironment is. It is important to note that the proximal tissue in this study is not classed as “healthy” as the segment is also collected from CD patient. Inflammation in CD is predominantly driven by an activated immune cell phenotype, however recent studies have shown the importance of stromal cells in driving the inflammatory phenotype and eventually fibrosis.

In this study, there was an interest in investigating the stromal cell profile in stricture and non-stricture segments; and the transcriptional changes that may be driving the inflammatory and fibrotic phenotype. The challenge in this case was the low frequency of stromal cells in the tissue which was compounded by the low cell numbers that passed QC and were suitable for analysis. Within the initial 14 clusters that were identified during the initial clustering, a further

sub-clustering analysis was performed on clusters 2, 4 and 12 that showed a stromal cell signature. Sub-clustering resulted in identification of stromal cell populations such as endothelial and epithelial cells, fibroblasts and a smooth muscle phenotype. Of interest were targets such as *FGF2*, *HDAC1* and *GREM1* that were identified as differentially expressed in the proximal, stricture and distal segments in the bulk RNA-seq. In addition, was their biological importance in regulating pathways linked to inflammation and fibrosis.

Epigenetic modulation by enzymes such as HDAC1 has been shown as an important process in fibrosis across various fibrotic conditions including cardiac and pulmonary fibrosis; emerging data suggests that inhibiting the function of HDAC enzymes alleviates fibrosis and increases tissue remodelling hence our interest in HDAC1 gene expression profile. The one challenge with HDAC1 as a target is because it is ubiquitously expressed by multiple cells, it becomes very challenging to target as this could result in systemic toxicity and off target effects. FGF2 gene expression on the other hand was more restricted in expression to endothelial /smooth muscle cells and fibroblasts. There is data that suggests that FGF2 signalling inhibits FGF2-mediated fibroblast activation and fibrosis hence the interest in FGF2 as a potential therapeutic target.

The scRNA-seq data showed an interesting expression profile for *GREM1* that was localised to the fibroblast cluster; inhibition of *GREM1* signalling has been shown to reduce intestinal fibrosis through inhibition of fatty acid oxidation in fibroblasts (489). Most importantly, there is evidence of cross talk between the TGF-β and BMP signalling pathways, both pathways canonically signal through the same transducers proteins (SMADs). Inhibition of gremlin 1, an antagonist of the BMP pathway may cause alteration in downstream signalling of the TGF-β pathway. It was important to further understand the role of gremlin 1 in progression of intestinal fibrosis and the effects of inhibiting gremlin 1 to TGF-β/BMP signalling. In the next chapter,

protein expression studies will be performed using multispectral imaging technology to understand the location of fibroblasts in the proximal, stricture and distal segments and fibroblast expression pattern of biomarkers of interest including gremlin 1, FGF2 and HDAC1.

CHAPTER 4

IDENTIFICATION OF FIBROBLAST MARKER EXPRESSION PATTERNS AND LOCALISATION OF FGF2, HDAC1 AND GREMLIN 1 IN THE PROXIMAL, STRICTURE AND DISTAL SEGMENTS OF RESECTED GUT IN CROHNS DISEASE

Overview of this chapter

This chapter provides an insight into the localisation of intestinal fibroblasts and expression pattern of target biomarkers of interest, these include gremlin 1, HDAC 1 and FGF2 in the proximal, stricture and distal segments of patients with Crohn's disease. Haematoxylin and Eosin staining was performed to determine tissue morphology. Immunohistochemistry (IHC) was used to determine localisation and distribution of markers of interest relative to expression of stromal cell markers.

4.1 Introduction

Progress has been made in the development of anti-inflammatory therapies (490), despite the success of these therapies in mitigating inflammation, they have not been effective in restoring the detrimental effects on tissue function and architecture caused by chronic inflammation resulting in tissue fibrosis. This remains a major clinical unmet need in organ fibrosis in general including fibrostenosing CD, which is a major complication of CD (491). Data has shown that resolution of inflammation does not always prevent disease progression to fibrosis, this is because the fibrotic process is modulated by multiple cell types and various signalling pathways which complicate the process of identification of potential therapeutic targets. It may also be a result of inflammation independent mechanisms driven by interactions between activated stromal cells such as myofibroblasts and extracellular matrix proteins (ECM) (450,492). Suggested inflammation independent molecular mechanisms include ECM-cell interaction, creeping fat (CrF), gut microbiota and metabolic reprogramming (493).

Bulk RNA-seq data presented in the previous chapter gave an insight into the transcriptome profile of the stricture, proximal and distal segments. The data showed a profibrotic gene expression profile in the stricture associated with an activated stromal cell phenotype which was less prominent in the proximal and distal segments. The upregulated genes associated with fibroblasts, myofibroblasts and pericytes included *COL1A1*, *COL1A2*, *VIM*, *PDGFRA* and *PODN* among others; Activated fibroblasts (myofibroblasts) have been reported as the main effector cells of the intestinal fibrotic microenvironment with profibrotic features including excess production of ECM proteins. Furthermore, the single cell RNA sequencing (scRNA-seq) data showed the presence of a stromal cell phenotype with an upregulated expression of *THY1*, *VIM*, *PDPN* and *ACTA2*; well documented markers known to be expressed on stromal cells (363,494). In addition, studies have shown stromal cells as key cellular drivers of organ fibrosis, particularly myofibroblasts as the major effector cells and producers of ECM proteins

(495)(1,97, this presented a justification to further investigate the expression patterns of myofibroblasts in the stricture, proximal and distal segments of resected intestinal tissue. Molecular profiling data obtained through sequencing provide an in depth understanding of the transcriptome profile. However, these techniques require tissue dissociation which results in distortion of tissue morphology and loss of spatial information that maybe associated to a specific disease phenotype. On that basis, stromal cell localisation and expression patterns of biomarkers of interest including gremlin1, HDAC1 and FGF2 were investigated in the proximal, stricture and distal segments using imaging techniques that preserved morphology of the tissue.

Figure 4.1A shows the layers of the intestine including the mucosa which is the inner most layer of the intestine consisting of a thin layer of epithelium and connective tissue made up of the lamina propria, villi and crypts. The submucosa surrounds the mucosa; it is highly vascularised and has high frequencies of immune and stromal cells for example, fibroblasts and myofibroblasts. Intestinal fibroblasts are located in the apex of the villi, fibroblasts and myofibroblasts as seen in figure 4.1B. The muscularis propria consists of an inner and outer layer of smooth muscle that directs the coordinated wave-like propulsive movement referred to as peristalsis of the intestine. The serosa is the outer most layer of the intestinal wall that contains a significant amount of collagen fibres and plays a key protective role.

Intestinal stromal cells are characterised by expression of phenotypic markers such as Thy1 (CD90), PDGFR α and β and vimentin (496), they have a negative expression of CD45 (immune cell marker), CD31(endothelial cell marker) and EpCAM (epithelial cell marker). Myofibroblasts which are the effector cells of fibrosis are located along the villi and crypt just beneath the sub epithelial layer, they are distinguished from other stromal cells such as fibroblasts and pericytes through their expression of alpha smooth muscle actin (α-SMA).

Our bulk RNA-seq data revealed differential gene expression of biomarkers implicated in fibrosis; gremlin 1 and FGF2 were upregulated at mRNA level in the stricture while expression of HDAC1 was downregulated in the stricture as compared to the proximal and distal segments. Studies in multiple organs such as the lung, kidney, liver and heart have shown the involvement of each of these biomarkers in the modulation of profibrotic mechanisms, prompting the interest to investigate the role of these biomarkers in the gut, how they drive the fibrotic phenotype in fibrostenosing CD and their potential as targets for anti-fibrotic therapy (465,497).

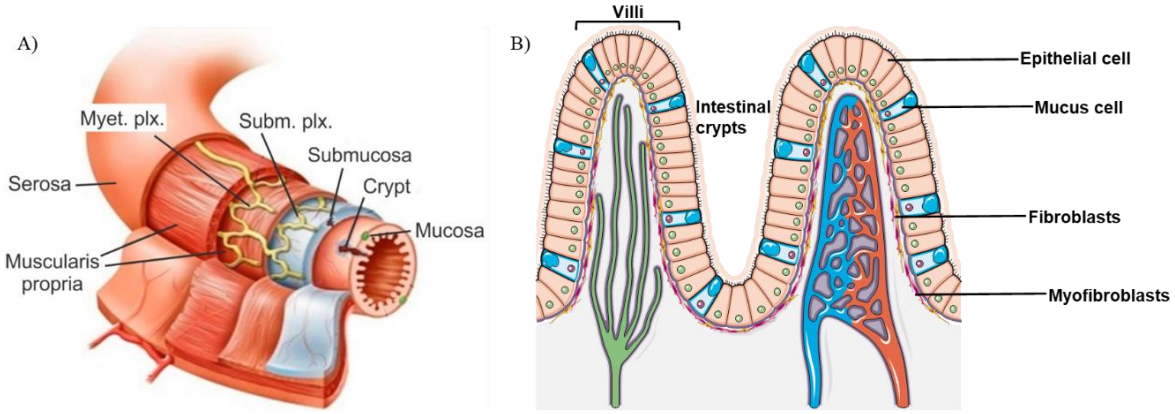


Figure 4.1: Schematic diagram showing layers of the GIT. A) The layered structure of the intestine with the mucosa as the inner most layer, submucosa is the next layer that surrounds the mucosa, muscularis propria surrounds the submucosa and made up of muscle tissue, the outer protective layer is the serosa. . Subm. plx. = submucosal plexus; Myet. plx. = myenteric plexus. **B)** Cross section of the mucosa containing the lamina propria, with villi and the gut epithelium. Stromal cells such as fibroblasts, intestinal subepithelial myofibroblasts and interstitial Cajal cells are located in this region. It is highly vascularized. *Images are adapted from Siri et al., 2020* and *https://smart.servier.com/.*

To further understand the function of gremlin1, FGF2 and HDAC1, protein expression studies were carried out relative to the spatial distribution of stromal cells in the proximal, stricture and distal segments. Understanding protein expression patterns of these biomarkers relative to the location of stromal cells may give an insight on their modulatory mechanisms of the

intestinal fibrotic microenvironment, and determine fibroblast cell fate in terms of activation, proliferation and differentiation into a more profibrotic phenotype likely responsible for the formation of strictures in CD.

FGF2 has been reported as a potent mitogen that plays key roles in tissue repair, it regulates cellular differentiation and migration into the stromal microenvironment. Its cellular targets include endothelial cells during angiogenesis, a process that is important in formation of tissue vasculature (498). TGF-β has been shown to induce stromal cells such as fibroblasts to express growth factors including FGF-2 (499,500). In the intestine, FGF2 is expressed intracellularly predominantly by stromal cells such as fibroblasts and myofibroblasts that are located in the intestinal submucosa, FGF2 has been shown to play an important role in cell proliferation, differentiation and repair of the epithelial barrier upon tissue in jury (501).

HDAC enzymes play a role in regulation of gene transcription through DNA modification mechanisms in particular removal of acetyl groups from histone/DNA complexes hence silencing gene transcription. HDAC1 has been implicated in the pathogenesis of multiple fibrotic diseases including kidney, lung and heart (502,503). An in vitro study conducted in kidney fibroblasts in which cells were treated with an HDAC1 inhibitor showed suppression in fibroblast proliferation and activation, expression of α-SMA and fibronectin was significantly reduced in treated cells as compared to control cells (504). Another in vitro study in lung derived fibroblasts showed inhibition of TGF-β induced fibroblast to myofibroblast transition (505). These multiple studies have presented HDAC1 as an attractive antifibrotic therapeutic target. HDAC1 is expressed intracellularly on multiple cell types but predominantly on intestinal epithelial cells, it has been reported to regulate homeostasis and intestinal epithelial barrier (506).

The third target of interest is gremlin 1, an antagonist of the BMP signalling pathway; studies in organ fibrosis such as kidney, heart and lung have shown increased expression of gremlin 1 which presents an unfavourable outcome in fibrotic diseases (489,507). A study by Duffy (2021) showed that increased expression of gremlin 1 was linked to increased fibroblast to myofibroblast transition, increased TGF-β expression and α-SMA levels. The modulatory mechanisms are not clearly understood; however, the literature suggests cross talks between TGF-β/BMP signalling axis involving the SMAD signalling proteins downstream both signalling pathways (509). Gremlin1 protein expression is not tissue specific however fibroblasts have been reported as the major expressers of gremlin 1, this is supported by the scRNA-seq data generated from this study shown in chapter 3 above. Cellular localisation appears cytoplasmic, it is also secreted into the cytosolic space and tends to attach to fibres within the ECM (510).

For this section of the study, FFPE tissue sections from the proximal, stricture and distal of resected gut were used to perform H and E staining to understand the tissue structure and architecture. IHC was used to evaluate the presence or absence of the protein and further to study protein expression patterns in the tissue. Multispectral immunofluorescence was used to study co-expression of HDCA1 and FGF2 with fibroblast markers.

4.2 Aims of this chapter

The specific aims of this chapter were to;

1. To identify the presence of fibroblasts in the proximal, stricture and distal segments by immunohistochemistry.
2. To identify the presence and spatial distribution of gremlin 1, HDAC1 and FGF2 in the proximal, stricture and distal segments.

4.3 Results

4.3.1 Structure and architecture of gut tissue

The structure of normal gut tissue is arranged in distinct layers which include from the inside the lumen, mucosa, submucosa, muscularis and serosa. The mucosa contains the epithelium and lamina propria with crypts and villi located in this layer. The next layer is the submucosa which is mainly made of connective tissue and fibres that support the structure of the tissue, the muscularis is the next layer made up of dense muscle layers followed by the outer layer which is the serosa that plays a protective role (511). The mucosa layer contains different sections, the crypts in normal gut tissue are arranged in well-defined tube-like structures parallel to each other with villi protrusions into the lumen, indicative of an undisturbed epithelial gut barrier. Cells that line the epithelium around the mucosal layer (crypts and villi) can be seen, these include goblet, Paneth, enterocytes and epithelial cells. These cells are responsible in maintaining gut homeostasis. Just beneath the epithelial layer in the sub-epithelium are resident fibroblasts, in an activated form, they proliferate and can be found even in the deeper layers of the gut (512).

4.3.2 Histology of gut resected tissue

Human resected gut tissue was obtained from patients undergoing resection surgery due to Crohn's disease (CD) strictures. Tissue was collected from the fibrotic region (stricture), a matched segment was collected from a region distal to the stricture (colon) and another from a region proximal to the stricture (ileum). The segments obtained from each patient were referred to as proximal, stricture and distal segments for this study respectively. In addition, resected gut obtained from "non-tumour" colon in patients undergoing colectomy surgery was used as a "non-CD control", as it exhibited characteristics of normal gut tissue as compared to CD tissue samples. The sections were stained by haematoxylin and eosin (H&E) to enable visualisation of tissue structure and architecture including cellular components such as the

cytoplasm and nucleus. Figure 4.2A shows non-CD control tissue in which the structure and architecture of the tissue was unaltered, the lamina propria of the control tissue contains a few immune cells. The intestinal layers were well defined to include the mucosa, submucosa and muscularis; figure 4.2 A shows a clear definition of the mucosa which is made up of the epithelium and the lamina propria which contains the crypts and villi. The next layers are the submucosa and muscularis which are made up of intestinal connective tissue.

Figure 4.2 B-D shows multiple tissue structure alterations that occur in CD. The H and E images were taken from the distal segment (figure 4.2.B), stricture (figure 4.2.C) and proximal segment (figure 4.2.D) from the same patient respectively. Across the three segments, there was distortion of the gut layers and severe damage to the epithelium. The black arrows in the diseased tissues point to irregular and missing crypts and villi which leads to the disrupted tissue structure seen in these tissue sections. Another feature that is typical of CD is infiltration of inflammatory cells which clustered together to form structures referred to as granulomas, this was seen across segments, a representative structure of the granuloma is in a red circle (figure 4.2.D). H and E staining provides insight into the cellular and tissue structure, however, the information derived is not sufficient to define cellular phenotype and, hence the need to use more advanced imagine techniques to characterise cell types in the tissue and protein expression patterns.

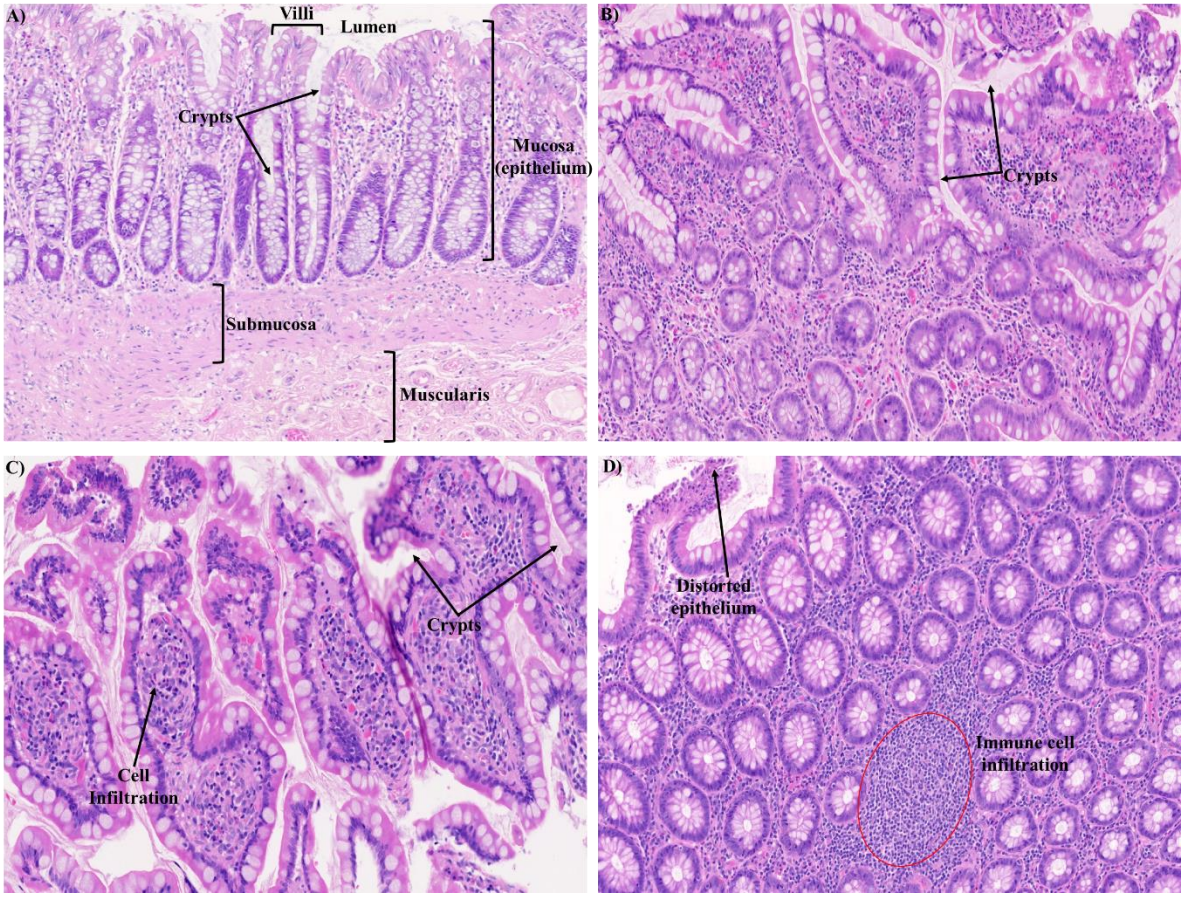


Figure 4.2: Histological appearance and structure of tissue in resected CD gut sections. Haematoxylin and eosin (H&E) staining was performed on FFPE sections. **A)** non tumour resection was used as "control" and shows the different layers of the intestine including the mucosa, submucosa and Muscularis, undistorted structure of the gut with clearly defined crypts. **B)** Resected tissue collected from the distal segment shows disrupted crypts, granuloma formation and immune cell infiltration **C)** Resected tissue collected from the stricture segment shows disrupted crypts and immune cell infiltration. **D)** Resected tissue collected from the proximal segment shows the presence of granulomas clustered and circled. Black arrow shows disruption to the epithelium. Imaging was performed on the Axio slide scanner (Zeiss); scale bar 100 μm

4.3.3 Fibroblasts identified in the proximal, stricture and distal segments of resected gut

Having identified a fibroblast population in the scRNA-seq data in both the proximal and stricture segments, the next step was to determine their presence and most importantly their location based on protein expression pattern. Identification of the fibroblast population in the scRNA-seq data was based on gene expression of well documented fibroblast markers such as

VIM, THY1, PDPN and ACTA2. Fibroblasts are positive for vimentin, Thy1 (CD90), PDGFRβ and podoplanin. α-SMA is expressed by activated fibroblasts (myofibroblasts), pericytes express all the mentioned are a summary of these common markers is shown in table 4.1.

Table 4.1: Common markers for identification of mesenchymal stromal cells

Markers	Fibroblasts	Myofibroblasts	Pericytes	Smooth muscle cells
Vimentin	+	+	+	+
CD90 (Thy -1)	+	+	+	-
Podoplanin	+	+	+	-
α-SMA	-	+	+	+
Desmin	-	-	+	+

4.3.4 Expression of vimentin in the proximal, stricture and distal segments

To get an insight into the location of intestinal stromal cells including fibroblasts, myofibroblast and smooth muscle cells in the proximal, stricture and distal segment, expression of a stromal cell marker was used for this purpose. Vimentin is a well-documented stromal cell marker; its expression can be used to locate stromal cells in the three segments. Therefore, formalin fixed paraffin embedded (FFPE) tissue sections were stained for vimentin by IHC (figure 4.3). IHC is a common technique used to detect expression of the protein of interest while ensuring that tissue structure, integrity and cellular characteristics are maintained.

Figure 4.3 shows the presence of vimentin+ cells (brown colour) in all the three segments, but their distribution was varied. Figure 4.3A shows vimentin+ cells distribution in the stricture, positive cells are distributed across the mucosa and submucosa, the brown staining extends to the deeper layers of the gut including the submucosa and muscularis. In comparison, positive vimentin staining in the proximal (figure 4.3B) and distal (figure 4.3C) is localised to the mucosal regions around the villi and crypts regions. Brown staining is still seen in the submucosa and muscularis but much less in comparison to the stricture. Further image analysis

for example using ImageJ to quantify positive staining was not performed for these tissues as only representative sections were stained. Figure 4.3D shows the isotype control staining that was performed on resection tissue obtained from the stricture segment, the isotype control did not show any brown staining which validated the positive vimentin signal in the images. Figure 4.3D was labelled to indicate the different intestinal layers including the mucosa, submucosa and muscularis.

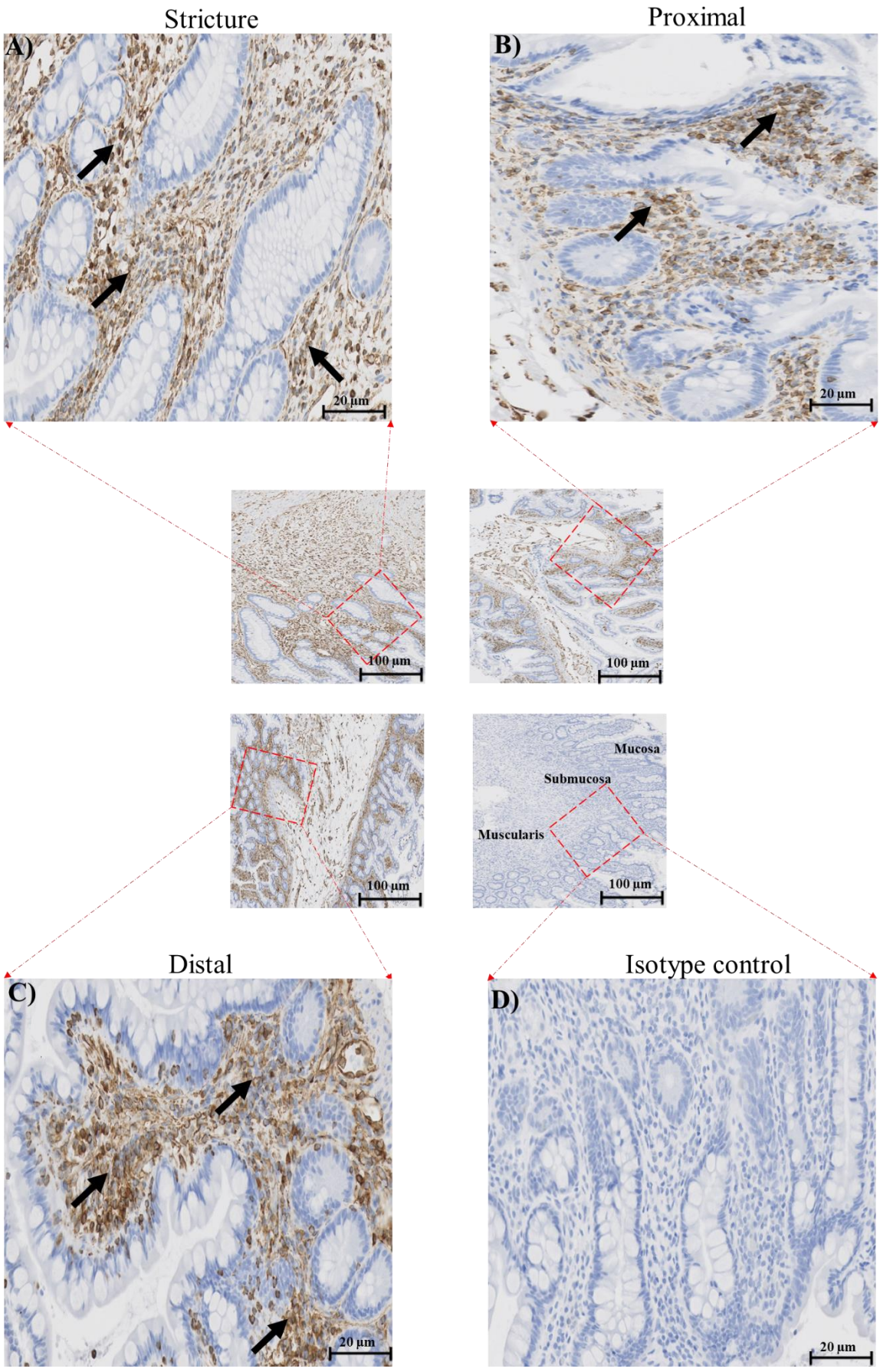

Stricture
Proximal
A)
B)
20 µm
20 µm
100 µm
100 µm
100 µm
Mucosa
Submucosa
Muscularis
100 µm
Distal
Isotype control
C)
D)
20 µm
20 µm

Figure 4.3: Vimentin^{+ve} cells identified in the proximal, stricture and distal segments. Representative images showing expression of vimentin in the three segments, black arrows are pointing to the brown positive staining **A)** Vimentin^{+} cells in the stricture **B)** Vimentin^{+} cells in the proximal **C)** Vimentin^{+} cells in the distal **D)** Representative image of the stricture segments stained with isotype control (mouse IgG1a), the image is labelled with the layers of the gut including the mucosa, submucosa and the muscularis. FFPE resected gut tissue sections cut at 5 μm were dewaxed and antigen retrieval performed at a pH of 9, peroxidase activity was blocked using BLOXALL reagent, sections were stained with unconjugated vimentin antibody at 0.78 μg/ml (Agilent, Dako) with the corresponding isotype control for 1 hour, visualisation of positive staining was performed using DAB while nuclei were stained with Mayers haematoxylin. Dehydration was performed and tissue was fixed with DPX for preservation. The arrows show vimentin positive staining as a brown colour while the cell nuclei appear blue. Imaging was performed using the Axio slide scanner (Zeiss) scale bars: 100 μm.

4.3.5 Expression of α-SMA in the proximal, stricture and distal segments

Expression of vimentin in section 4.3.1 showed the location of stromal cells within the sub mucosal layer. Fibroblasts, myofibroblasts, pericytes and smooth muscle cells express vimentin, alpha smooth muscle actin is used as a positive marker for pericytes, smooth muscle cells and activated fibroblasts (myofibroblasts) which have been reported to have increased expression of α-SMA (Shinde, Humeres and Frangogiannis, 2017). Sequential sections used for vimentin staining were stained for α-SMA to assess protein expression patterns. Figure 4.4 shows α-SMA+ cells detected in the proximal (figure 4.4A), stricture (figure 4.4B) and distal (figure 4.4C) segments. α-SMA protein expression pattern in the stricture was distributed in all the layers of the tissue including the mucosa, submucosa and the muscularis, this was similar to the staining pattern for vimentin in the stricture. On the contrary, although α-SMA+ staining was seen in all the layers of the proximal and distal segments, there are certain α-SMA+ regions marked in a red arrow in both the proximal and distal segments with brown staining that are not seen in corresponding regions for vimentin staining. This raised the question, whether all

α-SMA+ cells were myofibroblasts or were also smooth muscle cells especially in the muscularis layers. To address this, desmin staining was performed next.

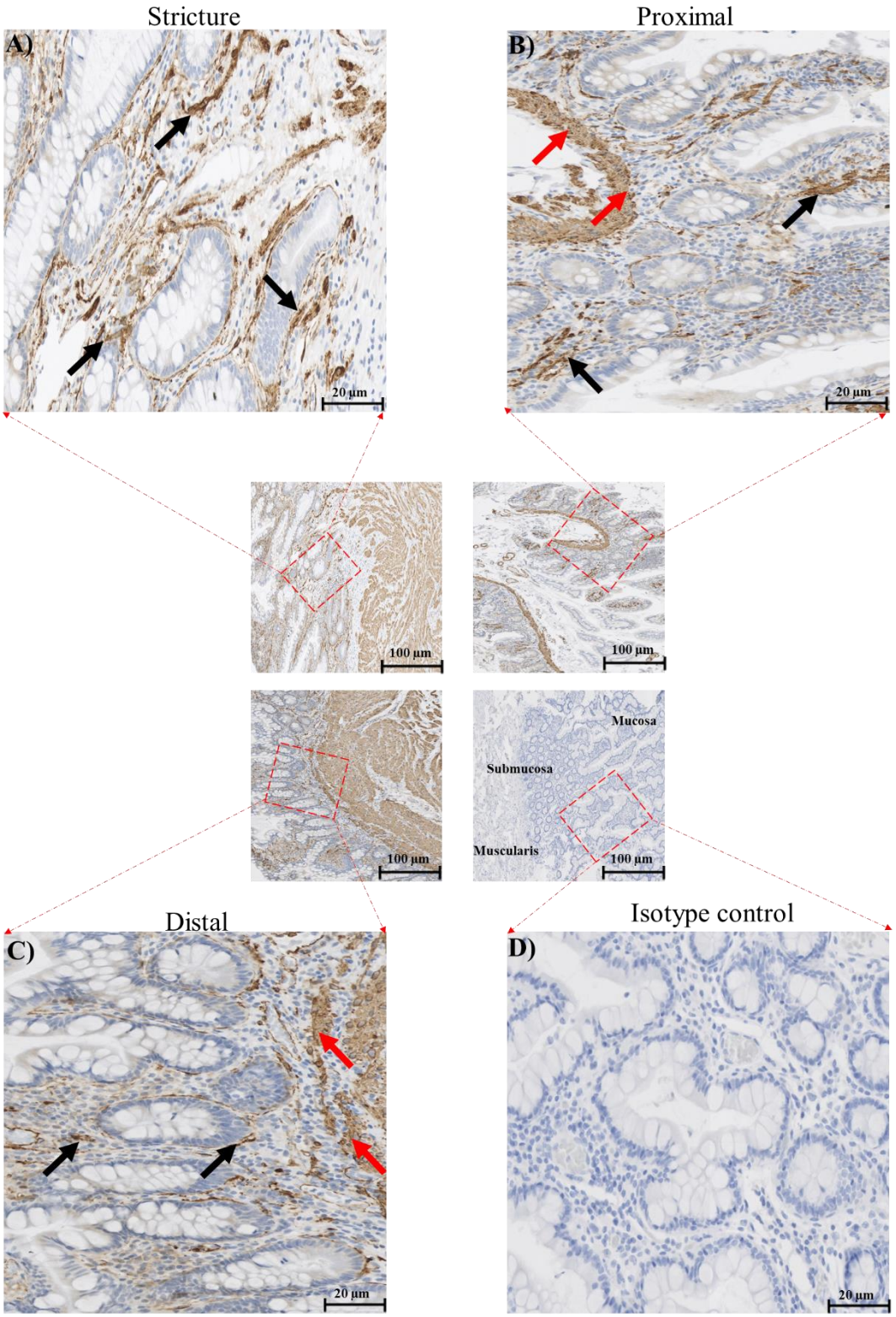


Figure 4.4: Expression of α-SMA spread across the deeper layers of the gut (submucosa and muscularis) in the stricture segment. Representative images showing expression of α-SMA in the **A)** stricture, **B)** proximal and **C)** distal segments. **D)** Representative image of stricture segment stained with isotype control (mouse IgG2a). FFPE resected gut tissue sections cut at 5 μm were subjected to antigen retrieval at a pH of 9 after dewaxing, the activity of peroxidases was blocked, sections were stained with unconjugated α-SMA antibody at 0.71 μg/ml (Agilent, Dako) and corresponding isotype control for 1 hour, visualisation of positive staining was performed using DAB while nuclei were stained with Mayers haematoxylin. Dehydration was performed and tissue was fixed to preserve in DPX. α-SMA expression appears as a brown colour as indicated with the black arrows, while the cell nuclei appear blue. The red arrows are positive brown staining regions seen for α-SMA staining which was negative for vimentin as seen in the previous image. Imaging was performed using the Axio slide scanner (Zeiss) scale bars; 100 μm.

4.3.6 Expression of desmin in the proximal, stricture and distal segments

Desmin is a muscle specific intermediate filament protein expressed by smooth muscle cells but not by fibroblasts and myofibroblasts. Therefore, it can be used to differentiate α-SMA positive myofibroblasts (desmin-negative) from smooth muscle cells expressing α-SMA (desmin-positive). Sequential tissue sections that were used for vimentin and α-SMA staining were also stained for desmin in matched tissues from the proximal, stricture and distal segments. Expression of desmin in figure 4.5 can be seen within the smooth muscle fibres particularly in the muscularis layers in all the three segments, highlighting presence of smooth muscle as indicated by the black arrows. Particularly, figure 4.5A shows a low expression of desmin within the stricture segment, there was no expression of desmin seen in the mucosa or submucosal layers however a low expression was seen in the muscularis as indicated by the black arrows. Figure 4.5B shows a localised expression of desmin within the submucosa in the proximal segment, brown staining in the mucosa is barely noticeable. On the contrary, there seems to be enhanced expression of desmin in the distal segment as compared to the stricture and proximal segments. Brown staining was distributed across the mucosa and submucosa as

indicated by the black arrows. Overall, there is no specific marker designated to identify fibroblasts and myofibroblasts; this requires a combination of various stromal cell markers. In that regards, images of the stricture segments were compared for expression of vimentin (figure 4.3), α-SMA (figure 4.4) and desmin (figure 4.5). Figure 4.6 shows enhanced expression of vimentin (figure 4.6A) and α-SMA (figure 4.6B) across all the segments in which most of these were desmin negative.

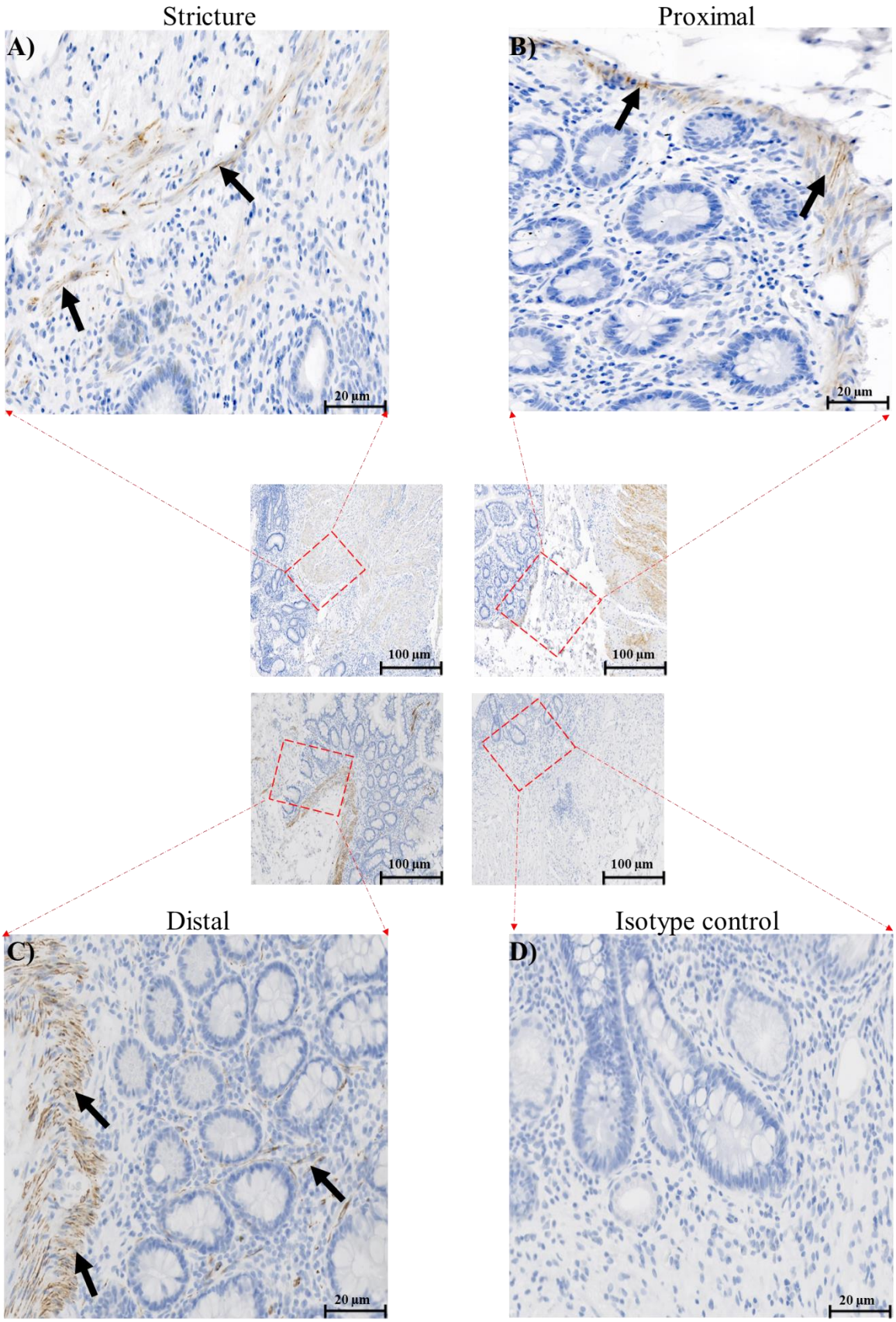

Stricture
Proximal
A)
B)
20 µm
20 µm
100 µm
100 µm
100 µm
100 µm
Distal
Isotype control
C)
D)
20 µm
20 µm

Figure 4.5: Expression of Desmin localised to the submucosa and muscularis layers in all the three segments. Representative images showing expression of Desmin in the **A)** stricture, **B)** proximal and **C)** distal segments. **D)** Representative image of stricture segment stained with isotype control (mouse IgG1a). FFPE resected gut tissue sections cut at 5 µm were subjected to antigen retrieval at a pH of 9 after dewaxing, the activity of peroxidases was blocked, sections were stained with unconjugated α-SMA antibody at 1.11 µg/ml (Agilent, Dako) and corresponding isotype control for 1 hour, visualisation of positive staining was performed using DAB while nuclei were stained with Mayers haematoxylin. Dehydration was performed and tissue was fixed to preserve in DPX. Desmin expression appears as a brown colour as indicated with the black arrows, while the cell nuclei appear blue. Imaging was performed using the Axio slide scanner (Zeiss) scale bars; 100 µm.

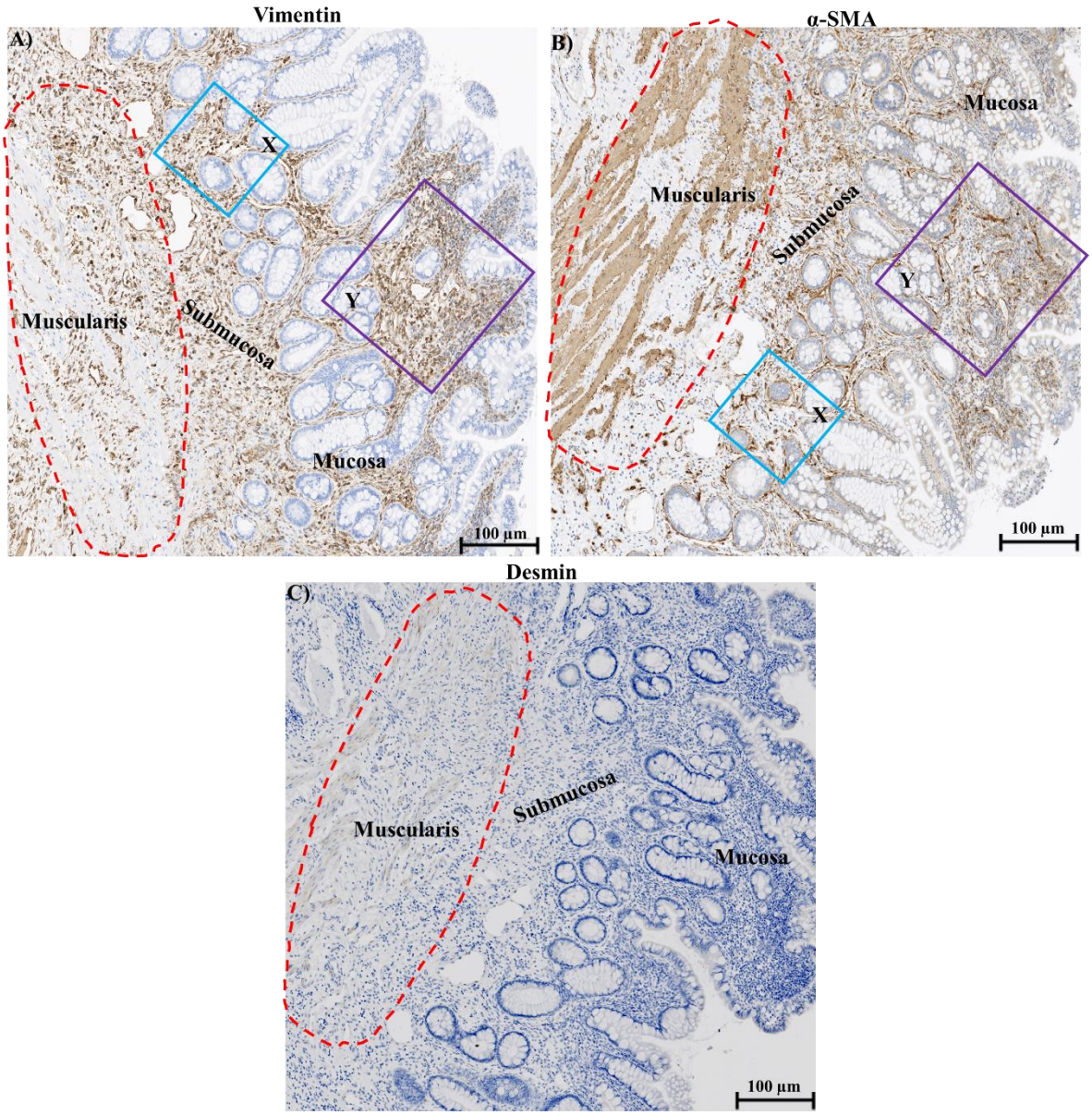


Figure 4.6: Protein expression of stromal cell markers in the stricture segment. Representative images from stricture segments showing comparison of stromal cell markers in the stricture, figure 4.6A) Vimentin, figure 4.6B) α-SMA and figure 4.6C) Desmin. The region highlighted in red displays the muscularis region, regions highlighted in blue (**X**) and purple (**Y**) is a direct comparison of expression of vimentin and α-SMA within matching regions. Protein expression appears as a brown colour while the cell nuclei appear blue. Imaging was performed using the Axio slide scanner (Zeiss) scale bars; 100 µm.

Fibroblasts and myofibroblasts are vimentin positive but desmin negative, therefore the regions positive for vimentin and negative for desmin show where the fibroblasts, myofibroblast and pericyte populations are localised. Myofibroblasts and pericytes are differentiated from fibroblasts through their expression of α-SMA, comparing protein expression of matched regions X and Y, it shows regions that are vimentin+ve in figure 4.6A but α-SMA-ve in figure 4.6B, this phenotype is characterised as fibroblasts. On the other hand, the regions also show cells that are positive for both markers which are characterised as myofibroblasts and pericytes. Characterisation of fibroblasts, myofibroblasts and pericytes requires the use of protein co-expression, the limitation to using IHC is that tissue can only be co-stained for a maximum of 2 markers hence the use of immunofluorescence as a better option.

Co-expression of phenotypic markers identified stromal cells in intestinal segments

To confirm the phenotype of stromal cells in intestinal segments, a combination of markers was used as there is no specific biomarker used for cellular profiling of fibroblasts, myofibroblasts or pericytes. IHC staining from the previous sections showed expression of vimentin and α-SMA that was seen in all intestinal layers including the mucosa, submucosa and muscularis. The single colour IHC staining provided an insight into the location of stromal cells, desmin was used as a biomarker to differentiate myofibroblasts and pericytes from smooth muscle cells. To further characterise fibroblast, myofibroblasts and pericytes, protein co-expression studies were ideal, this is a limitation when using IHC. In this section, an immunofluorescence (IF) panel containing CD90, α-SMA, podoplanin and a nucleus stain (DAPI) was designed to identify fibroblast populations in the proximal, stricture and distal segments.

These markers were chosen based on the scRNA-seq data previously shown in figure 3.16C in chapter 3 in which gene expression of *THY1, PDPN* and *ACTA2* were upregulated specific to the stromal cell cluster. Figure 3.16C showed that VIM was ubiquitously expressed by all stromal cells including epithelial and endothelial cells. Vimentin was excluded from the immunofluorescence panel due to its ubiquitous expression. A section stained with the isotype controls (labelled IC in figure 4.7) for the respective antibodies used in the IF panel was included.

Protein co-expression of CD90, podoplanin and α-SMA was performed in the proximal, stricture and distal segments, the images were taken from the regions around the intestinal crypts showing single and merged biomarker expression. As seen in figure 4.7, there is increased podoplanin expression in the distal as compared to the proximal and stricture segments. The green arrow in the merged image indicates cells that are co-expressing CD90, podoplanin and α-SMA. This phenotype is not clearly defined in the literature, α-SMA+CD90- cells were also identified in all the segments. On the other hand, expression of CD90 and α-SMA appears similar across all the three segments. As indicated by the red arrows, co-expression of CD90 and α-SMA cells can be seen, based on the literature, these double positive cells could be classified as a mixture of myofibroblasts or pericytes. CD90+ α-SMA- cells are also seen in all the three segments, these would be classified as fibroblasts (Hsia et al., 2016).

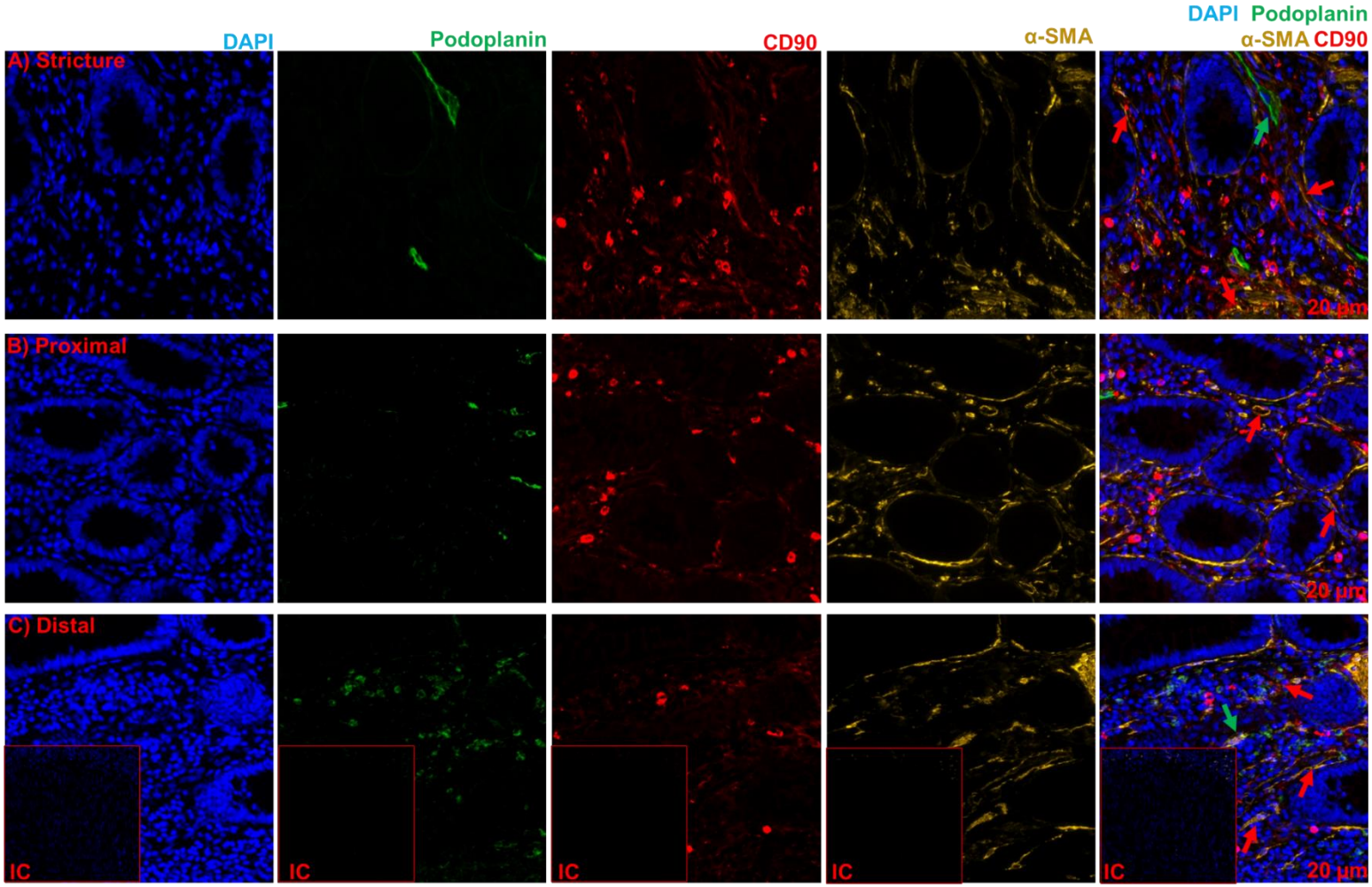

DAPI
Podoplanin
CD90
α-SMA
DAPI Podoplanin
α-SMA CD90
A) Stricture
B) Proximal
C) Distal
IC
IC
IC
IC
IC
20 µm
20 µm
20 µm

Figure 4.7: Co-expression of fibroblast markers using immunofluorescence (IF) staining in the proximal, stricture and distal segments. Antigen retrieval was performed on dewaxed FFPE sections in a basic antigen retrieval solution (pH9), unspecific binding was blocked by addition of casein, a mix of the unconjugated primary antibodies and the corresponding isotype controls was added onto the sections, antibodies included CD90 (rabbit), α-SMA (mouse IgG2a) and podoplanin (rat). Detection was performed by addition of secondary antibodies Alexa Fluor (AF) AF488 to detect podoplanin on the FITC channel, AF555 to detect CD90 on the Cy3 channel and AF647 to detect α-SMA on the Cy5 channel. Row 1; Images shown are sections stained with all three antibodies CD90, α-SMA, podoplanin and the nucleus stain. The images in each row are matched sections **A)** stricture **B)** proximal and **C)** distal. Isotype controls for each marker was included and labelled as IC. Imaging was performed on the Axio Zeiss confocal microscope (Zeiss, Germany); scale bar 20 μm.

4.3.7 Expression patterns of FGF2, HDAC1 and Gremlin1

4.3.7.1 FGF2 protein expression in resected gut of CD patients

FGF2 is a protein that is expressed by mesenchymal stromal cells located around the base of intestinal crypts and villi, it plays a key role in fibroblast proliferation and activation (Danopoulos et al., 2017b). To determine the expression of FGF2, IHC staining was performed on FFPE sections from the proximal, stricture and distal segments. An isotype control was included to establish a true positive signal. The black arrows in figure 4.8 show FGF2+ve staining which appears specific as seen in brown across all the 3 segments with very low background; stricture (figure 4.8A), proximal (figure 4.8B) and distal (figure 4.8C). The brown staining is localised to the lining of the sub epithelial layers of the villi within the intestinal mucosa, FGF2+ve staining also stretches out to the base of the crypts. FGF2+ve cells in this layer appear elongated like the shape of stromal cells. The staining pattern is similar to what has been shown in the previous section in which stromal cell markers were positive in similar

regions. In addition, expression of FGF2 appears to increase in the submucosal and muscularis layers where endothelial cells are predominantly located. Based on these images, it is not possible to identify major differences in FGF2 protein expression in the proximal, stricture and distal segments. However, the chromogenic staining provided an insight into the staining pattern of FGF2.

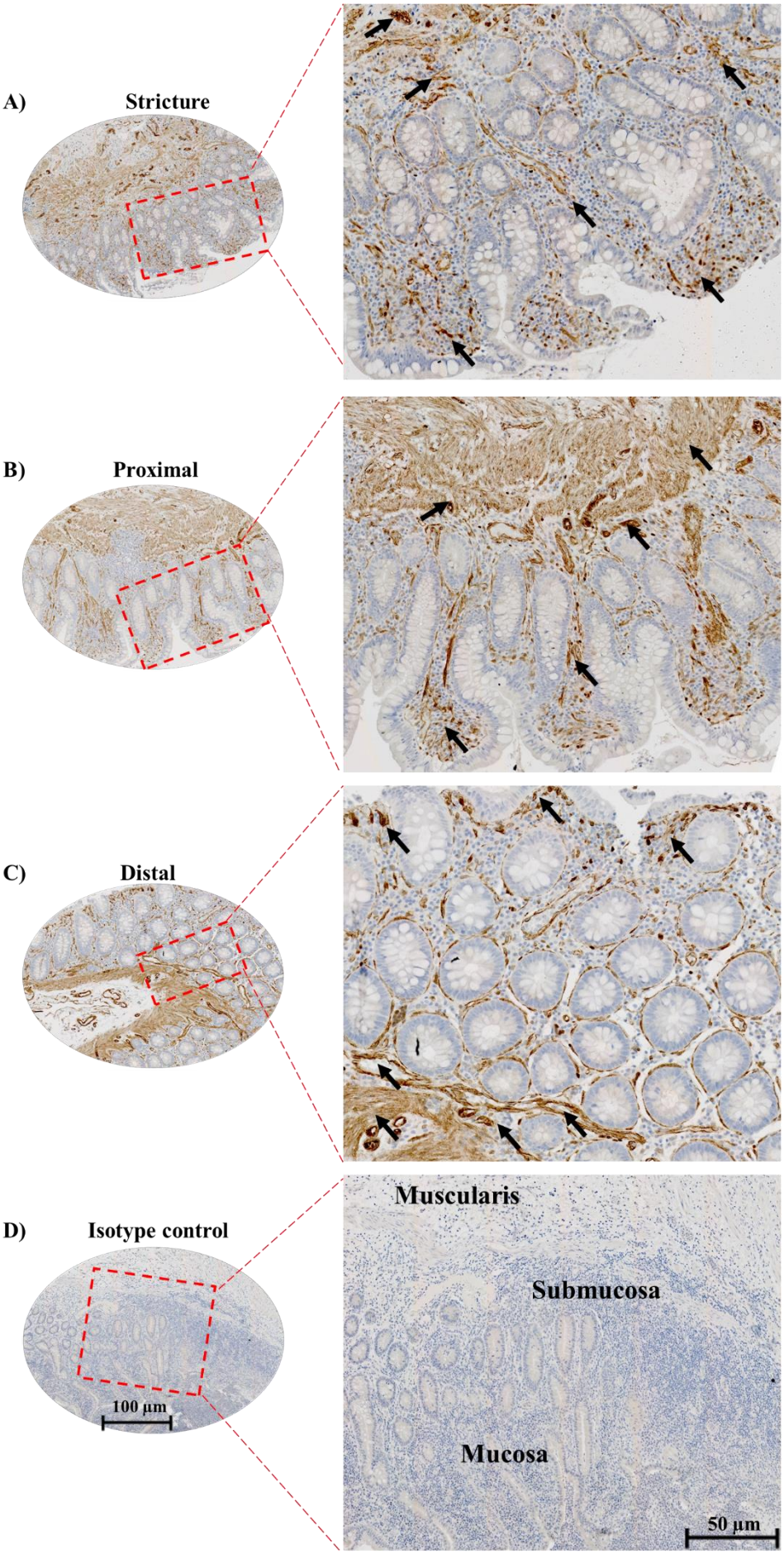

A)
Stricture
B)
Proximal
C)
Distal
D)
Isotype control
100 µm
Muscularis
Submucosa
Mucosa
50 µm

Figure 4.8: FGF2 protein expression in the proximal, stricture and distal segments of resected gut. FFPE were dewaxed and de-paraffinised, antigen retrieval was performed in basic (pH9) antigen retrieval solution. The activity of peroxidases was blocked and then sections were stained for 1 hour with unconjugated primary antibody mouse IgG2a FGF2 (Santa Cruz, USA), visualisation of positive staining was performed using DAB while nuclei were stained with Mayers haematoxylin. Tissue was dehydrated and fixed in DPX for preservation. FGF2 positive staining appears as a brown colour as indicated by the black arrows while the cell nuclei appear blue. Representative images shown include resected tissue collected from **A)** Stricture **B)** Proximal, **C)** Distal and **D)** isotype control (mouse IgG2a). Imaging was performed using the Axio slide scanner (Zeiss) scale bars: 100μm and 50μm respectively.

4.3.7.2 HDAC1 protein expression in resected gut tissue from CD patients

HDAC1 belongs to a family of enzymes whose main function is deacetylation of histones in epigenetic modulatory mechanisms that regulate gene transcription. HDAC1 is ubiquitously expressed in most tissues in the body including the intestine and in multiple cell types. An experimental model of HDAC1-/- in intestinal epithelial cells resulted in disruption of the epithelial barrier therefore revealing a key role of HDAC1 in maintenance of tissue homeostasis (Turgeon et al., 2013). Studies have determined the role of HDAC1 in intestinal fibrosis however, the data is predominantly in animal models (Zimmermann et al., 2007). To determine the expression of HDAC1, IHC staining was performed on FFPE sections from the proximal, stricture and distal segments. An isotype control was included to establish background from true positive signal. Figure 4.9 shows the presence of HDAC1 across all layers of the gut including the mucosa, submucosa and the muscularis as indicated by the brown colour showing positive staining. The black arrows point to representative regions of positive HDAC1 staining. As per the protein atlas, HDAC1 expression is not specific to a particular cell or tissue type, this is confirmed in figure 4.9 in which most cells expressed HDAC1 including tissue infiltrating immune and mesenchymal stromal cells. However, there is a pattern of protein expression around the intestinal epithelium within the villi and the base of the crypts. Comparison between FGF2 (figure 4.8) and HDAC1 (figure 4.9) showed that HDAC1 was

ubiquitously expressed as compared to the discrete staining observed for FGF2. The quality and specificity of the staining was determined based on the negative staining seen in the isotype control; HDAC1 negative cells can also be seen across the stained tissue, this affirms the specificity of the staining. From the images in figure 4.9, there are no major differences between the proximal, stricture and distal segments. However, the single chromogenic stain provides an understanding of HDAC1 protein expression pattern in human resected gut.

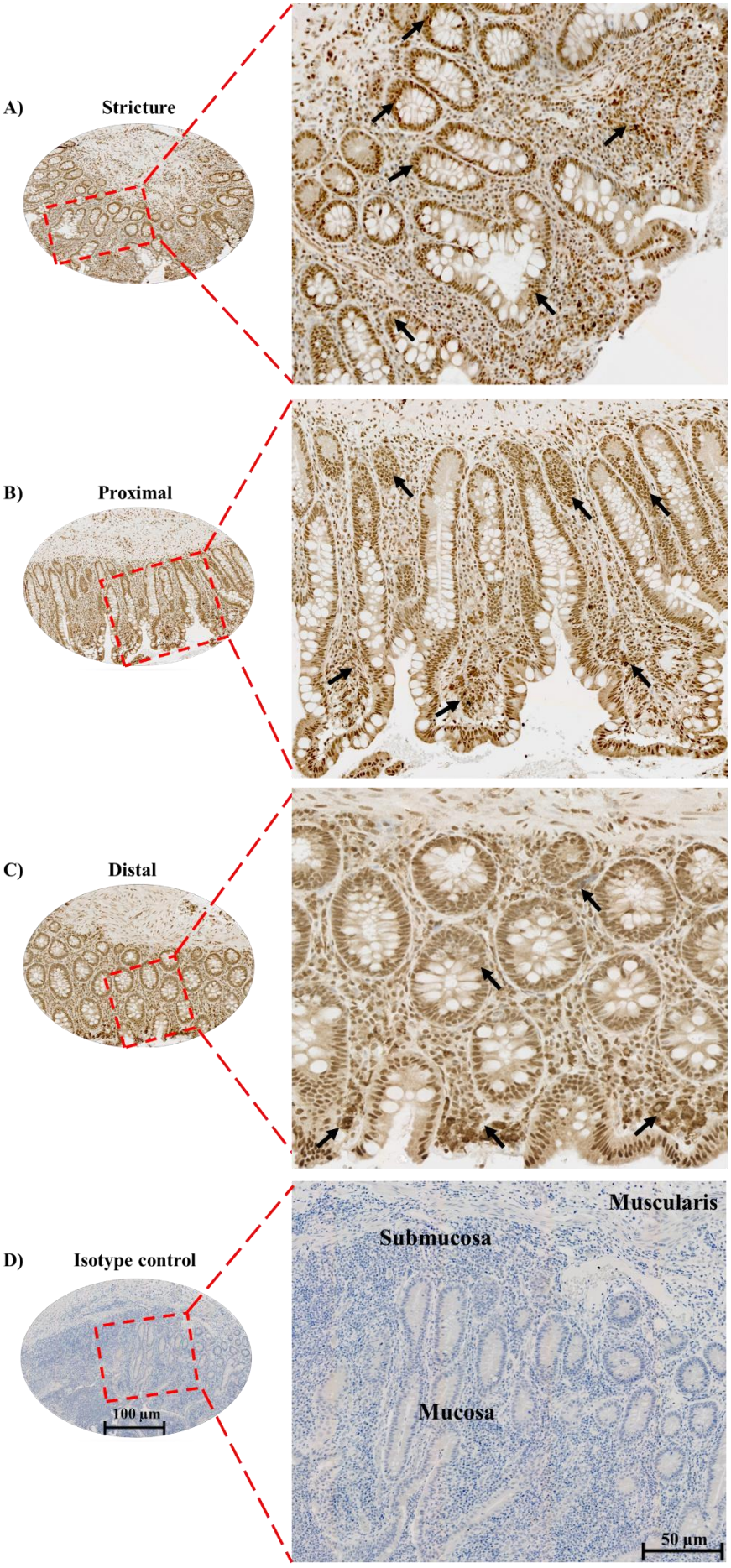

A)
Stricture
B)
Proximal
C)
Distal
D)
Isotype control
100 µm
Muscularis
Submucosa
Mucosa
50 µm

Figure 4.9: Expression of HDAC1 in the proximal, stricture and distal segments of resected gut. FFPE were dewaxed and de-paraffinised, antigen retrieval was performed in basic (pH9) antigen retrieval solution. The activity of peroxidases was blocked and sections stained for 1 hour with unconjugated primary antibody mouse IgG1a HDAC1(Santa Cruz, USA), visualisation of positive staining was performed using DAB while nuclei were stained with Mayers haematoxylin. Tissue was dehydrated and fixed in DPX for preservation. HDAC1 positive staining appears as a brown colour indicated by the black arrows while the cell nuclei appear blue. Representative images shown include resected tissue collected from **A)** Stricture **B)** Proximal, **C)** Distal and **D)** isotype control. Imaging was performed using the Axio slide scanner (Zeiss) scale bars: 100 μm and 50 μm respectively.

4.3.7.3 Gremlin 1 protein expression in resected gut of CD patients

Bulk RNA seq data showed an upregulation of GREM1 at mRNA level in the stricture. In addition, the scRNA-seq data showed expression of GREM1 specific to the cluster of cells identified as stromal cells. However, not all cells in that cluster expressed GREM1, thus the assumption was that there could be a population of stromal cells that expressed GREM1 and others that did not express GREM1. The expression of gremlin 1 at protein level was therefore studied by IHC in the proximal, stricture and distal segments. A few challenges were encountered during the optimisation process. First and foremost, there are very few antibodies designed to detect protein expression of gremlin 1 in tissue by IHC. The one that has been mostly used and validated is the anti-human/mouse gremlin 1 antibody (Cat#AF956 from R and D Systems). Even so, validation for this antibody has mainly been performed with mouse tissue and not in human. The second challenge was the absence of known positive control tissue that could be used during the IHC optimisation process. Gremlin 1 is secreted into extracellular matrix spaces, it therefore gets trapped in ECM fibres, thus it becomes quite difficult to ascertain the true positive staining from artefacts. The next section covers the steps taken to optimise conditions for detecting protein gremlin 1 by IHC staining.

4.3.8 Optimisation of gremlin 1 detection by IHC staining

The literature suggests a high expression of gremlin 1 in mouse kidney (Church et al., 2017). Initial staining was performed on FFPE sections from resected gut alongside those from mouse kidney as a positive control. FFPE sections were dewaxed, deparaffinised and antigen retrieval performed at a basic pH (pH9). BLOXALL reagent was used to block activity of peroxidases. Tissue was stained at a concentration of 3µg/ml of the primary antibody (anti mouse/human gremlin 1) with the corresponding isotype control for 1 hour, protein expression was visualised using DAB staining. No positive staining was seen in either the resected gut or the mouse kidney. Troubleshooting was performed in which the antibody concentration used may have been low to detect the antigen; the recommended concentration for use by the manufacturer was at 3-15 µg/ml. Therefore, the staining procedure was repeated in resected gut tissue with gremlin 1 antibody at a higher concentration of 10µg/ml and 15µg/ml, isotype control was included. The resulting staining for both concentrations showed a strong brown unspecific background.

To resolve the high unspecific background staining, two peroxidase blocking solutions were tested; BLOXALL which is an already prepared solution for peroxidase blocking and 3% hydrogen peroxide (prepared in house) were compared (Gao, Wang and Han, 2008). Staining was performed at an antibody concentration of 8 µg/ml, peroxidase blocking was performed for 30 minutes as compared to the 15 mins from the previous conditions. Unspecific binding was blocked using casein. As seen in figure 4.10, the background was significantly reduced with BLOXALL as compared to using hydrogen peroxide. The black arrows show positive staining for gremlin 1 (figure 4.10B and 4.10D). Although some background is still seen in the sections stained with the isotype control (figure 4.10A and 4.10C), the pattern for gremlin 1 positive stains was localised within the subepithelial layer around the crypts. This optimisation

informed the decision to use BLOXALL to reduce background for further gremlin 1 protein expression studies. Also, extra wash steps were included to get rid of unspecific binding.

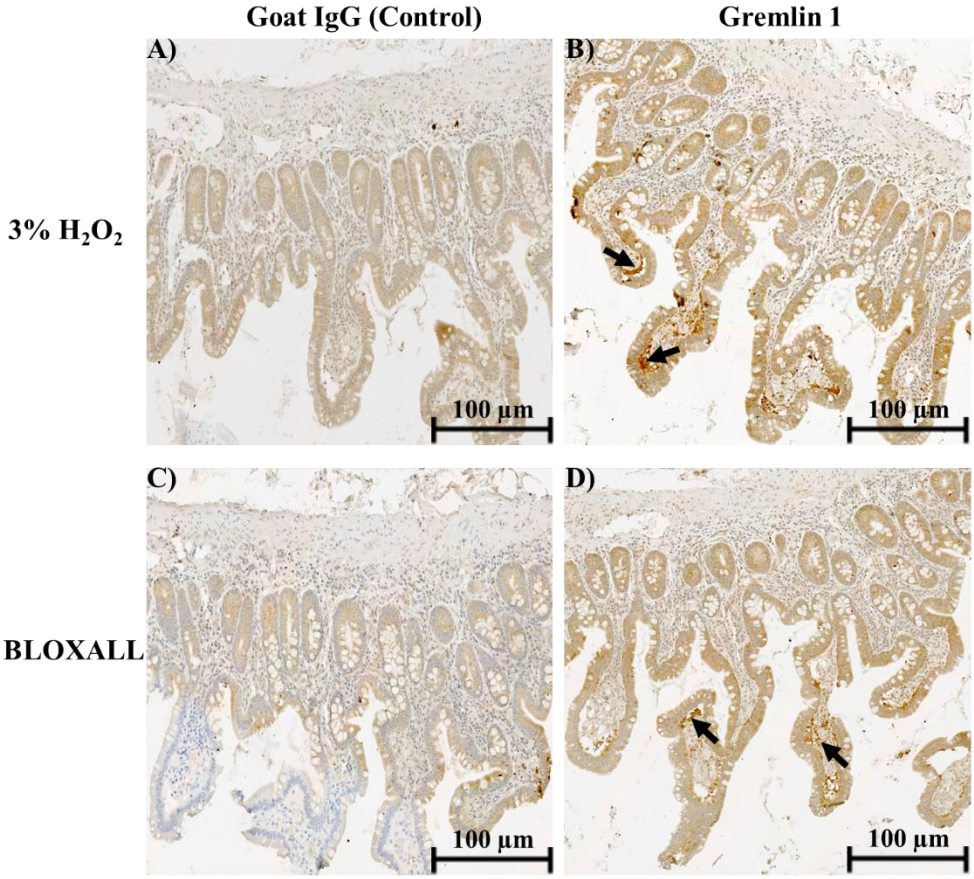


Figure 4.10: Comparison of two peroxidase blocking solutions for gremlin 1 IHC optimisation. FFPE sections were dewaxed and de-paraffinised, antigen retrieval was then performed in basic (pH 9) antigen retrieval solution. The activity of peroxidases was blocked with either BLOXALL or 3% hydrogen peroxide (H_2O_2); sections were stained for 1 hour with the unconjugated gremlin 1 and isotype control (8 µg/ml), visualisation of positive staining was performed using DAB while nuclei were stained with Mayers haematoxylin. Tissue was dehydrated and fixed to preserve in DPX. Gremlin 1 positive staining appears as a brown colour while the cell nuclei appear blue. **A)** Representative image showing isotype control where 3% H_2O_2 was used for blocking peroxidase activity. **B)** Representative image showing gremlin 1 stain where 3% H_2O_2 was used for blocking peroxidase activity. **C)** Representative image showing isotype control where BLOXALL was used for blocking peroxidase activity. **D)** Representative image showing gremlin 1 stain where BLOXALL was used for blocking peroxidase activity. Imaging was performed using the Axio slide scanner (Zeiss) scale bars: 100 µm.

4.3.8.1 Gremlin 1 protein expression in resected gut of CD patients

To understand the expression pattern of gremlin 1 in the proximal, stricture and distal segments, IHC was performed using the optimised protocol described in section 4.3.8 (use of BLOXALL as the peroxidase blocking solution with extra washes to limit unspecific binding. Figure 4.11 shows expression of gremlin 1 in all three segments as highlighted in areas where the black arrows are pointing, gremlin 1 positive cells are seen within the mucosal and submucosal layers. The staining pattern appears cytoplasmic and takes up a longitudinal shape like that of fibroblasts, myofibroblasts or pericytes. Based on staining intensity showing both weak and strong brown staining, certain cells appear to express increased levels of gremlin 1 as compared to others but this is not segment specific, the brown staining appears to be specific and the isotype control does not show background staining.

Figure 4.11A shows gremlin 1 protein expression in the stricture segment, brown staining is seen within the lining of the crypts and in the submucosal regions. There is very low protein expression in the proximal as seen in figure 4.11B in comparison to distal where more brown staining was seen. The protein expression by IHC is comparable to the scRNA-seq data shown in chapter 3 that revealed that GREM1 mRNA expression was specific to stromal cells, which agrees with published literature (Park, 2018). The low protein expression also suggested that not all stromal cells express gremlin 1, this data aligns with the sc-RNA seq data seen in chapter 3. With the images shown, it was not possible to determine the differences in protein expression and tissue localisation. However, the single chromogenic stain provides an insight into gremlin 1 protein expression levels in the resected gut.

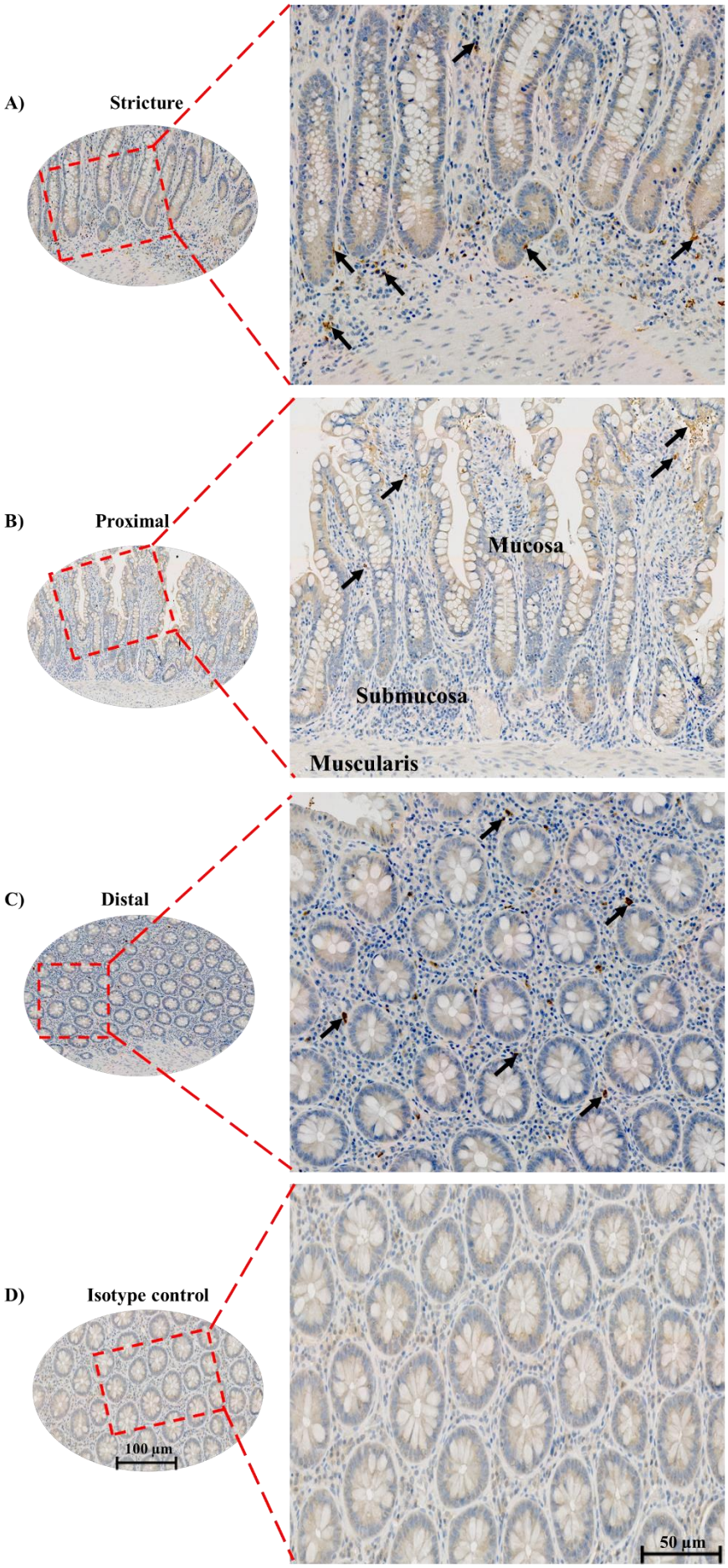

A)
Stricture
B)
Proximal
Mucosa
Submucosa
Muscularis
C)
Distal
D)
Isotype control
100 µm
50 µm

Figure 4.11: Expression levels of gremlin 1 in the proximal, stricture and distal segments of resected gut. Formalin fixed paraffin embedded tissue sections (n=1) were dewaxed and de-paraffinised, antigen retrieval was then performed in basic (pH9) antigen retrieval solution. The activity of peroxidases was blocked using BLOXALL for 30 minutes, sections were stained for 1 hour with unconjugated gremlin 1 primary antibody (8 μg/ml; R and D), visualisation of positive staining was performed using DAB while nuclei were stained with Mayers haematoxylin. Tissue was dehydrated and fixed in DPX for preservation. Gremlin 1 positive staining appears as a brown colour indicated by the black arrows while the cell nuclei appear blue. Representative images shown include resected tissue collected from **A)** stricture **B)** proximal, **C)** distal segments and **D)** Isotype control staining performed on distal segment. Imaging was performed using the Axio slide scanner (Zeiss) scale bars: 100 μm and 50 μm respectively.

4.4 Discussion:

Myofibroblasts have widely been reported as the main effector cell type that drive tissue fibrosis (436). In chapter 3 of this thesis, bulk RNA-seq and scRNA- seq data was used to characterise cellular and molecular phenotypes in the proximal, stricture and distal segments of resected gut in CD samples. In general, there is a lack of an in depth understanding of specific cell types, biological functions and mechanisms through which strictures in CD are formed. MSCs to which fibroblasts, myofibroblasts and pericytes are known to be a heterogenous population, one of the main challenges being faced in MSC studies is lack of specific cellular biomarkers for definition of phenotype. To date, a combination of surface markers is being used to define fibroblast populations, these include vimentin, CD90 (Thy-1), α-SMA and desmin to differentiate myofibroblasts and pericytes from smooth muscle cells.

In this study, FFPE sections were stained with stromal cell markers that in combination can be used to identify stromal cells, vimentin^{+ve} staining was seen in the proximal, stricture and distal segments, vimentin is ubiquitously expressed by cells of stromal origin and as seen from the

IHC staining, was expressed across all layers of resected gut including the mucosa, submucosa and muscularis. α-SMA was used to identify myofibroblasts and pericytes as both cell types express this protein; α-SMA^{+ve} staining can be seen across all the layers of the gut in the proximal, stricture and distal segments. Fibroblasts were defined by their expression of vimentin, myofibroblasts were defined by their expression of vimentin and α-SMA with negative expression of desmin. Desmin can be used to differentiate myofibroblasts (desmin^{-ve}) from pericytes and smooth muscle cells (desmin^{+ve}).

Another focus of this section of the study was to investigate protein expression patterns of FGF2, HDAC1, and gremlin 1. Increased FGF2 expression in tissue fibrosis has been shown to have a protective role in repair of the epithelium. FGF2 signals through FGFRs1, 2, 3 and 4 using the MAPK signalling cascade, the target genes activated downstream are important in wound healing and repair of the epithelium (513). Studies also show that FGF2 is an antagonist to TGF-β signalling in fibroblasts hence increased expression of FGF2 attenuates TGF-β induced fibroblast activation (514). A study conducted by Tiede (2009) in skin fibrosis showed the presence of FGF2 inhibits differentiation of stromal cells to myofibroblasts (515). From the IHC data, positive FGF2 staining was seen in the proximal, stricture and distal segments.

Epigenetic mechanisms have been reported as important in regulating gene transcription in fibrosis. HDACs are involved in epigenetic regulation through the removal of an acetyl group from the DNA histones. Differential HDAC1 expression was reported in both bulk and scRNA-seq data from this study. Studies have shown that overexpression of HDACs increases secretion of pro-inflammatory and pro-fibrotic molecules such as IL-1β, TNF-α and IL-6 (502,516,517). HDAC1 has been shown to regulate pro-inflammatory mechanisms in epithelial and fibroblast-like cells (447,518,519). In addition, data has shown overexpression of HDAC1 in fibrotic conditions such as IPF, protein expression has also been reported to be associated with the

pathogenesis of intestinal fibrosis. However, little is known about expression of HDAC1 in intestinal fibrosis, this was of interest to further explore the protein expression profile of HDAC1 in the proximal, stricture and distal segments.

This study has revealed a strong link between *GREM1* expression as seen in the RNA-seq data, overexpression of *GREM1* was a strong indicator of patients developing fibrostenosing CD strictures. Our scRNA-seq data showed gremlin 1 as specifically expressed in the cell population identified as stromal cells, a study by Ren (2019) showed fibroblast specific expression of *GREM1* RNA using in situ hybridization (ISH) (520). In this chapter, the aim was to further explore protein expression of gremlin 1. Immunohistochemistry staining revealed positive staining for human gremlin 1 which was enhanced in the distal and stricture as compared to the proximal. The IHC staining showed gremlin 1 expression mostly localised to the subepithelial layers of the crypts. The expression is quite weak across the three segments, this could be due to the quality of the antibody. Overall, there was no major differences in staining intensity or the pattern of staining across the three segments.

Gremlin 1 as an antagonist of the BMP pathway, it inhibits signalling of this pathway downstream (521). There have been reports of cross talks between the BMP and TGF-β signalling pathways; the two pathways share SMAD4, a key co-factor and regulator in downstream signalling. Therefore, inhibiting the BMP pathway upregulates the activity of TGF-β signalling downstream. TGF-β/SMAD is well documented as a key pathway for fibrosis regulating mechanisms that are important in fibroblast differentiation, activation and proliferation hence exacerbating tissue inflammation and fibrosis (522). Despite the evidence that indicates the importance of gremlin 1 in fibrosis, there is still very little understanding of the mechanisms through which gremlin 1 regulates intestinal fibrosis. Hence the need to further study the role of gremlin 1 in intestinal fibrosis, this is covered in the next chapter.

CHAPTER 5

DETERMING THE ROLE OF FIBROBLAST DERIVED GREMLIN 1 IN THE DEVELOPMENT OF INSTESTINAL FIBROSIS

Overview of this chapter

In the previous chapters I demonstrated expression of gremlin 1 by mesenchymal stromal cells at a transcript level with an upregulation in the stricture compared to the proximal and distal segments of intestinal resected gut. This section of the study focused on the role of gremlin 1 in regulation of pathways that contribute to a profibrotic microenvironment in intestinal fibrosis, it explores isolation of intestinal primary fibroblasts with the aim of studying the role of gremlin 1 in the development of intestinal fibrosis in vitro.

5.1 Introduction

Several mechanisms drive the fibrotic process, these processes are either shared across different organs or are tissue specific (436,523). Fibroblasts are the main effector cells of intestinal fibrosis; they can acquire an activated myofibroblast phenotype in the presence of profibrotic stimuli, with increased contractility higher expression of ECM proteins (524). In an inflamed microenvironment, fibroblasts are activated by profibrotic factors which include cytokines, chemokines and growth factors. The common profibrotic growth factors include TGF-β1 (TGF-β), FGF-2, PDGF-BB, IL-1β and TNF-α. These mediators in a profibrotic microenvironment regulate multiple pathways that lead to a highly activated and proliferating phenotype of myofibroblasts (436).

Bulk RNA and scRNA-seq data shown in chapter 3 revealed upregulation of profibrotic genes in the stricture as compared to the proximal and distal segments. The bulk RNA-seq data also showed an upregulation of genes linked to stromal cells and ECM production in the stricture, these included *COL1A1, COL1A2, FN1* and *ACTA2*. This pattern of gene expression in the stricture re-enforced what is already published in the literature on the important role stromal cells play in the pathogenesis of fibrotic diseases (436). Among the profibrotic mediators that were upregulated in the stricture was GREM1 (Gremlin 1). In addition, scRNA-seq data from this study showed expression of *GREM1* within a cluster of cells identified as stromal cells. The data showed a trend of increased expression of *GREM1* in the stromal cell cluster within the stricture segment as compared to those within the proximal segment. These findings were not conclusive because statistical analysis was not performed due to the low cell numbers used for analysis in the scRNA-seq data. Gremlin 1 was of particular interest due to its expression pattern both in the bulk and scRNA-seq data, and due to its expression in various fibrotic diseases where it has been identified both as an important mediator of fibrosis and as a potential

target for antifibrotic therapy, this was a justification to further understand the mechanisms through which gremlin 1 modulates a profibrotic phenotype in stromal cells (489,507,525).

Gremlin 1 is a highly conserved protein comprising 184 amino acids with a molecular weight of 20.7kDa (520), it acts as an antagonist of the bone morphogenetic protein (BMP) pathway. BMPs are a group of extracellular signalling proteins that belong to the TGF-β superfamily, these family of proteins play an important role in tissue homeostasis regulating cell growth and proliferation, wound healing, tissue remodelling and organ development (526,527). Gremlin 1 binds ligands of the BMP pathway, predominantly BMP 2 and BMP 4 with weaker binding to BMP 7 (521). The cognate receptors for BMP signalling molecules are BMPRIA, BMPRIB and BMPRII, with BMPRIB being reported as having the lowest expression across various tissues. The receptors are expressed on multiple organs such as the heart, liver, kidney, intestine and lung (528). BMP receptors are also widely expressed in multiple cell types, with intestinal expression localised to sub-epithelial cells within intestinal crypts (529,530).

TGF-β and BMP proteins are structurally related cell regulatory proteins that belong to the TGF-β superfamily of growth factors; these proteins modulate cellular function by signalling through TGF-β and BMP signalling pathways (460,531,532). TGF-β signalling is well documented as a promoter of profibrotic mechanisms in various organs such as the liver, kidney, heart and lung, whereas BMP signalling has been linked to antifibrogenic mechanisms (533–536). Studies were performed in a bleomycin-induced mose model of IPF comparing the wild type and Noggin +/LacZ (Nog +/LacZ) phenotype which is a known model of increased BMP signalling. The results from this study showed partial protection from bleomycin induced lung fibrosis and reduced collagen deposition in the Nog +/LacZ mice. The wild-type group showed increased activation of TGF-β signalling and repressed BMP signalling. The results from this study suggest cross talks between the two pathways, modulating balance between

TGF-β/BMP signalling, particularly increasing endogenous BMP signalling presents potential therapeutic targets for organ fibrosis (537).

Figure 5.1A shows TGF-β and BMP signalling pathways being maintained in a balanced state, changes in the microenvironment may result in a dysregulated state between the two pathways as seen in figure 5.1B. Gremlin 1 has been suggested to cause dysregulation in the TGF-β/BMP signalling pathways; the proposed mechanisms involves increased levels of gremlin 1 in a fibrotic tissue microenvironment results in inhibition of BMP signalling through binding to BMP2 and BMP 4; these are ligands of the BMP signalling pathway. Inhibition of BMP signalling results in an increase in TGF-β signalling (figure 5.1D). Both pathways signal through SMAD4 as a co-factor, this may suggest why inhibition of the BMP pathway results in increased signalling of the TGF-β pathway (509).

The signalling pattern for BMP and TGF-β pathways is illustrated in figure 5.1, in which the ligands bind to their cognate type II receptors forming complexes (figure 5.1C), leading to phosphorylation of type I receptors. Type I receptors then generate a transduction signal to activate the respective intracellular small mothers against decapentaplegic (SMAD) dependent signalling (figure 5.1D); SMAD 2/3 for the TGF-β pathway and SMAD 1/5/8 for BMP pathway with SMAD4 shared as a co-factor. SMAD proteins are phosphorylated and form a complex then translocate to the nucleus to activate expression of target genes. Upregulation of profibrotic genes such as *COL1A1*, *ACTA2*, *FN1* and *TIMP1* is seen downstream the TGF-β signalling pathway while antifibrotic genes such as Id1, Id2, Id3 and SMAD7 are upregulated following activation of the BMP signalling pathway (530). Gremlin 1 as an antagonist of BMP signalling is most likely to cause a dysregulation in TGF-β/BMP signalling axis. Therefore, upregulation of gremlin 1 as seen in our bulk RNA-seq data specifically in the stricture is a negative predictor for the development of fibrosis. In addition, the role of gremlin 1 in the

pathogenesis of intestinal fibrosis has not been fully elucidated. This section of the study was aimed at understanding the role of gremlin 1 in primary intestinal stromal cells and their role in the pathogenesis of intestinal fibrosis.

5.2 Aims of this chapter

The specific aims of this chapter were to;

1. To isolate and characterise primary stromal cells from proximal, distal and stricture resected gut tissue obtained from CD patients.
2. To study the role of gremlin 1 in the regulation of the TGF-β/BMP signalling pathways in primary fibroblasts
3. To understand the mechanism through which gremlin 1 modulates the stromal cell phenotype and its role in the pathogenesis of intestinal fibrosis.

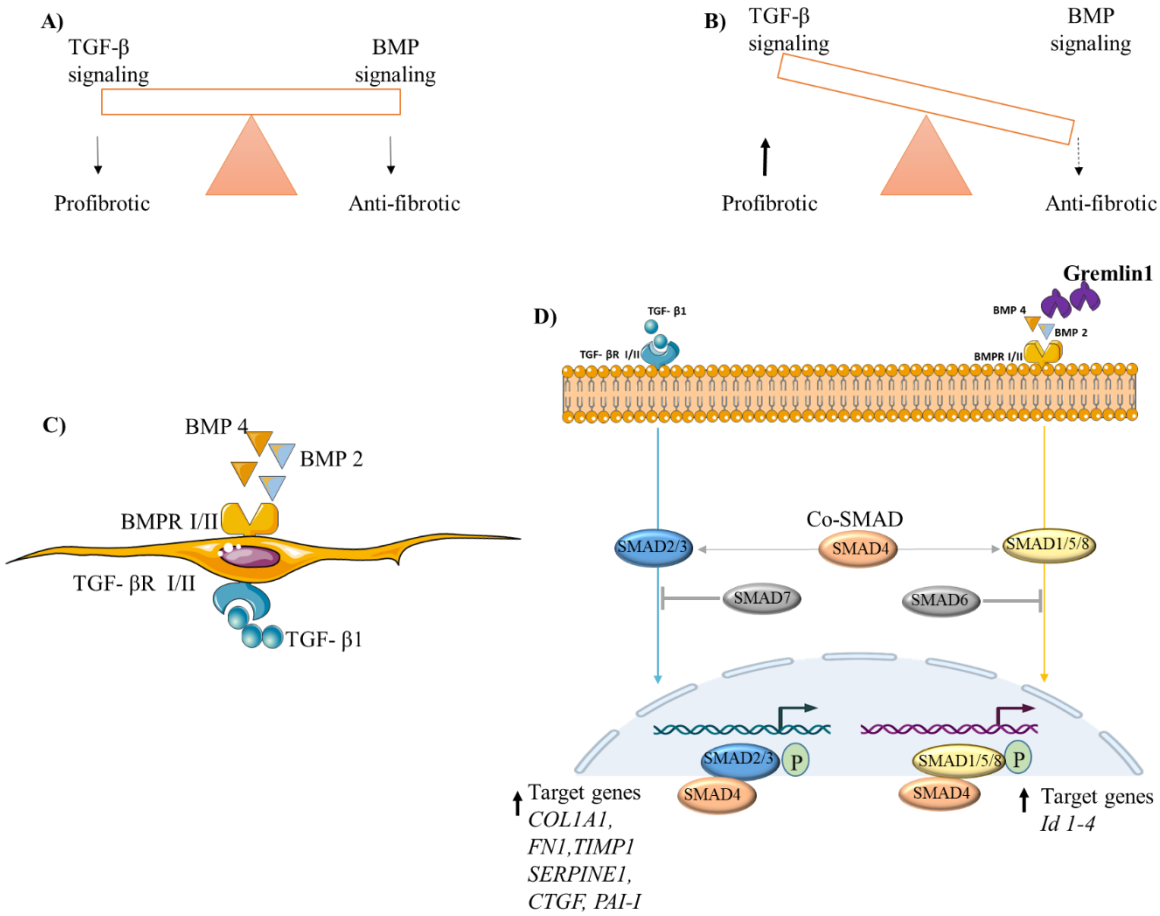


Figure 5.1: Schematic representation of regulation of the TGF-β/BMP signalling. A) Balanced signalling of TGF-β and BMP pathways in a physiological state with maintained tissue homeostasis **B)** Dysregulation in signalling between BMP/ TGF-β pathways; a decrease in BMP signalling due to inhibition by gremlin 1 results in increased TGF-β signalling and upregulation of target genes downstream the TGF-β pathway (dysregulation of the 2 pathways has been reported in fibrosis and cancer related diseases) **C)** TGF-β signals through its cognate receptors TGF-βRI/II while BMP pathway signals through ligands 2 and 4 and their cognate receptors BMPRI/II. **D)** Shows downstream mediators of signalling of the two pathways, where TGF-β signals through SMAD2/3 with SMAD 7 acting as an inhibitory factor. BMP pathway signals through SMAD1/5/8 with SMAD6 as the inhibitory factor. Both pathways use SMAD4 as a cofactor. Gremlin 1 as an antagonist of BMP pathway binds BMP2, 4 and 7 inhibiting signalling at the BMP- axis.

5.3 Results

5.3.1 Isolation of primary gut fibroblasts

To study the role of gremlin 1 in the modulation of TGF-β/BMP signalling in intestinal fibroblasts, a protocol for isolation of intestinal fibroblasts was optimised as there was no protocol specific for isolation of intestinal fibroblasts that had been developed within the lab at the time. There was an established protocol for isolation of fibroblasts from explanted liver tissue. The established "explant liver" protocol was used as a starting point for the optimisation because of the easier access to explant liver as compared to resected gut.

5.3.1.1 Isolation of primary gut fibroblasts using the "explant liver" protocol

Diseased explant liver tissue was collected and sliced into 3 - 5 mm pieces, before being washed in cold PBS and digested in collagenase IV and 1A. The tissue digest was layered on Percoll gradient and centrifuged to sediment debris and erythrocytes. The buffy coat that contains mainly leukocytes and stromal cells was collected (figure 5.2B), washed and resuspended in antibiotics containing 16% FBS +DMEM culture media, cells were left in culture at 370C while changing media every 2-3 days. After 4-10 days depending on the sample, spindle shaped fibroblast like cells appeared (figure 5.2 C-E). This was an effective method for liver fibroblast isolation from explanted livers of different aetiologies such as primary sclerosing cholangitis (PSC), alcohol related liver disease (ALD) and donor explant livers as shown in figure 5.2 below. The same protocol was applied on tissue from proximal, stricture and distal segments of gut resection samples where the starting material was approximately 0.5-1 gram from each segment. However, the use of this protocol was not successful in isolation of intestinal fibroblasts.

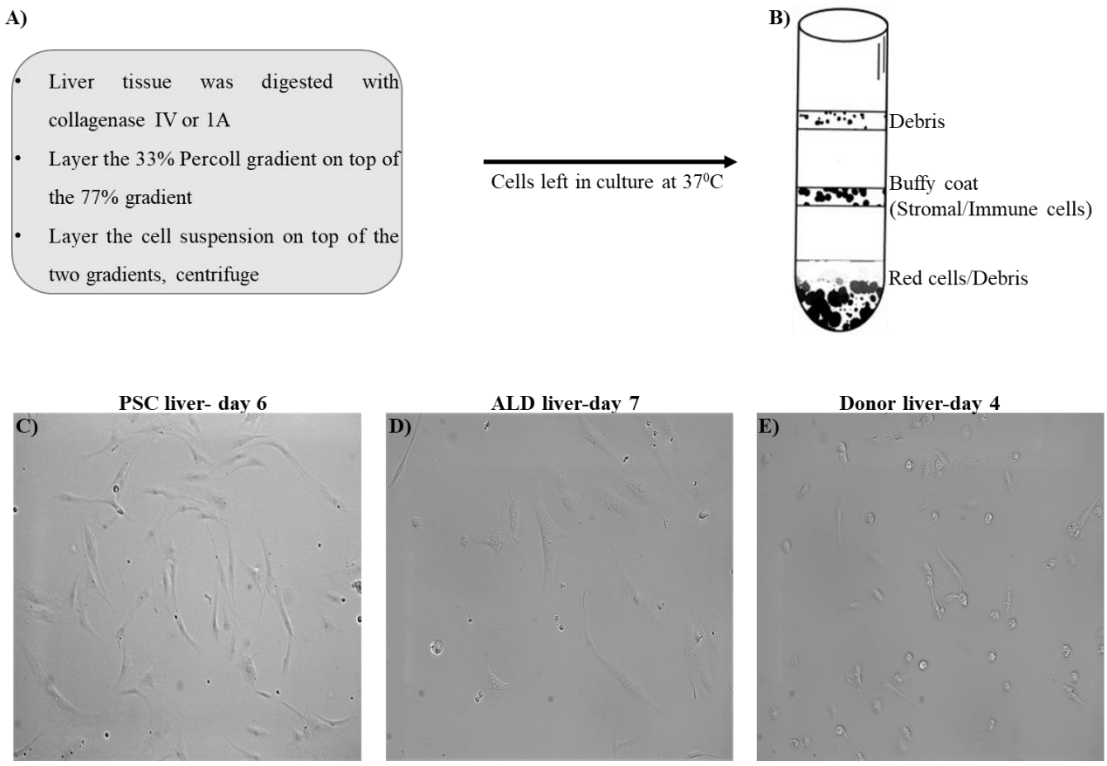


Figure 5.2: Isolation of primary fibroblasts from explanted liver tissue. A) Diseased explant was sliced into 3-5mm pieces and digested in serum free DMEM media containing collagenases IV (4,000 U/mL) and 1A (100 µg/mL), the tissue digest was layered on Percoll gradient at 77% and 33% respectively and centrifuged at 800 x *g* for 20 mins. **B)** Cells from the buffy coat were collected, washed and cultured at $2.5X10^6$cells per ml of media (16% FBS +DMEM) at 37^0C in CO_2. Spindle shaped cells with a fibroblast like morphology appear from day 4-10. Primary fibroblasts isolated from explanted liver from multiple conditions; **C)** PSC liver at day 6. **D)** ALD liver at day 7. E**)** Donor liver at day 4. Images captured on an inverted Zeiss bright field microscope n=2

5.3.1.2 Isolation of primary fibroblasts from gut resected tissue using an organ culture protocol

The isolation of primary gut fibroblasts using the "explant liver" protocol was unsuccessful, a literature search showed a published protocol by (538) on isolation of primary fibroblasts from small size tissue samples, biopsies collected from the eye were used to isolate primary fibroblasts (538). The same protocol was used by (539) to isolate gut resected fibroblasts. In this method, resected gut tissue from the proximal, stricture and distal segments was washed with ice cold PBS, cut into approximately 3-5 mm pieces with a surgical blade, pieces of tissue

were cultured in 16% FBS+DMEM media enriched with 2 ng/ml of human Fibroblast Growth Factor (hFGF), with renewal of media every 3-4 days (figure 5.3). In the protocol set up, tissue was placed on the plate while ensuring that the tissue was stuck to the culture plate to allow fibroblast to migrate from the tissue onto the plate. This was achieved by adding culture media just about covering the tissue while ensuring that the tissue did not float. This was performed on two resected samples containing tissue from the stricture, proximal and distal segments. Tissue in culture highlighted within the red circle and labelled T was monitored over 10-20 days for outgrowth of fibroblast from the tissue, however, it was a failed attempt to use the organ culture method for isolation of primary intestinal fibroblasts from resected gut.

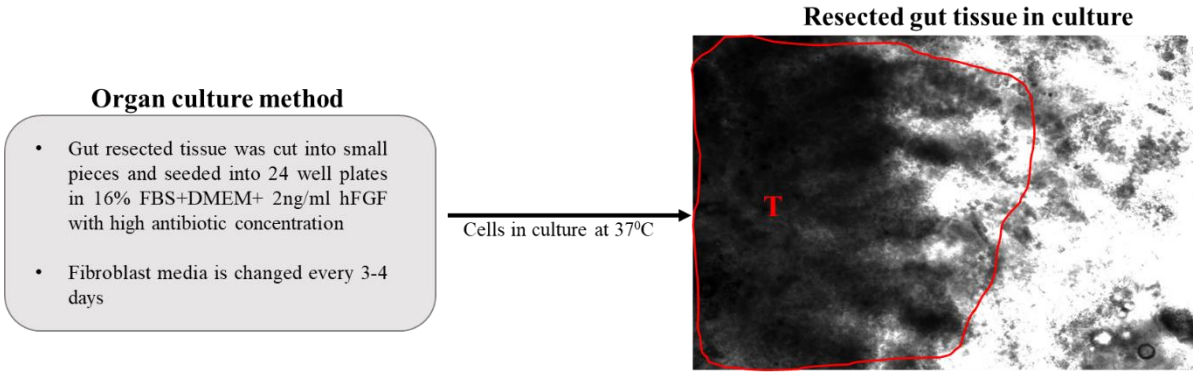


Figure 5.3: Organ culture method for intestinal fibroblast isolation. Resected gut tissue from the proximal, stricture and distal segments was sliced into 3-5mm pieces and placed into a 24 well plate while ensuring attachment of tissue to the plate. (500µl of 16% FBS + DMEM). Media was changed every 3-4 days while monitoring any sign of cell growth over 10-20 days. Images captured on an inverted Zeiss bright field microscope n=6 (2X proximal, stricture and distal)

5.3.1.3 Isolation of primary fibroblasts from gut resected tissue using "tissue enzymatic digestion" protocol

The unsuccess with the organ culture method prompted an urgent need to get a working protocol as access to resected samples was very rare. Samples were analysed by flow cytometry to assess the frequencies of stromal cells in single cells suspensions of stricture, proximal and

distal segments of resected gut tissue. Figure 5.4A-C show very high frequencies of epithelial cells (EpCAM labelling), as compared to endothelial cells in the proximal, stricture and distal segments. The next protocol was developed so as to strip the abundant epithelial cell layer in an effort to enrich the fibroblast population prior to subsequent cell culture.

In this protocol, HBSS and HEPES was included in the wash buffers to maintain cell viability. Resected tissue was further incubated with stripping buffer that contained EDTA, HBSS and HEPES with the aim to strip the epithelial layer to reduce frequencies of epithelial cells and enrich fibroblasts. Tissue was digested in an enzymatic cocktail of collagenase and Dispase enzymes to ensure complete dissociation of any interlocking bonds in tissue extracellular matrix proteins and to achieve optimum cell yield (details provided in Chapter 2). The resulting cell suspension was filtered, washed and left in culture in 16% FBS+DMEM media enriched with 2 ng/ml of human Fibroblast Growth Factor (hFGF), and the media was being changed every 3-4 days (figure 5.5B) while monitoring for fibroblast growth. Two resected samples were processed using this protocol, but all attempts were unsuccessful, figure 5.4C shows dying cells in culture.

The three protocols that had been tested for the isolation of primary gut fibroblasts including the explant liver protocol, the organ culture protocol and enzymatic digestion protocol were unsuccessful. The next step in the optimisation process was to use a protocol that minimised cell loss especially during multiple wash steps while preserving cell viability. This was referred to as a “hybrid” isolation protocol where mechanical and enzymatic digestion was combined.

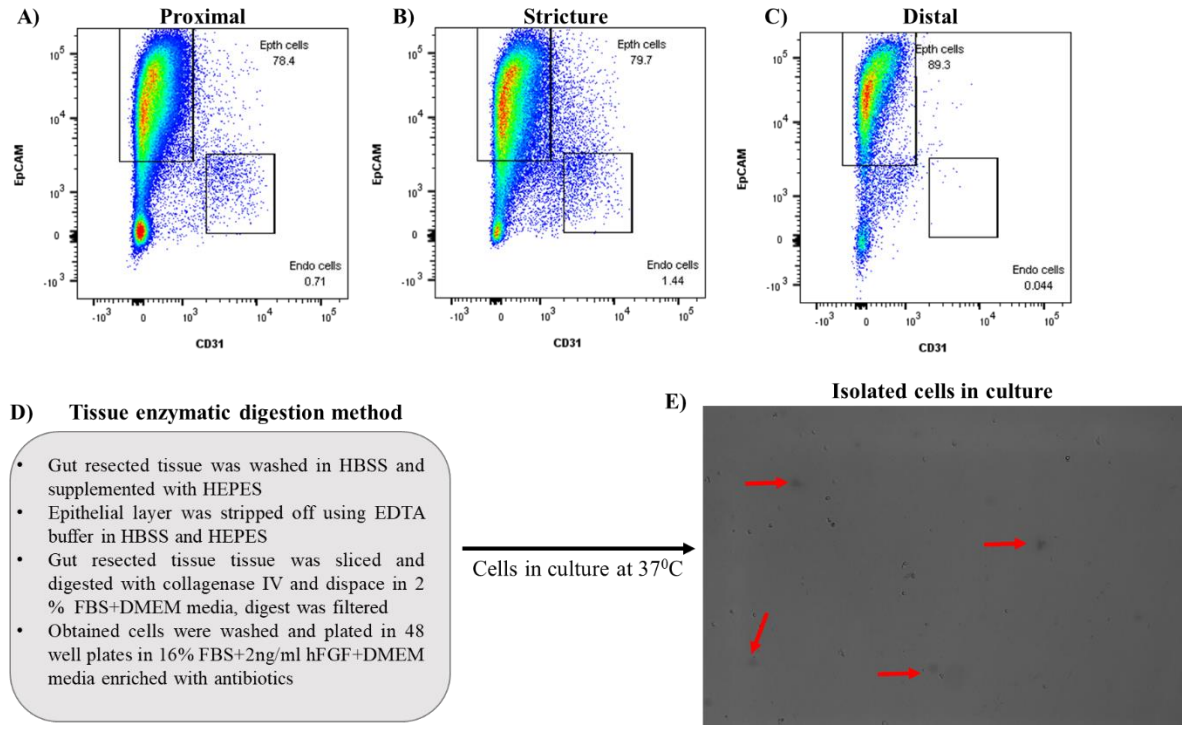


Figure 5.4: Isolation of intestinal fibroblasts using the tissue enzymatic digestion protocol. A-C) Plots showing flow cytometry performed on cells isolated from resection tissue (proximal, stricture and distal) segments. The gates show epithelial and endothelial cells respectively, EpCAM was used to identify epithelial cells while PECAM-1(CD31) was used to identify endothelial cells. **D-E)** Resected gut tissue from the proximal, stricture and distal segments was sliced into 3-5mm pieces and washed in calcium and magnesium free HBSS supplemented with HEPES. Tissue was then incubated with stripping buffer containing HBSS, EDTA (5mM), HEPES (15mM) and 10% FBS and then incubated at 37^0C for 15 mins to strip the epithelial layer; tissue was then digested with an enzymatic cocktail of Dispase and Collagenase IV at 37^0C for 30 minutes. The tissue digest was filtered through a 70 μm filter, cells were collected by centrifugation, an extra wash was performed using 2% FBS+PBS. Cells were then counted and cultured (37^0C CO_2) at $1X10^6$ cells/ml in 16% FBS + 2ng/ml hFGF + DMEM media using a 48 well plate. Cell growth was monitored over a period of 10-20 days n=6 (2X proximal, stricture and distal).

5.3.1.4 Isolation of primary fibroblasts from gut resected tissue using a "hybrid" method of enzymatic and mechanical digestion

The hybrid method combined mechanical and enzymatic digestion protocols. In this method, instead of using collagenases IV and 1A that had previously been used for the liver and enzymatic digestion protocol, Liberase (highly purified collagenase I and II) was used.

Liberase has been reported as a purified enzyme blend that is free of endotoxins, provides maximum tissue dissociation and preserves cell viability (540,541). Resected tissue from the proximal, stricture and distal segments was cut into 3-5mm pieces, tissue was transferred into a C MACS tube containing 2% FBS+DMEM media with Liberase enzyme blend. The C MACS tube with tissue in it was placed on the Miltenyi gentleMACs octo dissociator. This machine contains software that enables setting up a program that is best suited for disaggregation of the tissue type being digested based on tissue stiffness. An incubation protocol was set up on the octo dissociator that included gentle mechanical force to dissociate tissue while an inbuilt rotor that enabled constant mixing of the tissue sample. The incubation protocol included heaters that were set to maintain the tissue at 370C during tissue digestion.

Tissue was digested for 20-30 minutes depending on stiffness. The tissue digest was filtered through a 70 µm strainer, and the resulting cell suspension was washed and left in culture in hFGF enriched media (16% FBS + 2ng/ml hFGF + DMEM). The cell cultures were monitored for cell growth and media changed every 3-4 days. A brief illustration of the hybrid protocol is in figure 5.5. After isolation, primary intestinal fibroblasts were cultured for several days until they were confluent and transferred to a 6 well plate, then to a T25 flask to increase numbers. The cells had a morphological flat and spindle shape (figure 5.5 C-E) which was characteristic of fibroblasts. Once the cells reached 90% confluency, they were split at a ratio of 1:3 and cryopreserved in a 5% DMSO:95% FBS solution ready for future experiments. Using a combination of enzyme digestion and gentle mechanical dissociation (figure 5.5), the hybrid protocol resulted in a 54 % success rate in isolation of intestinal fibroblasts. Of the 11 isolations that were performed including tissue collected from the proximal, stricture and distal segments, primary fibroblasts were collected from 6 samples (3X stricture and 3X proximal). More isolations could have been performed; however, the limiting factor was limited access to resected tissue which was on average one sample per month. Table 5.1 shows a summary of

the samples that were used for the isolation process. Important to note that resected gut samples we collected from the proximal, stricture and distal regions, no healthy tissue was collected for the study.

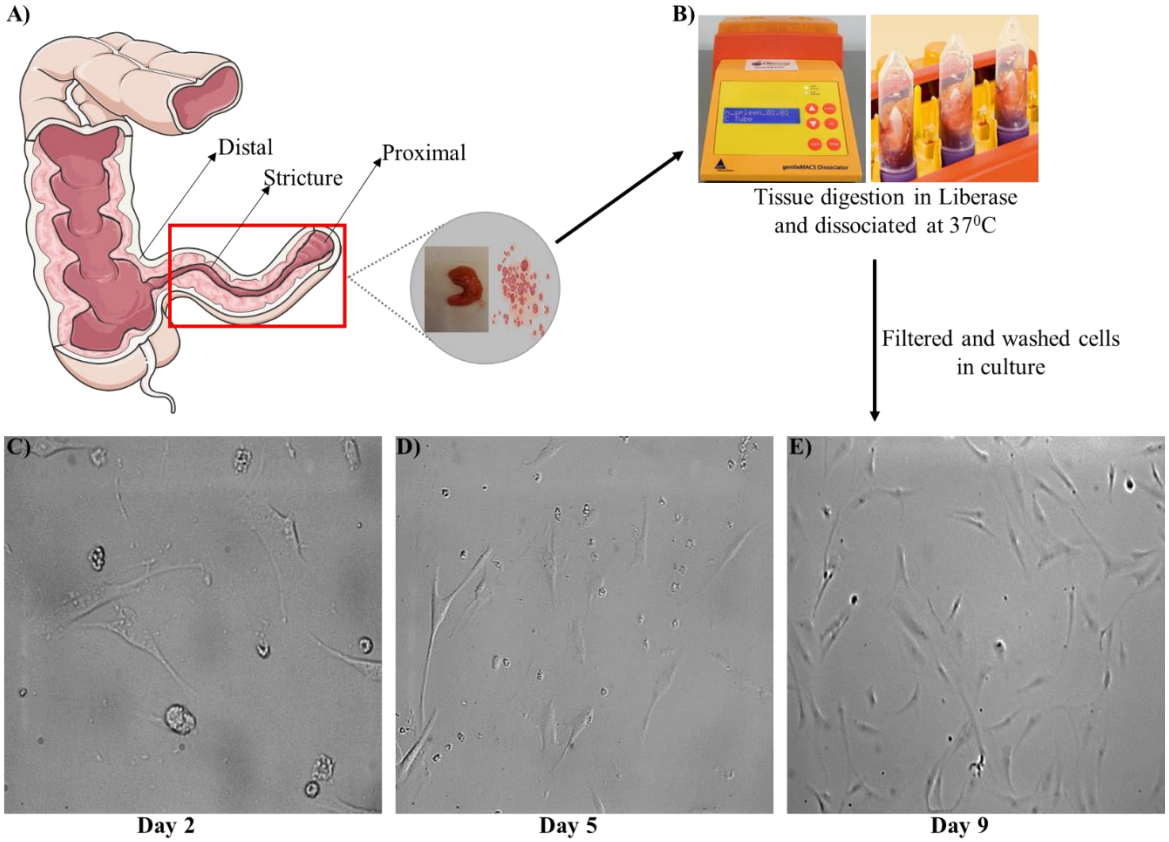


Figure 5.5: Isolation of intestinal fibroblasts from gut resected tissue using the hybrid protocol **A)** Specific sites from which each tissue segment is collected **B)** Tissue sections were sliced into 3-5mm pieces and placed in C MACS tubes containing DMEM media +2% FBS+ Liberase (40µg/mL). The tissue was transferred onto the gentleMACs Octo dissociator with heaters set at 37^0C and a gentle mechanical dissociation was performed for 20-30 minutes. Cell digest was filtered and centrifuged. The pellet was resuspended in 2% FBS + PBS for an extra wash; cells were then counted and seeded at $2X10^6$ cells/ml of DMEM + 16% FBS + 2ng/mL hFGF supplemented with antibiotics in a 24 well plate. The images were taken following a successful isolation of fibroblasts from a stricture sample, **C)** Representative image of primary fibroblast at day 2. **D)** Representative image of primary fibroblast at day 5 **E)** Representative image of primary fibroblast at day 9. Images captured on the Zeiss inverted bright field microscope (n=11).

Table 5.1: FCD cohort samples, experiments performed
(Some samples were used for scRNA-seq or IHC, therefore could not be used for isolation of fibroblasts)

Sample ID	Proximal (g)	Stricture (g)	Distal (g)	Fibroblast Isolation performed	Protocol used	Fibroblasts isolated	Thawed and recovered	Experiments performed
FCD001				Yes	Liver explant protocol	No		
FCD002				Yes	Liver explant protocol	No		
FCD003					Organ culture	No		
FCD004					Organ culture	No		
FCD005	2.26	2.76	1.43	Yes				
FCD006	1.969	No tissue	1.63	Yes				
FCD007	Only blood samples collected							
FCD008	3.1	3.186	3.949	Yes				
FCD009	1.438	1.391	2.149	Yes		Yes (stricture)	Yes (stricture)	Fibroblast characterisation
FCD010	0.95	2.119	1.631	Yes		Yes (Proximal)	No	
FCD011	Only blood samples collected							
FCD012	1.19	1.81	0.9	Yes		Yes (stricture)	Yes (stricture)	Fibroblast characterisation
FCD013	0.668	1.39	1.68			RNA isolation	RNA isolation	RNA isolation
FCD014	2.113	1.48	0.97			RNA isolation	RNA isolation	RNA isolation
FCD015	1.668	1.389	1.484	Yes		Yes (stricture and proximal)	Yes (stricture)	Fibroblast characterisation
FCD016	2.936	3.5	4.585	Yes		Yes (Proximal)	No	
FCD017	1.56	1.4	0.8			RNA isolation	RNA isolation	RNA isolation
FCD018	2.21	1.389	1.46			RNA isolation	RNA isolation	RNA isolation

FCD019	2.89	0.98	1.25			RNA isolation	RNA isolation	RNA isolation
FCD020	1.668	1.389	1.484			RNA isolation	RNA isolation	RNA isolation
FCD021	1.184	0.968	No tissue			RNA isolation	RNA isolation	RNA isolation
FCD022	0.96	1.1	1.32					Single cell RNA sequencing
FCD023	1.5	1.2	1.8					Single cell RNA sequencing
FCD024	1.12	1	1.2					Single cell RNA sequencing
FCD025	1.8	2.8	2.51			FFPE sections	FFPE sections	FFPE sections

5.3.2 Characterisation of fibroblasts and myofibroblasts

The hybrid method was used to successfully isolate intestinal fibroblasts, these were frozen down for future experiments. The next step was to confirm cellular phenotype because fibroblasts have been shown to exhibit cellular plasticity. Isolated fibroblasts from the proximal (n=3) and stricture (n=3) that had been frozen down were thawed and cells left in culture re-suspended in fibroblast media (16% FBS+1% Penicillin-Streptomycin+ DMEM). The aim was to check phenotype using expression of stromal cell markers such as CD34, vimentin, PDGFRα and β, podoplanin, Thy1.1 (CD90) and α-SMA. Frozen fibroblasts stored in liquid Nitrogen were thawed; however, a large proportion of cells died during the thawing process. It is not clear why this happened but a plausible explanation could have been the stress that was induced on the cells during the cryopreservation process. Of the 6 samples that were thawed (3X stricture, 3X proximal), three samples (fibroblasts from the stricture) were recovered in culture while the other three samples (fibroblasts from the proximal) were lost as a large proportion underwent cell death. The recovered fibroblasts were left in culture while changing media every 2-3 days. Unfortunately, the primary fibroblasts did not proliferate and expand any further.

As published in the literature, fibroblasts are cells of mesenchymal origin, they express markers such as CD34 and vimentin. Fibroblasts also express surface and intracellular such as PDGFRα and β, podoplanin, Thy1.1 (CD90) in humans and Thy1.2 in mice. Fibroblasts, myofibroblasts and pericytes express similar markers; the three types can be distinguished from each other through their expression of FSP-1 as fibroblasts and myofibroblasts express FSP-1 while pericytes do not (542), unfortunately, FSP-1 was not included in the flow cytometry phenotyping panel. Myofibroblasts are distinguished from fibroblasts using α-SMA, myofibroblasts express high levels of α-SMA while fibroblasts do not. Myofibroblasts are known as the pathogenic phenotype and the highest producers of ECM proteins (543) .

Epithelial and endothelial cells were characterised by their expression of EpCAM and PECAM-1 (CD31) respectively, these markers were included in the panel. Below is a schematic diagram that summarises some of the markers used in identification of mesenchymal cells.

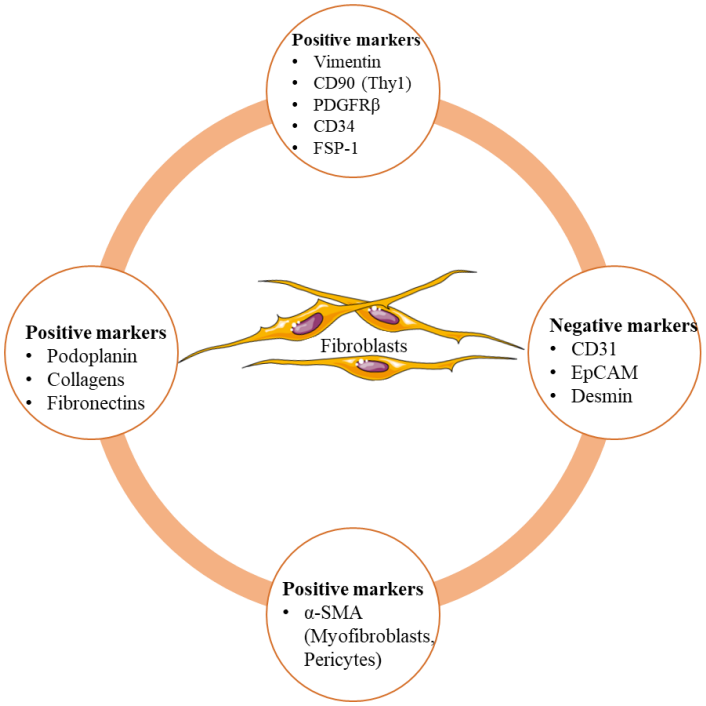


Figure 5.6: Schematic summary of mesenchymal stromal cell identification markers. Positive markers that are used to identify fibroblasts include; **Surface proteins** (Vimentin, CD90, PDGFRβ, CD34, podoplanin, FSP-1); **Negative marker**s for fibroblasts: CD31, EpCAM, α-SMA, Desmin
α-SMA identifies myofibroblasts and pericytes
ECM proteins (collagens and fibronectins) expressed by MSCs.

After testing multiple methods, I finally optimised a working protocol for the isolation of primary intestinal fibroblasts with success in 6 samples (3 stricture and 3 proximal). As mentioned above, isolated fibroblasts were cryopreserved and stored in liquid nitrogen for future experiments. The cells were thawed to perform further experiments; unfortunately, only 3 of the 6 samples were recovered (cell from the stricture samples), the other three samples (cells from the proximal samples) were lost due to cell death during the thawing process. The recovered fibroblasts were left in culture while changing media every 2-3 days. The cells in culture did not proliferate and expand any further, hence it was not possible to phenotype these cells due to very low cell numbers. It is not clear why the cells did not recover after the thawing process but this may be attributed to the stress induced on the cells during cryopreservation.

Supply to resected tissue was limited, I was unable to perform any further fibroblast cell isolations. The next option was to characterise a cell line to confirm the similarity in phenotype to fibroblasts, and then use it as an in vitro cell culture model to perform gremlin 1 studies. Cell lines have been used as in vitro cell culture systems especially in situations where there is limited access to primary cells. The next section covers the characterisation of a gut derived cell line to be used as an in vitro cell culture model for further studies on gremlin 1.

5.3.3 Culture and expansion of colon fibroblast cell lines

Cell lines have long been used as in vitro experimental models to study gene expression profiles and pathway regulation in fibrosis. As mentioned in the previous section, there was limited and unreliable access to gut resection tissue hence inadequate primary fibroblasts for further experimental studies. The use of cell lines for further experimental studies was a good option. Two intestinal cell lines were obtained from the American Tissue Culture Collection (ATCC): Hs255T, a human gut derived fibroblast cell line from an adenocarcinoma donor, and CCD-18Co, a human gut derived fibroblast cell line obtained from a normal donor. The two cell lines were chosen because they are both intestinal derived and morphology was described as fibroblasts by ATCC. The Hs255T cell line was aimed at representing the diseased fibroblast phenotype while CCD-18Co was aimed at representing the normal fibroblast phenotype.

Figure 5.7 shows morphology of the two cell lines in culture, Hs255.T at passage 2 and CCD-18Co at passage 6 respectively. The Hs255.T was cultured under conditions recommended by the ATCC as detailed in the materials and methods chapter. However as seen in figure 5.8 A-B, the cells did not proliferate and always looked stressed, the ATCC had recommended culture media using 10% FBS+1% PS (Pen-Strep) + EMEM (eagle's minimum essential medium). The cells in culture looked stressed under these conditions. Changes were made to the culture media replacing EMEM with DMEM, this did not rescue cells from their non-proliferative and stressed state. The Hs255T cell line was discontinued from further experiments. CCD-18Co

proliferated well in culture using culture conditions recommended by the ATCC. Figure 5.7 C-D shows CCD-18Co cell morphology that is flat and elongated similar to that of primary fibroblasts.

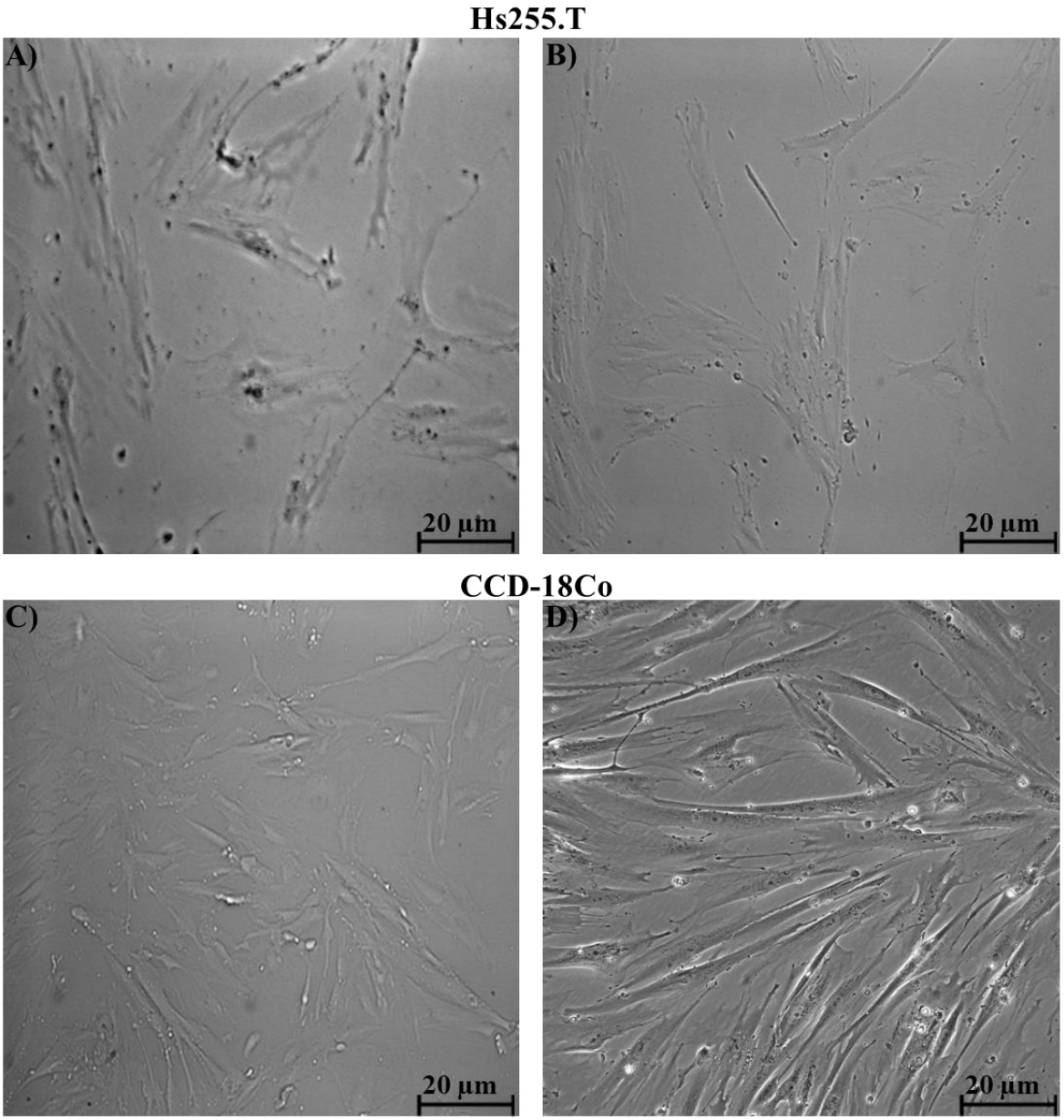


Figure 5.7: Human intestinal derived fibroblast cell lines in culture at 37^0C 5% CO_2. A-B) shows Hs255.T fibroblast cell line from an adenocarcinoma donor at passage 2, cells were cultured in complete medium (10% FBS+DMEM+1% Penicillin-Streptomycin) **C-D)** shows CCD-18Co intestinal derived fibroblast cell line from a normal donor at passage 6, cells were cultured in complete medium (10% FBS+EMEM+1%Penicillin-Streptomycin). Images were taken at X20 magnification using a Zeiss Axio inverted microscope (bright field settings)

5.3.4 Characterisation of CCD-18Co

5.3.4.1 CCD-18Co express surface markers of a fibroblast like phenotype

To characterise CCD-18Co as an in vitro cell culture model to study the role of gremlin 1 in intestinal fibrosis. CCD-18Co cells in culture were stained for expression of biomarkers that identify cells of mesenchymal origin specifically fibroblasts (543). Using flow cytometry, CCD-18Co were detached into a single cell suspension and labelled with fluorochrome conjugated antibodies to identify expression of surface markers expressed by fibroblasts. These included CD34 as a marker of mesenchymal origin, CD90 (Thy1), PDGFR-β, and Podoplanin. These markers are predominantly expressed on the surface of fibroblasts and are used in combination for fibroblast phenotyping. The flow cytometry panel also included CD45 which is a marker specific to immune cells, EpCAM known to identify Epithelial cells and CD31 which is a known endothelial cell marker. The panel included markers to validate CCD-18Co as a homogenous fibroblast cell population and not contaminated with other cell types.

Using flow cytometry, CCD-18Co were found to be a homogenous population as shown in the plots in figure 5.8A. The absence of CD45, EpCAM and CD31 expression showed that CCD-18Co did not express markers of immune cell, epithelial and endothelial cell lineages thus affirming the purity of the cell line. Despite published literature that show expression of HLA-DR on primary fibroblasts (544), CCD-18Co did not express HLA-DR on their surface. CCD-18Co also did not express the mesenchymal marker CD34. Most surprising was the lack of expression of PDGFRβ contrary to that routinely observed for fibroblasts. CCD-18Co cell population had a homogenous expression of CD90 and podoplanin (plots in figure 5.8B).

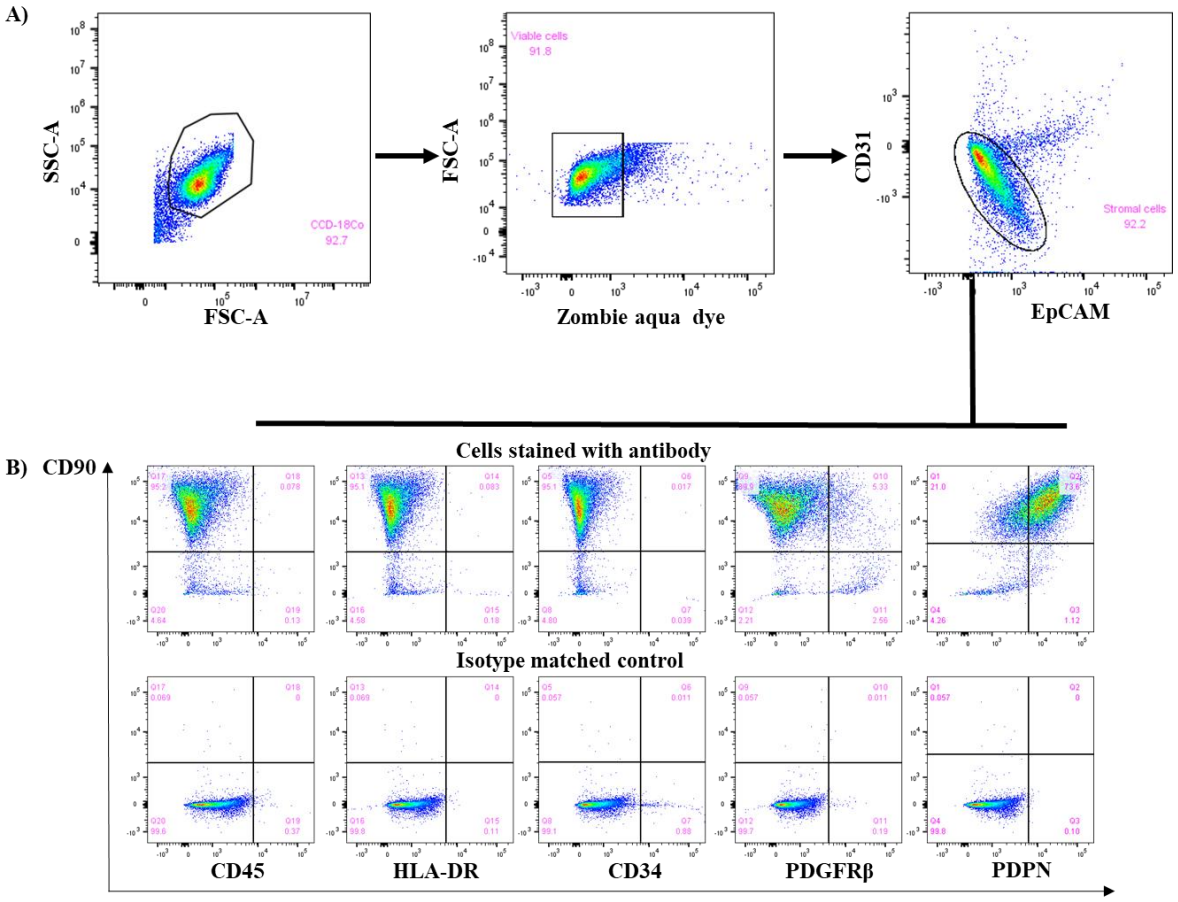


Figure 5.8: Flow cytometry gating to show identification of fibroblast like phenotype in CCD-18Co A) Side scatter area versus forward scatter area plot was used to select the cell population, live cells were selected based on Zombie Aqua (live/dead) dye uptake, then stromal cells defined as $CD31^{-ve}$ $EpCAM^{-ve}$ cells respectively. **B)** CCD-18Co were gated for expression of fibroblast like markers, fibroblasts are defined as $CD45^{-ve}CD90^{+ve}HLA\text{-}DR^{-ve}CD34^{-/+}PDGFR\beta^{-/+}$ $PDPN^{+ve}$.

5.3.4.2 CCD-18Co express surface markers of a fibroblast like phenotype

The flow cytometry graphs shown in figure 5.8 showed the expression of some fibroblast markers by CCD-18Co especially CD90 and podoplanin. To further verify the phenotype of CCD-18Co, cells were labelled with antibodies to identify expression of surface markers and intermediate filament proteins, these included vimentin (type III intermediate filament) expressed on stromal cells, α-SMA which is a marker of activated fibroblasts, CD90 and PDGFRβ, their expression pattern was studied by multiplex immunofluorescence imaging. Immunofluorescence was used to study the expression pattern, cellular distribution, and co-

expression of the different stromal cell markers in CCD18Co. As seen in figure 5.9, vimentin which is an intermediate filament protein expressed on mesenchymal cells was expressed on CCD-18Co showing intracellular staining within the intermediate filaments. There was positive staining for both CD90 and PDGFRβ, with the protein expression pattern for CD90 observed within the nucleoplasm and plasma membrane whereas PDGFRβ was cytoplasmic. CCD-18Co also showed expression of α-SMA which stained the actin filaments. Overall, there was co-expression of vimentin, CD90, PDGFRβ and α-SMA on CCD-18Co. This protein expression pattern agrees with the flow cytometry data and confirms the identity of CCD-18Co as fibroblasts. As a control measure, isotype matched controls were included for all the antibodies that were used to determine the presence of non-specific binding and background effect.

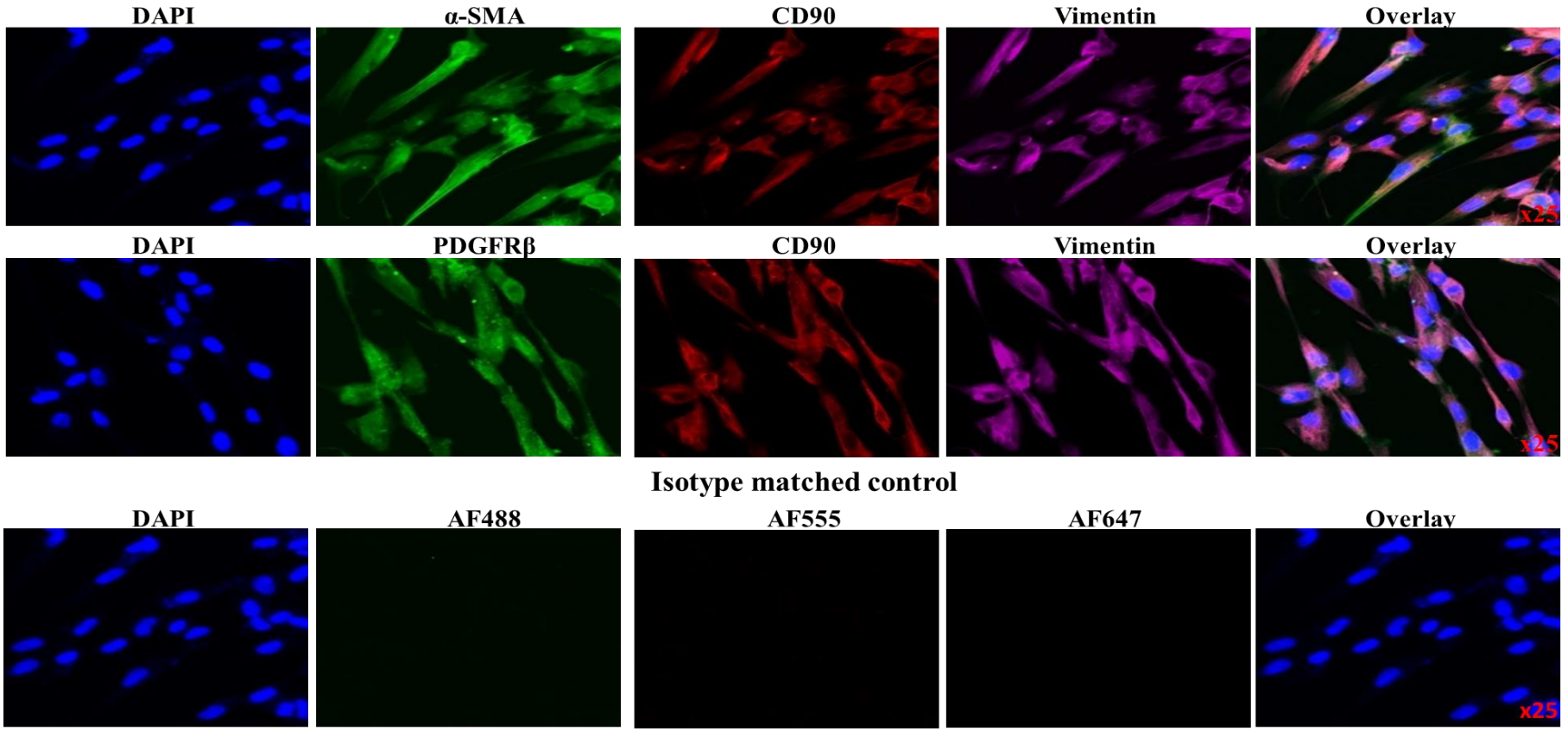


Figure 5.9: CCD-18Co expressed α-SMA, PDGFRβ, CD90 and vimentin. CCD-18Co were plated in IBIDI slides and multiplex staining for fibroblast markers CD90 visualised on the Alexa Fluor(AF)555 channel (red), vimentin visualised on the AF647 channel (purple), PDGFRβ and α-SMA visualised on the AF488 channel (green), Nucleus visualised on the DAPI channel (blue). **Row 1** contains nucleus (DAPI), α-SMA, CD90, vimentin and the overlay of the 3 markers; **Row 2** contains nucleus, PDGFRβ, CD90, vimentin and the overlay of the 3 markers; **Row 3** contains isotype controls (Mouse IgG1a, Mouse IgG2a and Rabbit). Images were taken at a magnification of X25 on Zeiss LSM 780.

5.3.4.3 CCD-18Co respond to profibrotic factors and express ECM related genes

To investigate the response of CCD-18Co to profibrotic stimuli in vitro, cells were stimulated with known profibrotic factors, these included TGF-β, IL-13 and IL-36 (545,546). CCD-18Co were stimulated at 1ng/mL, 10ng/mL and 100ng/mL for 24hrs. Stimulated cells were lysed and RT-qPCR was performed to measure mRNA levels of *COL1A1, ACTA2, FN1* and *TIMP1*. Results from 3 repeat experiments seen in figure 5.10 showed that while stimulation of CCD-18Co with IL-13 had no effect on expression of FN1 and TIMP1, there was a significant upregulation in expression of ACTA2 (p=0.007**) at 10ng/ml and a downregulation effect in expression of COL1A1 (p=0.007**) at 1ng/ml. Stimulation with IL-36 on the other hand showed a different effect to IL-13, there was no significant effect in expression of FN1 and ACTA2, although there was a downregulation effect in expression of COL1A1 (p=0.008**) and TIMP1 (p=0.02*) respectively.

As published in the literature, TGF-β is a potent profibrotic factor as seen in its effect on CCD-18Co in figure 5.10. There was significant upregulation in expression of *COL1A1, ACTA2, FN1* and *TIMP1* at all concentrations except at the 10ng/ml concentration where although COL1A1 was increased compared to the unstimulated cells, it was not statistically significant (figure 5.10A). This data showed that CCD-18Co respond to stimulation by profibrotic factors but most especially to TGF-β which is well studied as a key profibrotic factor in various fibrotic diseases. The experiment showed TGF-β profibrotic effect on CCD-18Co at 24 hr timepoint with significant effect of profibrotic genes including COL1A1, FN1, TIMP1 and ACTA2. The next step was to assess effect at both 24 and 48 hrs to assess any increase in the assay window.

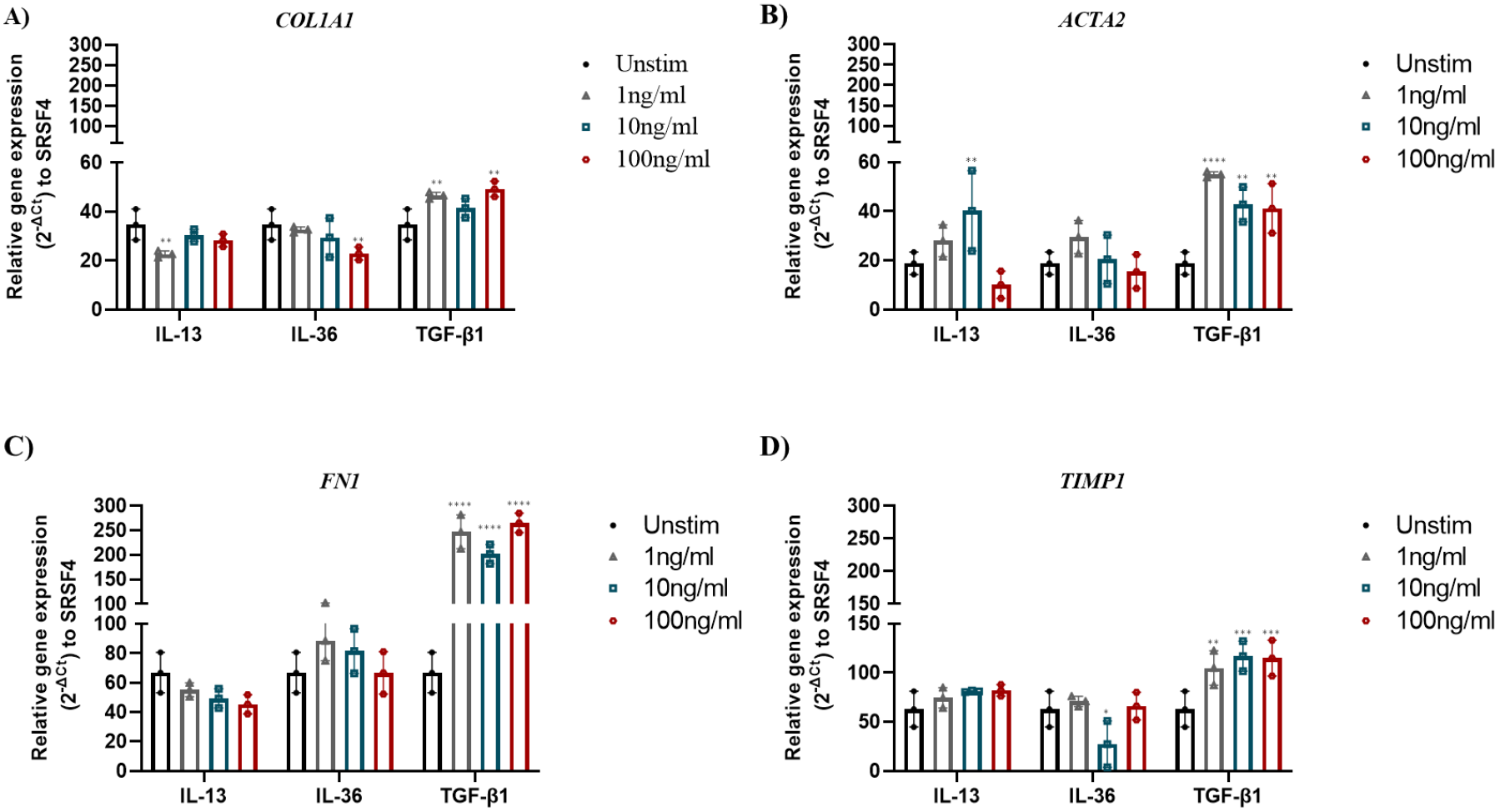

A)
COL1A1
Relative gene expression
(2-ΔCt) to SRSF4
300
250
200
150
100
60
40
20
0
IL-13
IL-36
TGF-β1
Unstim
1ng/ml
10ng/ml
100ng/ml
B)
ACTA2
Relative gene expression
(2-ΔCt) to SRSF4
300
250
200
150
100
60
40
20
0
IL-13
IL-36
TGF-β1
Unstim
1ng/ml
10ng/ml
100ng/ml
C)
FN1
Relative gene expression
(2-ΔCt) to SRSF4
300
250
200
150
100
100
80
60
40
20
0
IL-13
IL-36
TGF-β1
Unstim
1ng/ml
10ng/ml
100ng/ml
D)
TIMP1
Relative gene expression
(2-ΔCt) to SRSF4
300
250
200
150
150
100
50
0
IL-13
IL-36
TGF-β1
Unstim
1ng/ml
10ng/ml
100ng/ml

Figure 5.10: CCD-18Co respond to profibrotic mediator stimulation with increased mRNA expression of ECM related genes. CCD-18Co were cultured in a 24 well plate at $5x10^4$ cells/ml up to 70% confluence in 10% FCS+EMEM (media). Cells were serum starved for 12hrs in 0.2% BSA+EMEM and stimulated with recombinant cytokines (TGF-β, IL-13 and IL-36) at 1ng/ml, 10ng/ml, 100ng/ml and cultured for 24hrs at 37^0C and 5% CO_2. RNA was extracted, cDNA was produced, and qPCR performed. Data are presented as relative gene expression to endogenous control *SRSF4*, $2^{-\Delta Ct}$ of *COL1A1*, *TIMP1*, *ACTA2* and *FN1*. **A)** Plot shows relative gene expression of *COL1A1* in response to stimulation by IL-13, IL-36 and TGF-β at 1, 10, 100 ng/ml respectively **B)** Plot shows relative gene expression of *ACTA2* in response to stimulation by IL-13, IL-36 and TGF-β at 1, 10, 100 ng/ml respectively **C)** Plot shows relative gene expression of *FN1* in response to stimulation by IL-13, IL-36 and TGF-β at 1, 10, 100 ng/ml respectively **D)** Plot shows relative gene expression of *TIMP1* in response to stimulation by IL-13, IL-36 and TGF-β at 1, 10, 100 ng/ml respectively. Gene expression in stimulated cells was compared relative to unstimulated cells; *SRSF4* was used as the endogenous control. Data represented as median with 95% CI; statistics performed using 2way ANOVA multiple comparison (Tukey) performed.

5.3.4.4 Increase in ECM related gene expression with TGF- β stimulation at 48 hours

CCD-18Co response to stimulation by profibrotic factors was assessed and confirmed in the previous section. A subsequent experiment was performed to investigate changes in expression of COL1A1, ACTA2, FN1 and TIMP1 in CCD-18Co for a later time point. The experimental set up was similar to the previous section. However, for this experiment, stimulation was performed using TGF-β only at a wide range of concentrations including (1ng/ml, 10ng/ml, 50ng/ml, 100ng/ml and 200ng/ml) and gene expression measured at 24 and 48 hrs. As seen in figure 5.11, overall gene expression of COL1A1, ACTA2, FN1 and TIMP1 relative to the house keeping gene was upregulated at 48hr time point as compared to 24hr. Expression of COL1A1 by CCD-18Co was statistically significant at the 48hr time point for all concentrations with p values (1ng/ml; p value=0.03*), the higher concentrations showed a statistical significance with a p value <0.0001****. There was an increase in expression of ACTA2 at the 48hr time point although it was not statistically significant. Expression of FN1

by CCD-18Co was statistically significant at the 48hr time point at the 1ng/ml concentration (p value=0.004**) and at 10ng/ml (p value= <0.0001****). Expression of TIMP1 was increased significantly at the 48hr time point for all conditions with p values of 0.005** for the unstimulated CCD-18Co and <0.0001**** for stimulated cells. This data showed an overall increase in gene expression at 48hr time point and a sustained ECM related gene expression and activated cellular state of the CCD-18Co cells, which is a similar gene expression pattern for activated fibroblasts. The TGF-β effect seen in this experiment in which mRNA levels of *COL1A1*, *FN1* and *TIMP1* are higher at 10ng/ml versus the top concentration of 200ng/ml shows a hermetic effect. Hormesis is biphasic dose-response that displays low concentration stimulation and a high concentration inhibitory effect. Within the hormetic zone, the biological response to low concentration exposure to TGF-β achieved the desired effect of upregulation of downstream target genes on the contrary, the top concentration of TGF-β at 200ng/ml showed an inhibitory effect that could be attributed to toxicity of the stimuli (547,548).

Another point to note is that figure 5.11B shows no major changes to expression of *ACTA2* across all concentration in at 24 and 48 hrs respectively. It presents the argument if the 48 hr time point is enough to induce fibroblast to myofibroblast transition (FMT). A later timepoint of 72 hrs could have been tested to assess FMT as suggested by (549) in the literature but this was not tested, it is a limitation to the study. Overall, the ECM gene expression data obtained from the above experiments indicated an active phenotype of CCD-18Co upon stimulation with profibrotic mediators particularly TGF- β, the next step was to assess CCD-18Co proliferative capacity.

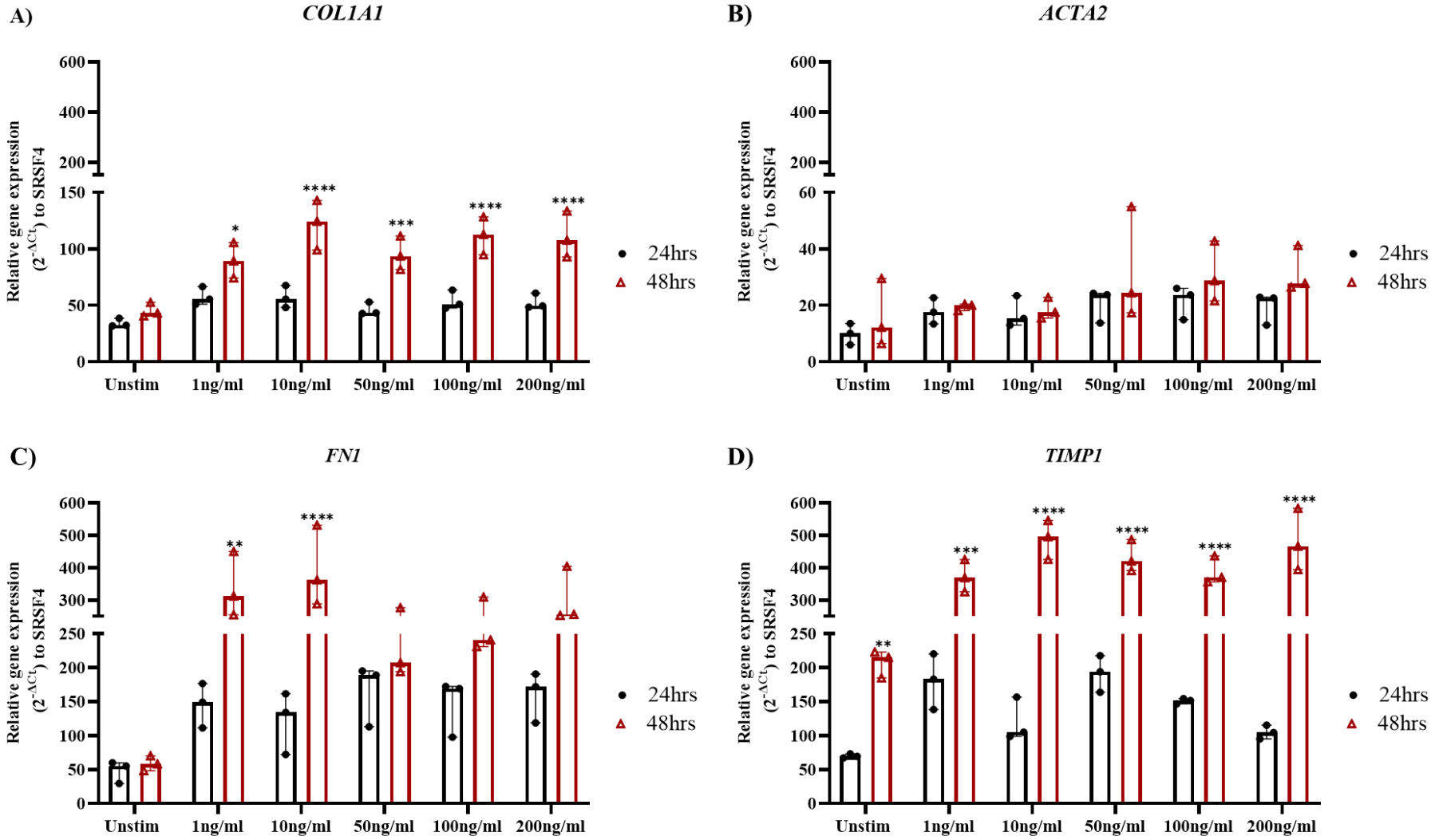

A)
COL1A1
Relative gene expression
($2^{-\Delta Ct}$) to SRSF4
600
400
200
150
100
50
0
*
****
***
****
****
Unstim
1ng/ml
10ng/ml
50ng/ml
100ng/ml
200ng/ml
24hrs
48hrs
B)
ACTA2
Relative gene expression
($2^{-\Delta Ct}$) to SRSF4
600
400
200
60
40
20
0
Unstim
1ng/ml
10ng/ml
50ng/ml
100ng/ml
200ng/ml
24hrs
48hrs
C)
FN1
Relative gene expression
($2^{-\Delta Ct}$) to SRSF4
600
500
400
300
250
200
150
100
50
0
**
****
Unstim
1ng/ml
10ng/ml
50ng/ml
100ng/ml
200ng/ml
24hrs
48hrs
D)
TIMP1
Relative gene expression
($2^{-\Delta Ct}$) to SRSF4
600
500
400
300
250
200
150
100
50
0
**
***
****
****
****
****
Unstim
1ng/ml
10ng/ml
50ng/ml
100ng/ml
200ng/ml
24hrs
48hrs

Figure 5.11: CCD-18Co have a higher expression of ECM related genes at 48hrs upon stimulation with TGF-β. CCD-18Co were cultured in a 24 well plate at $5x10^4$ cells/ml up to 70% confluence in 10% FCS+EMEM (media). Cells were serum starved for 12hrs in 0.2% BSA+EMEM and stimulated with TGF-β at 1ng/ml, 10ng/ml, 50ng/ml 100ng/ml and 200ng/ml and cultured for 24 and 48 hrs respectively at 37^0C and 5% CO_2. RNA was extracted, cDNA was produced and qPCR was performed. Data are presented as relative gene expression to endogenous control *SRSF4,* $2^{-\Delta Ct}$ of *COL1A1, TIMP1, ACTA2* and *FN1*. **A)** Plot shows relative gene expression of *COL1A1* in response to stimulation by TGF-β at 1, 10, 50, 100, 200 ng/ml respectively **B)** Plot shows relative gene expression of *ACTA2* in response to stimulation by TGF-β at 1, 10, 50, 100, 200 ng/ml respectively **C)** Plot shows relative gene expression of *FN1* in response to stimulation by TGF-β at 1, 10, 50, 100, 200 ng/ml respectively **D)** Plot shows relative gene expression of *TIMP1* in response to stimulation by TGF-β at 1, 10, 50, 100, 200 ng/ml respectively. Gene expression at 24 hrs was compared to 48 hrs; *SRSF4* was used as the endogenous control. Data represented as median with 95% CI, n=3, 2way ANOVA multiple comparison (Sidak) performed.

5.3.4.5 CCD-18Co proliferation increases upon stimulation with profibrotic factors

Profibrotic factors in a fibrotic microenvironment modulate fibroblast proliferation, cellular proliferation is one of the key mechanisms through which fibroblasts drive pathogenesis and formation of strictures in CD (308). This assay was aimed at testing the effect of profibrotic factors on the proliferative capacity of CCD-18Co. Cells were seeded in 96 well plates at a low and high density of 6,250 and 25,000 cells per well respectively. CCD-18Co were stimulated with IL-1β, PDGF-BB and TGF-β at 100ng/ml for 20, 36 and 48 hrs in culture at 37°C/ 5%CO2. A concentration of 100ng/ml concentration was chosen based on previous experiments where this concentration of TGF-β elicited the greatest effect on target gene expression. Cell proliferation was assessed using the MTT assay, in which metabolically active cells (thus proliferating cells) reduced tetrazolium salt (substrate) in the MTS reagent to form a dark coloured formazan dye.

Cell proliferation was measured after 4 hrs of adding the MTT reagent onto cells at the 20, 36 and 48 hr time points respectively, by use of an ELISA plate reader, the absorbance was measured at an optical density of 490nm. Data analysis was performed using linear regression analysis in which case the R2 values were determined all of which were >0.97 meaning there was a correlation to the proliferative effect across the different stimulations. The data showed in figure 5.12 suggests increased proliferation of CCD-18Co upon stimulation with all the three profibrotic mediators (IL-1β, PDGF-BB and TGF-β). The proliferative effect seen on CCD-18Co was time dependent with increase in proliferation most noticeable at 48hrs. The data aligned with what has already been published in the literature on the increase in proliferation of fibroblasts in a profibrotic microenvironment (425). It is therefore confirmation of a similar behavioural pattern between primary fibroblasts and CCD-18Co.

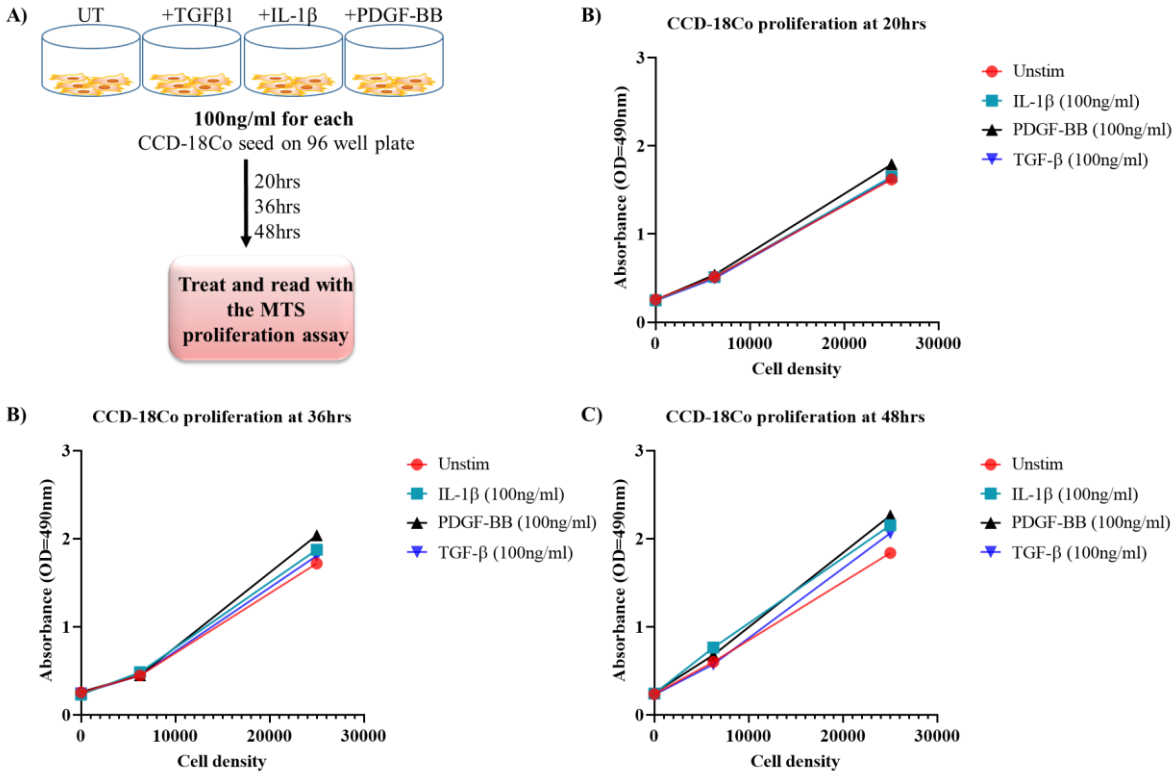


Figure 5.12: Increase in proliferation of CCD-18Co in the presence of profibrotic mediators

A) Experimental set up of the proliferation assay, cells were seeded at different cell densities (0, 6250, 25000 cells/200µl per well) and stimulated as; (unstim- no stimulant, CCD-18Co+IL-1β at 100ng/ml; CCD-18Co+ PDGF-BB at 100ng/ml; CCD-18Co+ TGF-β at 100ng/ml). Cells were then cultured at 37°C/ 5%CO_2 over 20hrs, 36hrs and 48hrs. MTS reagent was added for 4 hrs and absorbance read at OD 490nM using an ELISA plate reader. Graphs show CCD-18Co proliferation at the different time points. **B)** proliferation at 20 hrs **C)** proliferation at 36 hrs **D)** proliferation at 48 hr time points respectively. Measurement was performed using the MTS assay, each point represents a mean of 3 replicates n=3. Three independent experiments performed in triplicate. Linear regression analysis was performed to determine the effect of profibrotic factors on CCD-18Co cell proliferation

5.3.4.6 Primary liver fibroblasts proliferate upon stimulation with profibrotic factors

To test the effect of profibrotic stimuli and to validate the proliferation assay, primary fibroblasts from explant primary sclerosing cholangitis (PSC) liver samples at passage 6 were stimulated with TGF-β and IL-1β at 100ng/ml in a 96 well plate. PSC primary fibroblasts were used due to the limited access to primary intestinal fibroblasts. Although it is different tissue types, fibroblast phenotypes are similar across tissue. Cells were cultured at 37^0C and CO2

over 20, 36 and 48 hrs respectively, the MTS reagent was added to determine the effect of profibrotic mediators on the proliferation of primary fibroblasts. Although statistical analysis was not performed on results from the proliferation assay as there were no repeat experiments done due to a limited access to primary fibroblasts, figure 5.13 shows a trend with increase in CCD-18Co proliferation at 48hrs. It also shows that stimulation of primary fibroblasts by profibrotic factors including IL-1β and TGFβ show an increase in proliferation of primary PSC fibroblasts. The trend seen in PSC primary fibroblasts give an insight on how profibrotic factors modulate proliferation, this could be translated to primary intestinal fibroblasts.

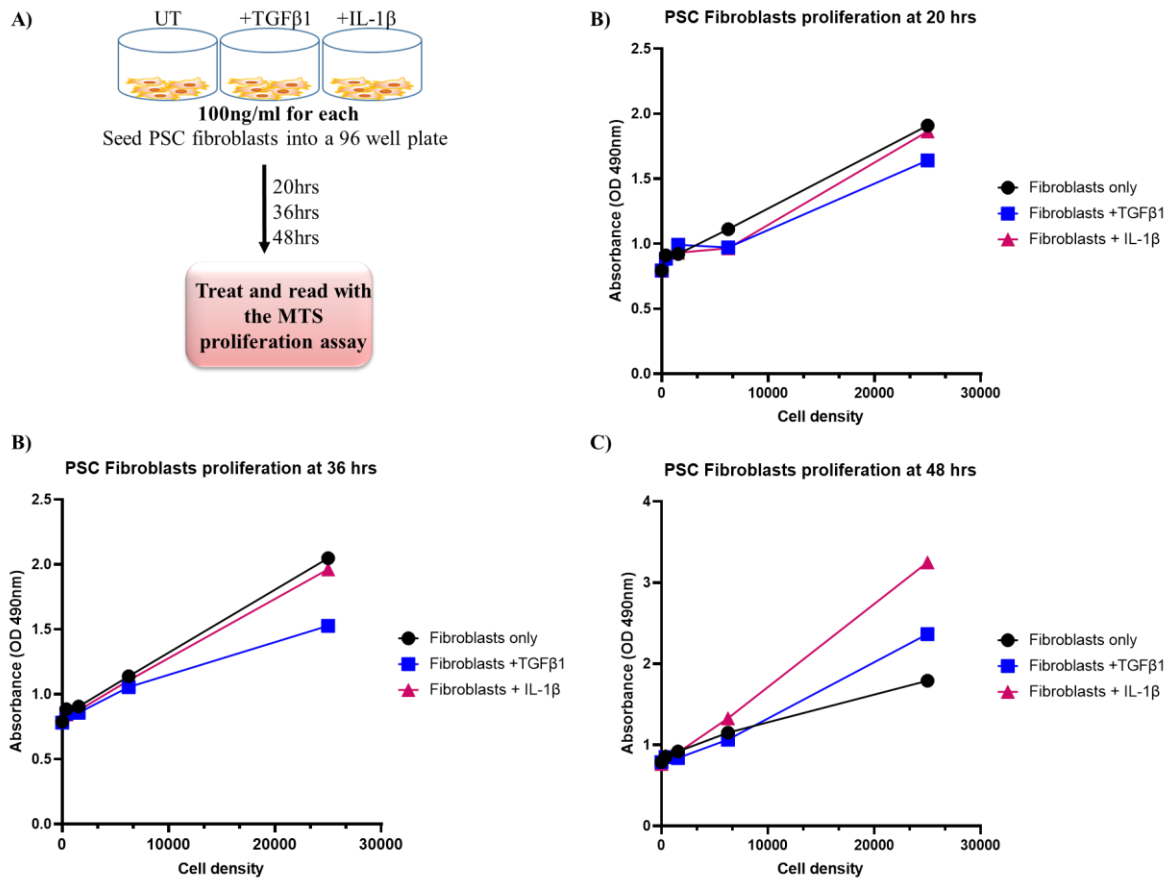


Figure 5.13: Presence of profibrotic mediators increase proliferation in primary fibroblasts from PSC explant liver. A) Cells were seeded at different cell densities (0, 390, 1562, 6250, 25000 cells/200µl per well) and stimulated as; (fibroblasts only- no stimulant, fibroblasts + TGF-β; fibroblasts + IL-1β) at 100ng/ml. Cells were cultured at 37°C/ 5% CO_2 over 20hrs, 36hrs and 48hrs. MTS reagent was added for 4 hrs and absorbance read at OD 490nM using an ELISA plate reader. **B-D)** Graphs show fibroblast proliferation at the 20hr, 36hr and 48hr time points respectively measured by MTS assay, each point represents a mean of 3 replicates n=1. Statistical analysis was not performed due to few repeats limited by available samples

5.3.4.7 CCD-18Co show gene expression of BMP receptors in vitro

The BMP pathway signals through BMPRIA, BMPRIB and BMPRII, and the associated signal transduction pathway is, in the most part, shared with other members of the TGF-β superfamily. To understand the dynamics in gremlin 1 regulation of the BMP/TGF-β signalling axis, it was important to establish expression levels of BMP receptors by CCD-18Co, dermal fibroblasts have been shown to express BMP receptors in a study performed by Fessing (2010). CCD-

18Co were stimulated with TGF-β at 1, 10, 50, 100 and 200 ng/ml and left in culture over 24 and 48 hrs respectively. The expression levels of *BMPRIA*, *BMPRIB* and *BMPRII* were analysed using RT-qPCR. Gene expression data shown in figure 5.14 indicated very low expression of the 3 receptors relative to the house keeping gene by CCD-18Co. There was no statistical significance in expression between stimulation at 24 and 48 hrs except for the 10ng/ml concentration effect on expression of *BMPRIA* shown in figure 5.14A (p value= <0.008**).

In summary, the data generated this for showed that CCD-18Co are phenotypically similar to primary fibroblasts. CCD-18Co are morphologically flat and spindle shaped which is a similar morphology to that of primary fibroblasts. CCD-18Co express CD90 (Thy1), PDGFRβ, vimentin and α-SMA, well documented stromal cell markers. The cells respond to profibrotic mediators with increased expression of profibrotic target genes upon stimulation with TGF-β predominantly at the 100 ng/ml concentration. CCD-18Co also showed expression of BMP receptors which are key molecules in BMP receptor signalling. These findings affirm the similarity in phenotype between CCD-18Co and primary fibroblasts, with challenges faced in obtaining gut primary fibroblasts, there was confidence in utilising CCD-18Co as an in vitro cell line model to study the role of gremlin 1 in regulating BMP/TGF-β signalling and how this contributes to formation of strictures in CD.

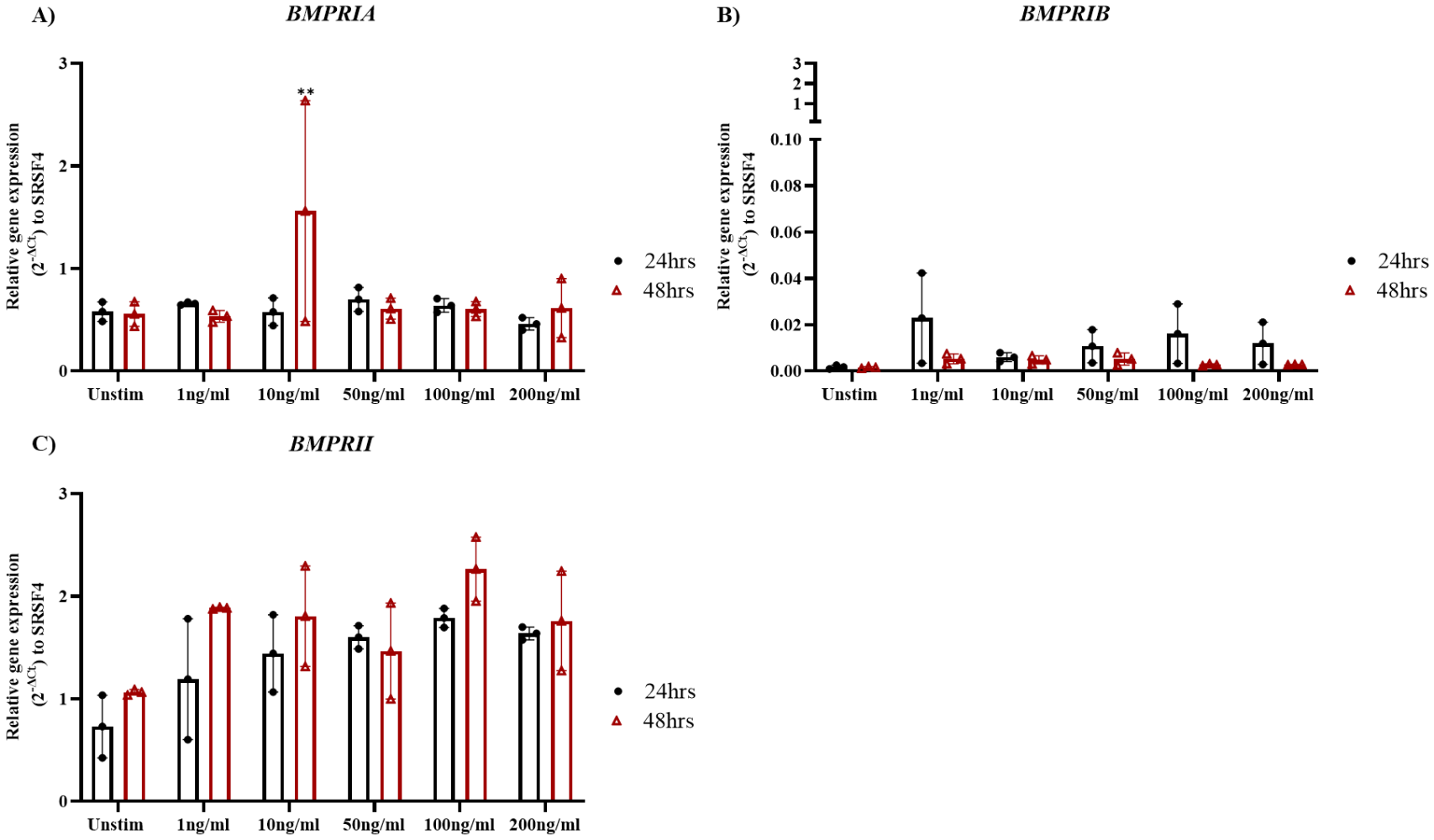

A)
BMPRIA
Relative gene expression
(2^-ΔCt) to SRSF4
**
Unstim
1ng/ml
10ng/ml
50ng/ml
100ng/ml
200ng/ml
24hrs
48hrs
B)
BMPRIB
C)
BMPRII

Figure 5.14: CCD-18Co have similar expression patterns of BMP receptors at 24 and 48 hrs. CCD-18Co were cultured in a 24 well plate at $5x10^4$ cells/ml up to 70% confluence in 10% FCS+EMEM (media). Cells were serum starved for 12hrs in 0.2% BSA+EMEM and stimulated with TGF-β at 1ng/ml, 10ng/ml, 50ng/ml 100ng/ml and 200ng/ml and cultured for 24 and 48 hrs respectively at 37^0C and 5% CO_2. RNA was extracted, cDNA was produced and RT-qPCR was performed. Data are presented as relative gene expression to endogenous control *SRSF4*, $2^{-\Delta Ct}$ of *BMPRIA, BMPRIB* and *BMPRII* **A)** Plot shows relative gene expression of *BMPRIA* in response to stimulation by TGF-β at 1, 10, 50, 100, 200 ng/ml respectively **B)** Plot shows relative gene expression of *BMPRIB* in response to stimulation by TGF-β at 1, 10, 50, 100, 200 ng/ml respectively **C)** Plot shows relative gene expression of *BMPRII* in response to stimulation by TGF-β at 1, 10, 50, 100, 200 ng/ml respectively. Gene expression at 24 hrs was compared to 48 hrs; SRSF4 was used as the endogenous control Data represented as median with 95% CI, n=3, 2way ANOVA multiple comparison (Sidak) performed.

5.3.5 Gremlin 1 and its role in BMP/ TGF-β signalling

Increased levels of gremlin 1 have been reported in multiple fibrotic conditions, the protein has been linked to mechanisms that promote fibroblast proliferation and migration (Yang et al., 2021), the next section covers studies performed to understand the role of gremlin 1 in fibrosis while using CCD-18Co as an in vitro model.

5.3.5.1 CCD-18 express *GREM1* at transcript level at different time points

To assess baseline expression of GREM1 mRNA, CCD-18Co at a dilution of 5x104 cells/ml were seeded in a 24 well plate and left in culture for 24 and 48 hrs respectively. Unstimulated cells were lysed, and RT-qPCR performed on cDNA to check for GREM1 mRNA expression. The data as represented in figure 5.15 shows that while CCD-18Co express GREM1 there was no significant difference in expression of GREM1 between cells left in culture for 24 or 48 hrs. There is no information in the literature to indicate mRNA expression levels at different time points.

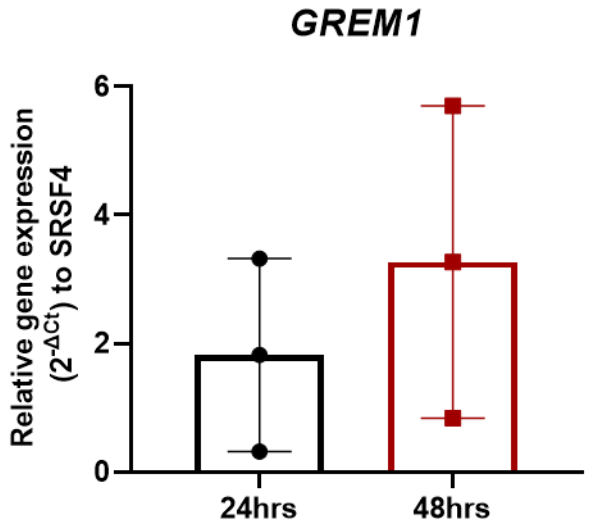


Figure 5.15: Unstimulated CCD-18Co express *GREM1* at transcript level. CCD-18Co were cultured in a 24 well plate at $5x10^4$ cells/ml and cultured for 24 and 48 hrs respectively at 37^0C and 5% CO_2. RNA was extracted, cDNA was produced and RT-qPCR performed. Data are presented as relative gene expression to endogenous control *SRSF4*, $2^{-\Delta Ct}$ of *GREM1*. Plot shows relative gene expression of *GREM1* in unstimulation CCD-18Co. *SRSF4* was used as the endogenous control. Data presented as median with 95% CI, n=3, t test (non- parametric tests) Mann-Whitney test was performed

5.3.5.2 CCD-18 express *GREM1* upon stimulation with TGF-β at transcript level

Having established *GREM1* expression at baseline, *GREM1* expression was assessed following stimulation with TGF-β. Cells were seeded, then stimulated with TGF-β at 1, 10, 20, 50, 100 and 200 ng/ml. Cells were left in culture for 24 and 48 hrs respectively, RNA was extracted and RT-qPCR performed on cDNA to check for *GREM1* expression. Figure 5.16 shows the levels of *GREM1* at 24 and 48 hr time points post TGF-β stimulation at different concentrations. The highest *GREM1* expression is seen at 24 hrs with a 100 ng/ml TGF-β stimulation. The lowest expression level of gremlin 1 is seen at the highest concentration of 200ng/ml, this could be explained by increased toxicity of TGF-β at the highest concentration. This could have been confirmed by performing a toxilight assay to assess toxicity; however, this was not performed at the time of the study. The TGF-β effect across the different concentrations is not significant due to the variability in the data as seen in the error bars. These

data showed transcription of GREM1 in CCD-18Co both at baseline and upon stimulation with TGF-β. The next step was to assess translation by assessing protein levels of gremlin 1. One of the main challenges with detection of gremlin 1 at protein level is the lack of reliable commercially available antibodies that can be used for protein detection using techniques such as western blot. Gremlin 1 protein secretion levels were measured in supernatants collected from CCD-18Co cell cultures using ELISA.

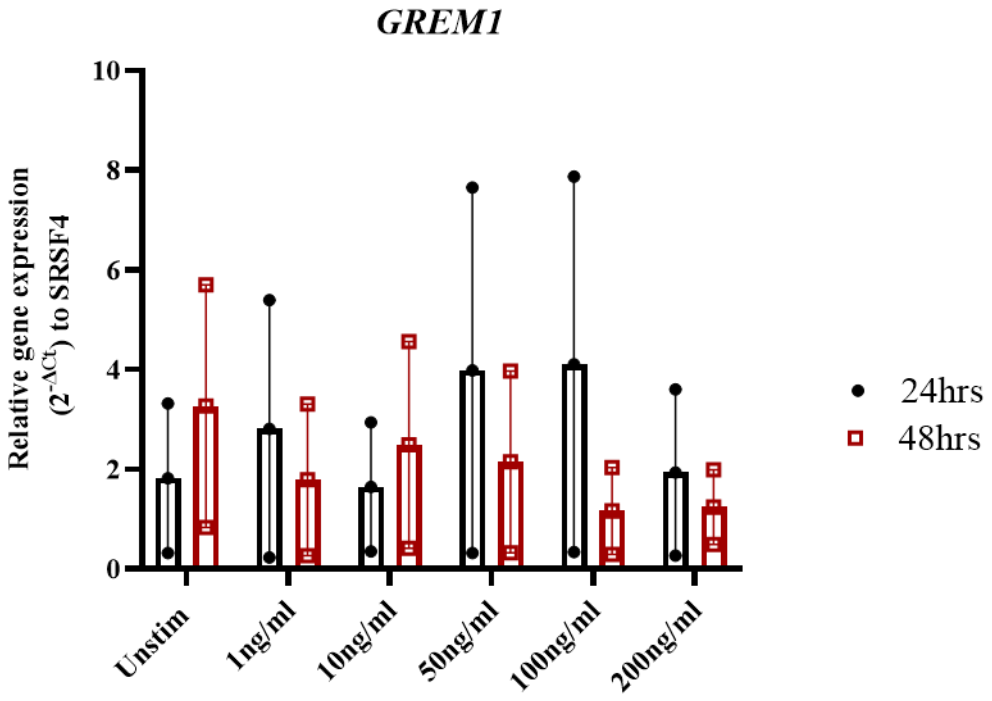


Figure 5.16: CCD-18Co express *GREM1* in response to TGF-β stimulation at 24hrs. CCD-18Co were cultured in a 24 well plate at $5x10^4$ cells/ml up to 70% confluence in 10% FCS+EMEM (media). Cells were serum starved for 12hrs in 0.2% BSA+EMEM and stimulated with TGF-β1 at 1ng/ml, 10ng/ml, 50ng/ml 100ng/ml and 200ng/ml and cultured for 24 and 48 hrs respectively at 37^0C and 5% CO_2. RNA was extracted and RT-qPCR was performed. Data was presented as relative gene expression to endogenous control SRSF4, $2^{-\Delta Ct}$ of *GREM1*. Plot shows relative gene expression of *GREM1* in response to stimulation at 24 and 48 hr time course. Gene expression at 24 hrs was compared to 48 hrs; *SRSF4* was used as the endogenous control. Data presented as median with 95% CI, n=3, 2-way ANOVA multiple comparison (Kruskal-Wallis test) performed

5.3.5.3 Increased levels of gremlin 1 protein secreted in supernatants of CCD-18Co with TGF-β stimulation at 72hrs

To determine translation of GREM1 by CCD-18Co to protein and to check for extracellular gremlin 1 secretion, CCD-18Co were seeded and stimulated with profibrotic factors including TGF-β at 1ng/ml, 10ng/ml and 50ng/ml in parallel with Plasminogen activator inhibitor-1 (PAI-1) and fibroblast growth factor (FGF2) at 50ng/ml. TGF-β and PAI-1 were used as they have been well documented as potent profibrotic factor (551) while FGF2 is reported as an inhibitor of profibrotic mechanisms (552). CCD-18Co were cultured for 24, 48 and 72 hrs respectively, and supernatants were collected for gremlin 1 ELISA assay as summarised in the experimental setup (figure 5.17A).

Data generated from the ELISA assay in figure 5.17B showed low levels of gremlin 1 in supernatants of unstimulated CCD-18Co, which was similar to the level of protein seen in CCD-18Co stimulated with 50ng/ml of FGF2, with a slight increase in protein levels between 24 and 48 hrs but a decrease towards 72 hr time points. On the other hand, PAI-1 stimulated cells had a high expression of gremlin 1 at 24 hrs, however there was a gradual decrease in protein detected at 48 hrs and even further at 72 hrs. There was an overall increase in gremlin 1 expression in TGF-β stimulated CCD-18Co, protein levels in supernatants increased at 48 hrs and furthermore at 72 hrs for all concentrations. The highest increment in protein expression was seen in supernatant of cells stimulated with TGF-β at 50ng/ml. Statistical analysis was not performed as there were no repeats performed for this experiment due to a lack of supply of ELISA plates that was a direct consequence of the pandemic (COVID). This protein expression data is an added validation and confirmation of mRNA expression data in figure 5.16 above, it shows that CCD-18Co respond to potent profibrotic factors such as TGF-β to express *GREM1* at transcript level, can translate and secrete gremlin 1 into extracellular spaces.

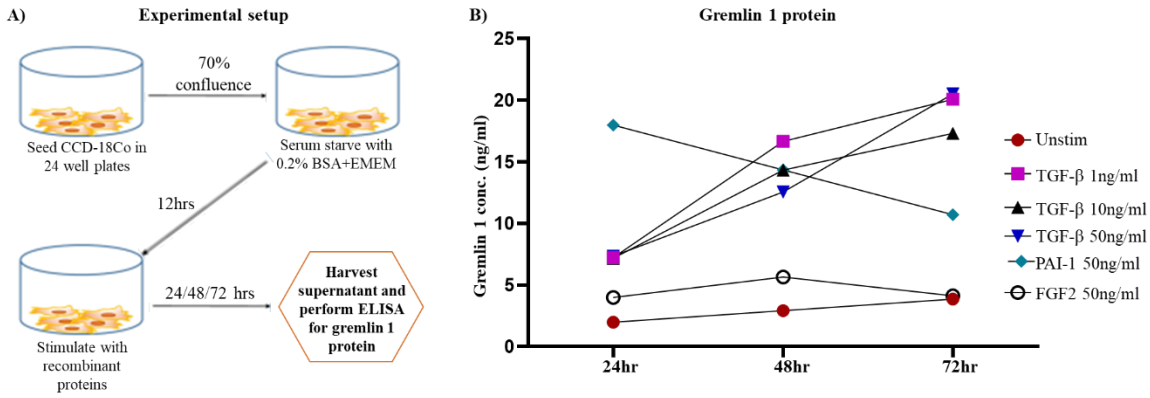


Figure 5.17: CCD-18Co secrete gremlin 1 upon stimulation with profibrotic mediators.

A) Summary of the experimental set up, CCD-18Co were seeded at $5x10^4$ cells/ml per well in 24 well plates until 70% confluence, cells were starved of serum by replacing with 0.2% BSA+ EMEM media. Cells were then stimulated with TGF-β (1ng/ml, 10ng/ml and 50ng/ml), PAI-1 at 50ng/ml and FGF2 at 50ng/ml, unstimulated cells were included. Cells were left in culture at 37^0C; 5% CO_2 and supernatants were collected at 24, 48 and 72 hrs. Gremlin 1 protein levels were determined through an ELISA assay. **B)** The graph displays gremlin 1 protein expression levels in supernatants collected from unstimulated and stimulated CCD-18Co at 24, 48 and 72 hr time points. Each condition was processed and analysed in triplicate n=1

The data presented in the sections above shows expression of gremlin 1 both at transcript and protein level by CCD-18Co. To explore the regulatory role of gremlin 1 further based on the observation that *GREM1* expression was upregulated in the stricture in our bulk RNA-seq data, with expression deriving predominantly from stromal cells as seen in the scRNA-seq data (chapter 3). There have also been several studies that suggest that gremlin 1 plays a key role in organ fibrosis (489,553) . The underlying mechanisms however are not well understood especially in intestinal fibrosis and therefore the next section explores over expression of *GREM1* in CCD-18Co in vitro and seeks to determine how this may interfere with signalling downstream the BMP/TGF-β axis.

5.3.6 Over expression of GREM1 in CCD-18Co using lentivirus 3-vector protocol

To generate GREM1 over expressing CCD-18Co cell line, gene transduction was performed using a lentiviral approach. Cloning using lentivirus was an already established method in the lab (Weston group), and a member of the research team performed the cloning process. A brief outline of the cloning process and production of virus particles is outlined in figure 5.18. The GREM1 sequence was inserted into the pWPI-Puro plasmid by means of a Pmel restriction site. As the pWPI-Puro plasmid also encodes green fluorescent protein (GFP) from a separate promoter, we were able to monitor transduction efficiency of both GREM1 and the empty vector control by GFP fluorescence.

In addition to the vector plasmids, two viral packaging plasmids were included as is the case with third generation manufacture of lentiviruses; one encodes for Rev element while the other encodes for gag and pol. These lentivirus packaging plasmids improved the efficiency of transduction and transgene expression; Rev element was important in export of RNA to the cytoplasm for virion packaging while the gag-pol element is important in virion assembly by mediating protein-protein interaction leading to formation of spherical particles, virus plasma membrane binding and recruitment of envelop proteins used for packaging genomic RNA (554,555). The three plasmids (vector plasmid and 2 packaging plasmids) were inoculated into E.*Coli* to allow viral assembly and formation of multiple clones, viral DNA was then collected, purified and sequenced to ensure that the transgene adequately integrated into the viral DNA. Once the viral sequence was confirmed, plasmid DNA was transfected into HEK293T cells to allow for expansion of the viral particles containing GREM1 transgene and control. Transfected HEK293T cells were lysed and the virion particles purified and stores at -800C ready for used in downstream experiments involving overexpression of GREM1.

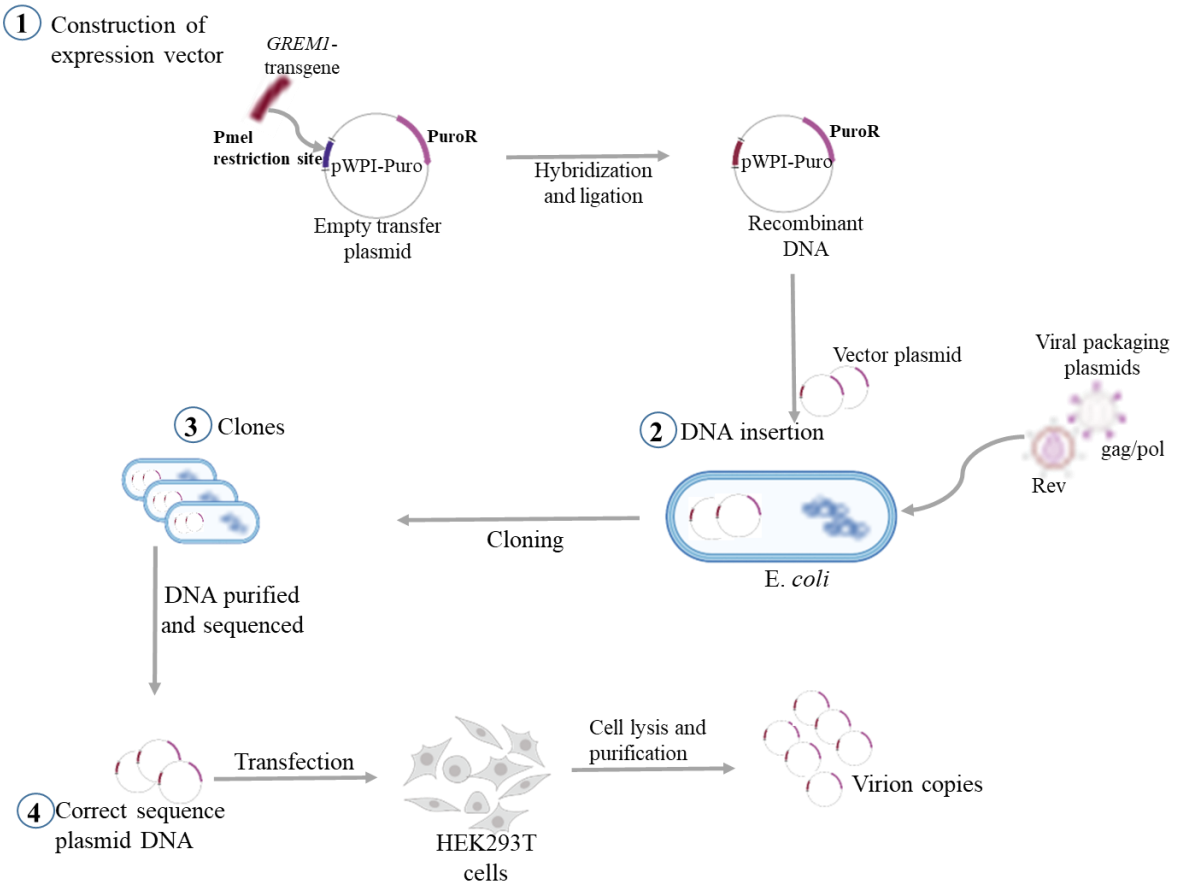


Figure 5.18: Preparation of GREM1-GFP tagged and control lentiviral particles.

Step 1: construction of expression vector, Pmel restriction enzyme cleaves the Pmel restriction site when the *GREM1* transgene or control anneal into the empty transfer plasmid, the single stranded DNA molecule with the new transgene (*GREM1* or control) underwent a process of hybridization and ligation to form a stable double stranded DNA molecule in the vector plasmid. **Step 2:** Vector plasmid and viral packaging plasmids were inserted into *E.coli* for viral assembly and manufacture of clones. **Step 3:** *E.coli* is lysed and plasmid DNA was collected, purified and sequenced to ensure the correct insertion of *GREM1* **Step 4:** plasmid DNA with the correct target sequence of *GREM1* was then transfected into HEK293T cells to generate more copies of virion particles for both over expression and control.

5.3.6.1 Efficiency of lentiviral mediated GREM1 overexpression in CCD-18Co

To check transduction efficiency of lentivirus plasmids containing the inserted GREM1 transgene in CCD-18Co, infected cells (either GFP vector alone or GFP vector transduced with GREM1 encoding vector) in culture at 5 days post transduction were imaged under a fluorescent microscope for green fluorescent protein (GFP) expression which acts as a reporter

gene and an indicator of successful transduction. Green fluorescence was seen in both CCD-18Co-GFP-control (figure 5.19B) and CCD-18Co-GFP-GREM1 over expressors (figure 5.19C). To check for the quantity of GREM1 transcript product between CCD-18Co-GFP-GREM1 over expressors and control, semi-quantitative PCR was performed. Amplification was carried out at 20 and 30 cycles respectively. The PCR products were loaded onto a 2% agarose gel alongside a DNA ladder. As seen in figure 5.19D, the band size for the expected PCR product which was 567 base pairs was detected using the Chemidoc imaging system. The band intensity in DNA from GREM1 over expressing CCD-18Co appeared stronger than the control. The absolute intensity values obtained from ImageJ as seen in figure 5.19E further confirm the successful transduction and over expression of GREM1 in CCD-18Co.

In addition, GREM1 mRNA levels were determined by RT-qPCR, figure 5.19F shows relative expression level of GREM1 in CCD-18Co GREM1 over expressers, the data shows a 2-fold change increase in expression as compared to control cells. No major changes in expression for both GREM1 and GFP was seen in control cells, there is still expression of GREM1 most likely associated to endogenous GREM1.There was a noticeable upregulation in GREM1 expression in CCD-18Co cells transduced with GREM1 plasmid.

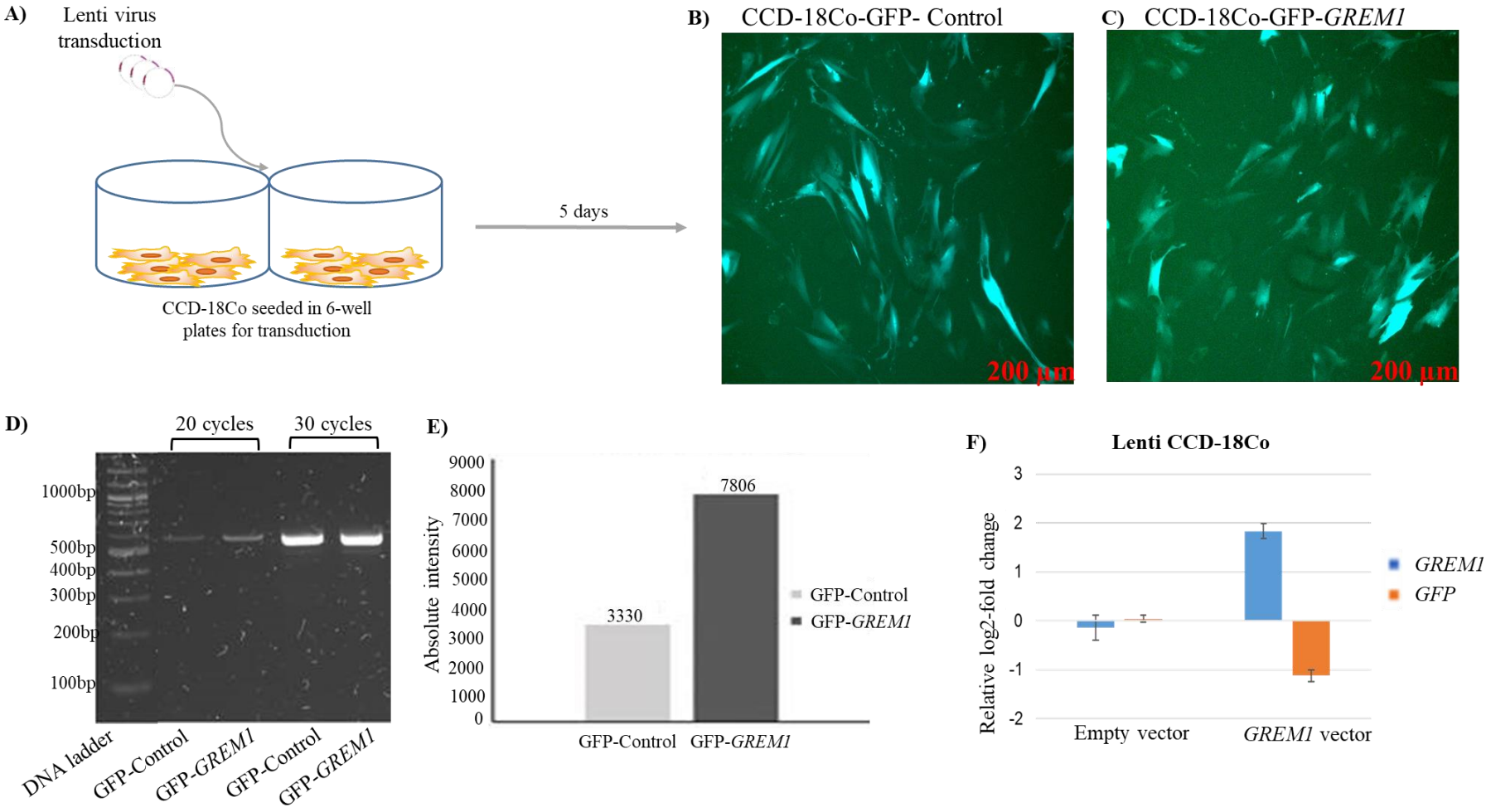

A)
Lenti virus transduction
CCD-18Co seeded in 6-well plates for transduction
5 days
B) CCD-18Co-GFP- Control
200 µm
C) CCD-18Co-GFP-GREM1
200 µm
D)
20 cycles
30 cycles
1000bp
500bp
400bp
300bp
200bp
100bp
DNA ladder
GFP-Control
GFP-GREM1
GFP-Control
GFP-GREM1
E)
Absolute intensity
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
3330
7806
GFP-Control
GFP-GREM1
GFP-Control
GFP-GREM1
F)
Lenti CCD-18Co
Relative log2-fold change
3
2
1
0
-1
-2
Empty vector
GREM1 vector
GREM1
GFP

Figure 5.19: Transduction efficiency of *GREM1* lentivirus particles into CCD-18Co. A) CCD-18Co were seeded at $4X10^5$ cells per well for 12 hrs; lentivirus with transgene insert (*GREM1* and an empty plasmid (control) both tagged with GFP were added into the respective wells (5µl per well), cells were left in culture at 37^0C 5%CO_2. Media containing the lentivirus was replaced after 48 hrs. **B)** Transduced CCD-18Co-GFP-control cells were imaged under a Zeiss inverted fluorescent microscope to assess transduction efficiency **C)** Transduced CCD-18Co-GFP-*GREM1* cells were imaged under a Zeiss inverted fluorescent microscope to assess transduction efficiency **D)** Representative 2% agarose gel image showing *GREM1* DNA amplified using semi-quantitative PCR at 20 and 30 cycles respectively, the bands appearing correspond to the expected PCR product of 567 base pairs for the *GREM1* sequence, imaging was performed on the Chemidoc imaging system **E)** Graph showing band intensities at 20 cycles for both the control and *GREM1* over expressing CCD-18Co **F)** Graphs shows relative expression as a fold change value in expression of both *GFP* and *GREM1* in over expressing cells and control; SYBR green fluorescent dye used for qPCR.

5.3.6.2 Sorting of GFP+ transduced CCD-18Co to improve purity

To improve purity and enrich GREM1 transduced CCD-18Co, fluorescence activated cell sorting (FACS) was performed on transduced cells based on their expression of GFP. Transduced cells in culture were trypsinised, washed in 2% FBS+PBS and resuspended into single cell suspensions in preparation for cell sorting. As seen in figure 5.20, gates were set up based on side scatter and forward scatter, this was performed on the Flow sorter (BD FACSAria), doublets were excluded based on forward scatter height and area parameters, GFP positive cells were then selected. CCD-18Co cells that were not transduced with GFP were used to set gates in order to select GFP positive cells (figure 5.20 Row A). Of the cells that were selected, 12% were GFP positive from the CCD-18Co- control samples (figure 5.20 Row B) while 9.3% were positive for the CCD-18Co-GREM1 sample (figure 5.20 Row C). The GFP positive sorted cells were washed and re-plated back in culture. Transduced cells adhered onto the plate; however, the cells did not proliferate any further when left in culture at 370C

CO2. The loss in proliferative capacity may be attributed to mechanical cell stress cellular stress as a result of cell sorter induced cell stress (SICS) given how large in cell size CCD-18Co are (556). The cell stress may have induced cellular senescence hence losing cell proliferative capacity. The loss in cell proliferative capacity meant there were insufficient numbers to perform any further experiments on the sorted GFP positive cells.

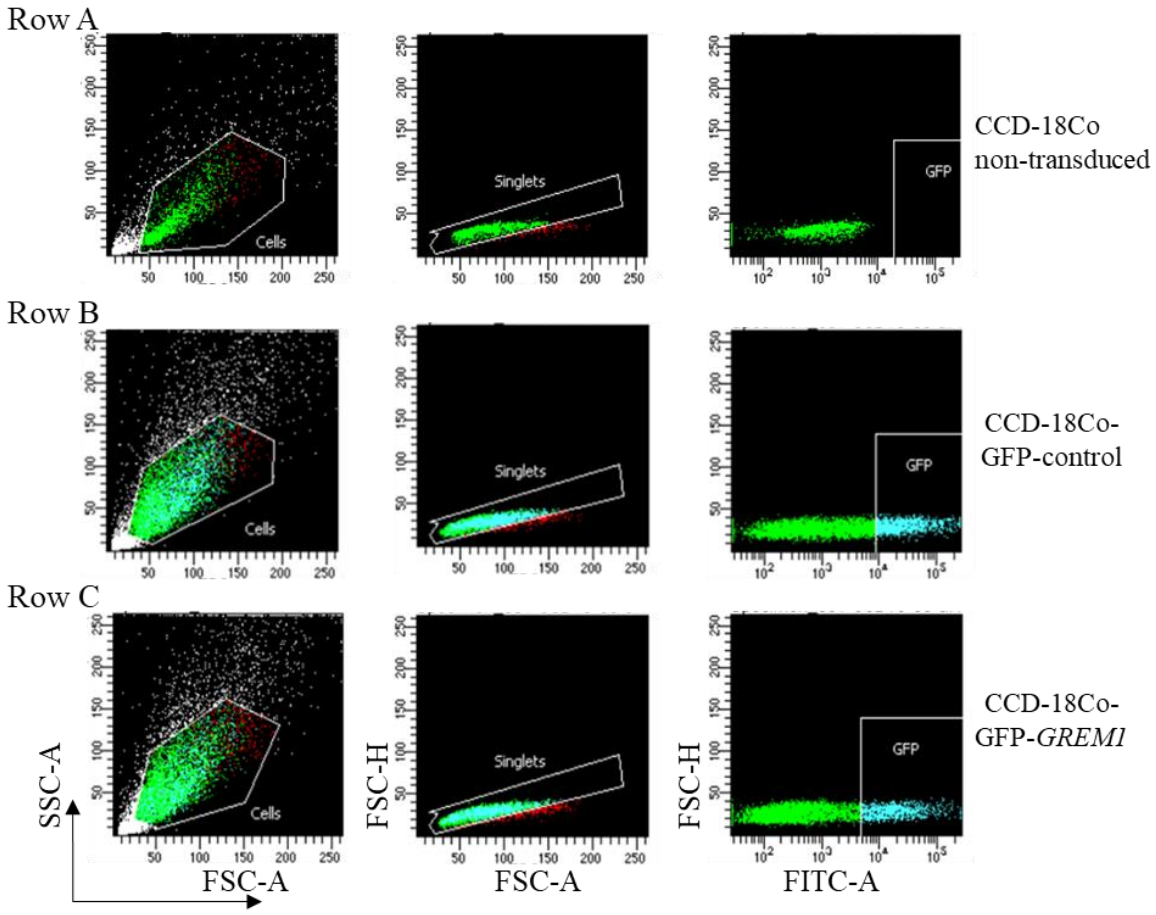


Figure 5.20: Gating strategy for sorting of CCD-GFP positive cells. Transduced CCD-18Co were detached using trypsin and washed ready for sorting using the BD FACSAria cell sorter. **Row 1** shows non transduced CCD-18Co that were used as background to determine the positive GFP signal. **Row 2** shows transduced CCD-18Co-GFP-control cells that were selected based on side and forward scatter, singlets were selected based on forward scatter height and area, GFP^{+ve} labelled cells were selected based on FITC positive signal and the gate as determined by unlabelled cells. **Row 3** shows transduced CCD-18Co-GFP-GREM1 cells that were selected, doublets excluded based on side and forward scatter, GFP^{+ve} labelled cells were selected based on green fluorescence (FITC) positive signal, the negative gate was set up using un-transduced CCD-18Co.

5.3.7 Optimisation of siRNA transfection to knock down GREM1 in CCD-18Co

From previous experiments, I had shown expression of endogenous gremlin 1 at transcript and protein level. However, due to the loss in proliferative capacity of the GREM1 overexpressed CCD-18Co, an alternative approach was tested in which endogenous GREM1 knock down model was generated to study effects on downstream signalling of the TGF-β signalling pathway. In this section, optimisation was carried out to knock down GREM1 expression in CCD-18Co, using four synthetic small interfering RNA (siRNAs). The four siRNAs are designed each with a unique sequence to specifically target different regions of GREM1.

5.3.7.1 Small interfering RNA (siRNA) knockdown model I: use of Lipofectamine transfection reagent did not knock down GREM1

First optimisation experiment involved seeding CCD-18Co in a 24 well plate at 50,000 cells per well. Transfection with the respective GREM1 siRNAs (1-4) at a concentration of 10 nM and 20 nM was performed. The transfection reagent used was Lipofectamine 3000 (30µl/well) and the corresponding P3000 reagent (10µl/well) in Opti-MEM media (Low serum). Cells were left in culture at 370C for 48hrs; lysed and GREM1 expression assessed using RT-qPCR, GREM1 gene expression was analysed. The knock down effect of GREM1 gene expression relative to control siRNA is shown in figure 5.21A (siRNA1), figure 5.21B (siRNA2), figure 5.21C (siRNA3) and figure 5.21D (siRNA4). Overall, the graphs show minimal efficiency in knocking down GREM1 in CCD-18Co. Figure 5.21A shows a slight knock down effect at a 0.6 ratio with the use of siRNA1 at 20nM relative to control siRNA. Figure 5.21B shows a reverse effect in which there is an upregulation of GREM1 at both concentrations of siRNA2 relative to the non-transfected CCD-18Co. Figures 5.21C and figure 5.21D did not show a

knock down effect when siRNA3 and siRNA4 were used. Overall, the experiment did not show a significant knock down effect of GREM1.

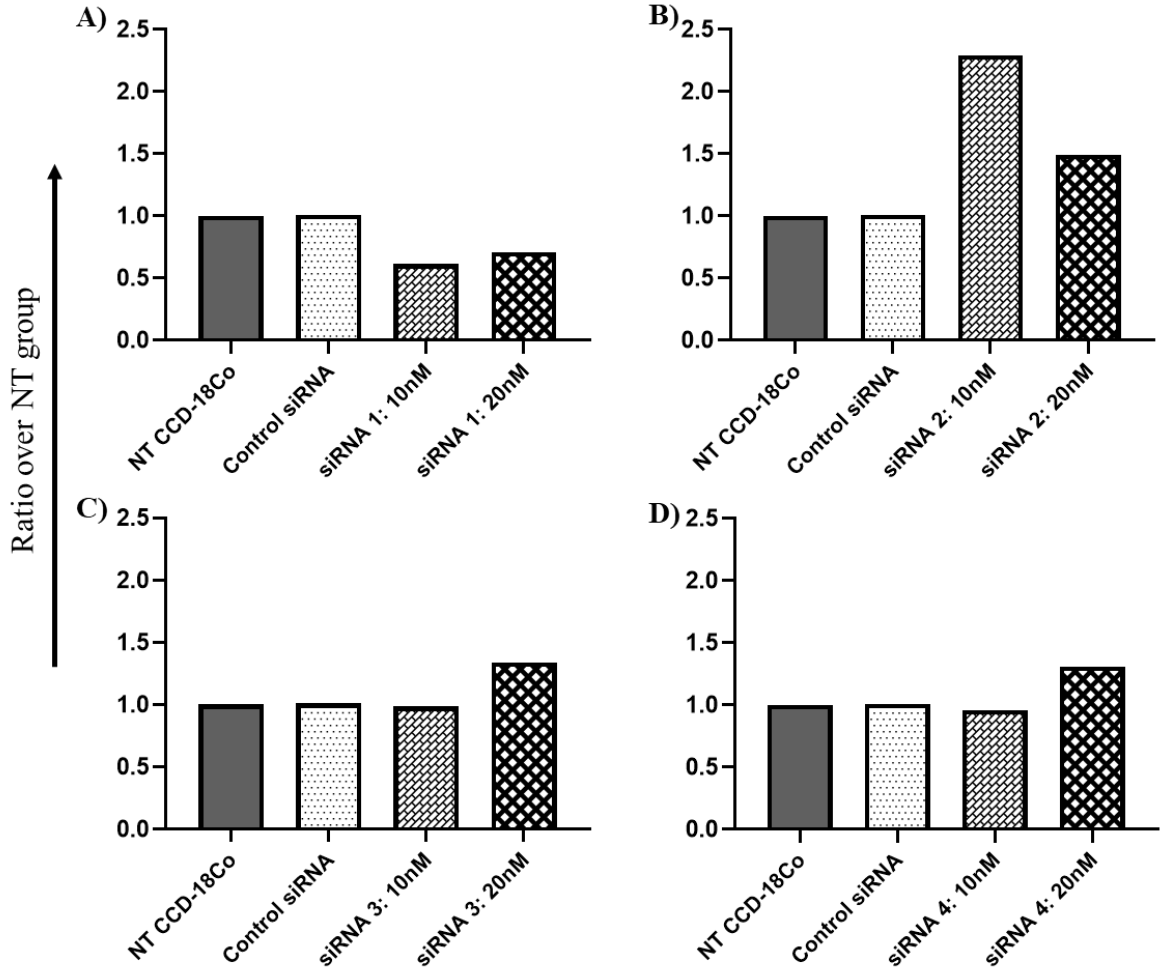


Figure 5.21: Assessing the knock down effect of siRNAs to suppress *GREM1* expression when using lipofectamine as a transfection reagent. CCD-18Co were seeded at $5X10^4$/cells per well for 12 hrs to adhere. Cells were transfected with siRNAs (1-4) including control siRNA and non-transfected cells (NT), lipofectamine reagent was added in Opti-MEM media (low serum). Cells were left in culture for 48hrs. Cells were lysed, RNA extracted, and RT-qPCR was performed on cDNA to check *GREM1* expression. Graphs show effect of using *GREM1* siRNA1 **(A),** siRNA2 **(B),** siRNA3 **(C)** and siRNA4 **(D)** at 10nM and 20nM on *GREM1* gene expression compared to control siRNA and non-transfected cells n=1

5.3.7.2 Small interfering RNA (siRNA) knockdown model II: use of Qiagen transfection reagent showed a GREM1 knock down effect

From the previous experiment, the expected knock down effect was not achieved. I suspected the low success right may be attributed to incompatibility between the siRNA and the transfection reagent hence reducing the cellular uptake of siRNA to achieve effective knock down effects. In this experiment set up, the transfection reagent was changed to the HiPerFect reagent from Qiagen. It was easier to trouble shoot with this because the same supplier provided both siRNA and transfection reagent. The experimental set up was similar to that described above with CCD-18Co seeded in a 24 well plate at 50,000 cells per well before being transfected with the respective GREM1 siRNAs (1-4) at a concentration of 10 nM and 20 nM respectively, non-targeting siRNA at 20 nM was added as a control. HiPerFect transfection reagent (4.5µl/well) was added into the respective wells as per the manufacturers' recommendation. Cells were left in culture at 370C for 48hrs; Cells were lysed and by RT-qPCR, GREM1 gene expression was analysed. Despite changes to the transfection reagent in this experiment, figure 5.22 shows low efficiency in knock down of GREM1 expression in CCD-18Co with the use of siRNA2 (figure 5.22B), siRNA3 (figure 5.22C) and siRNA4 (figure 5.22D) at the 48hr time point. There was a trend that showed a 0.4 (60%) knock down ratio with the use of siRNA1 (20nM) at 48 hrs post transfection with siRNA1 as seen in figure 5.22A. With this trend, further experiments testing out different conditions were tested to optimise the knock down efficiency.

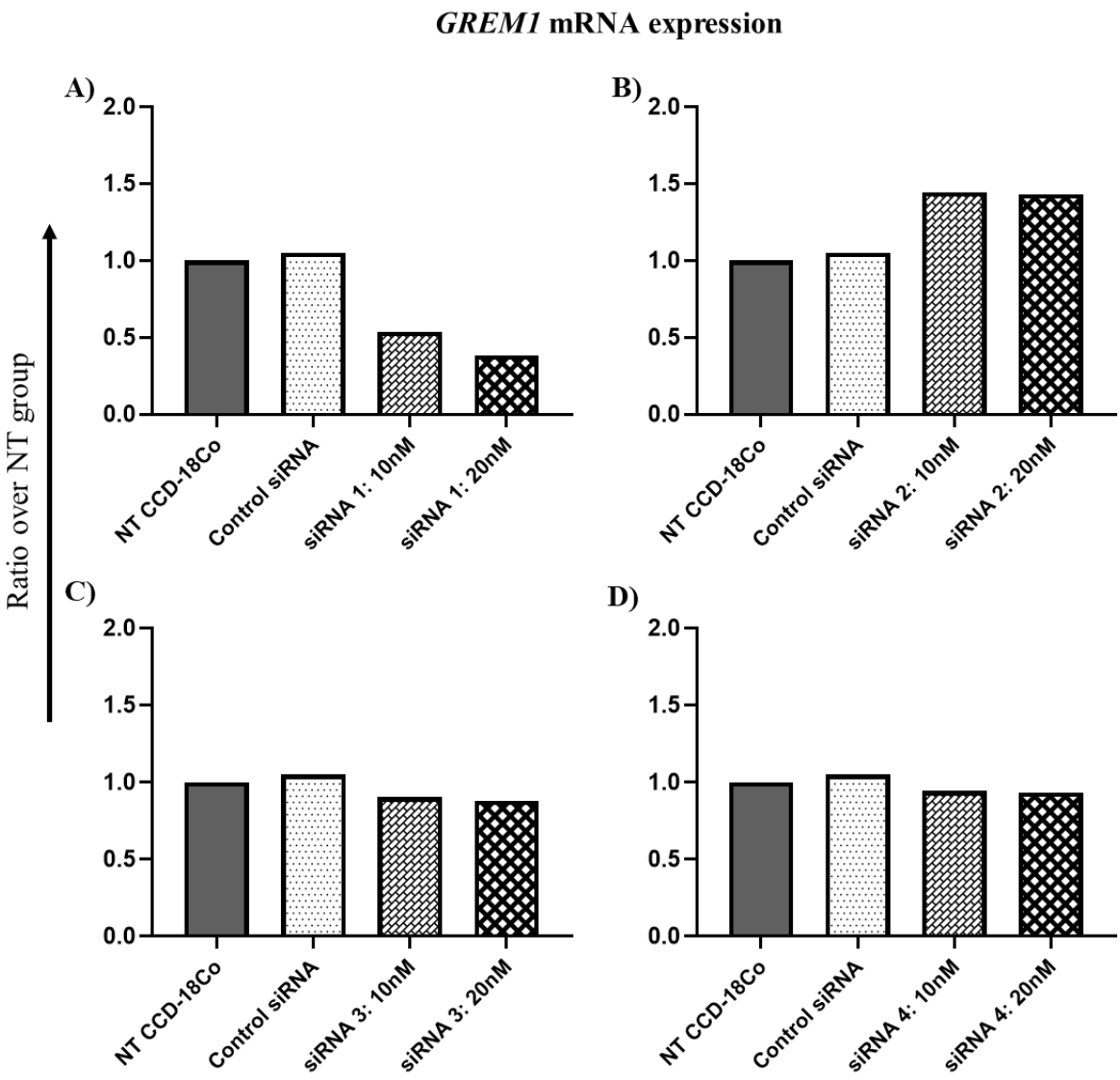


Figure 5.22: Assessing different siRNAs efficiency to suppress *GREM1* expression when using HiPerFect as a transfection reagent. CCD-18Co were seeded at $5X10^4$/cells per well for 12 hrs to adhere. Cells were then transfected with siRNAs (1-4) including control siRNA and non-transfected cells (NT), HiPerFect transfection reagent (Qiagen) at 4.5µl was added with EMEM without serum. Cells were left in culture for 48hrs. Cells were then lysed, RNA extracted and RT-qPCR was performed on cDNA to check *GREM1* expression. Graphs show effect of using GREM1 siRNA1 **(A),** siRNA2 **(B),** siRNA3 **(C)** and siRNA4 **(D)** at 10nM and 20nM on *GREM1* gene expression compared to control siRNA and non-transfected cells n=1.

5.3.7.3 Further optimisation of GREM1 siRNA1 at a higher concentration of 60nM

The data shown in figures 5.21 and 5.22 showed a trend in which the highest knock down effect of GREM1 expression was seen in CCD-18Co transfected with 20nM of siRNA1 at 48hrs post transfection. To improve knock down efficiency and increase the assay window, siRNA1 was used at 20nM and 60nM respectively while maintaining the volume of HiPerFect transfection reagent used at 4.5µl for 48 hrs. Figure 5.23A showed a 0.7 ratio relative to the non-transfected cells with siRNA1 used at 20nM for 48 hrs. This was a 30% knock down effect compared to the 60% knockdown that was shown in the previous experiment. Using siRNA1 at 60nM showed a 0.6 (40%) knock down ratio at 48 hrs post transfection with siRNA1 as seen in figure 5.23B. The results were not consistent in this experiment and the previous one. However, there was a trend that showed a slight increase in knockdown effect at a higher concentration of 60nM as compared to 20nM. The next experiment was set up to improve knock down efficiency by altering volume of the transfection reagents at different time points.

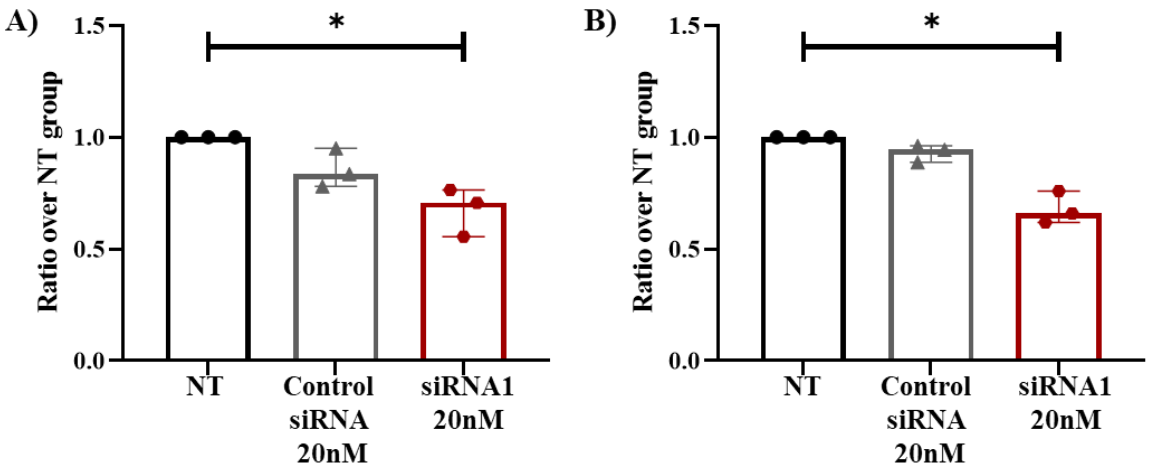


Figure 5.23: Assessing siRNA1 efficiency at 20nM vs 60nM to suppress *GREM1* expression. CCD-18Co were seeded at 5X10^4/cells per well for 12 hrs to adhere. Cells were then transfected with siRNA1 (20nM and 60nM respectively), control siRNA (60nM) and non-transfected cells (NT), HiPerFect transfection reagent (Qiagen) at 4.5µl was added with EMEM without serum. Cells were left in culture for 48hrs. Cells were then lysed, RNA extracted, and RT-qPCR was performed on cDNA to check *GREM1* expression. **A)** Graph shows effect of using *GREM1* siRNA1 at 20nM at 48 hrs **B)** Graph shows effect of using *GREM1* siRNA1 at 60nM at 48 hrs. Control siRNA was used at the higher concentration at 60nM. Data presented as median with 95% CI, statistics was performed using one-way ANOVA and multiple comparisons were performed using the Kruskal-Wallis test (*P<0.05) n=3

5.3.7.4 Increase in GREM1 knock down using siRNA1 at a later time point of 72 hours

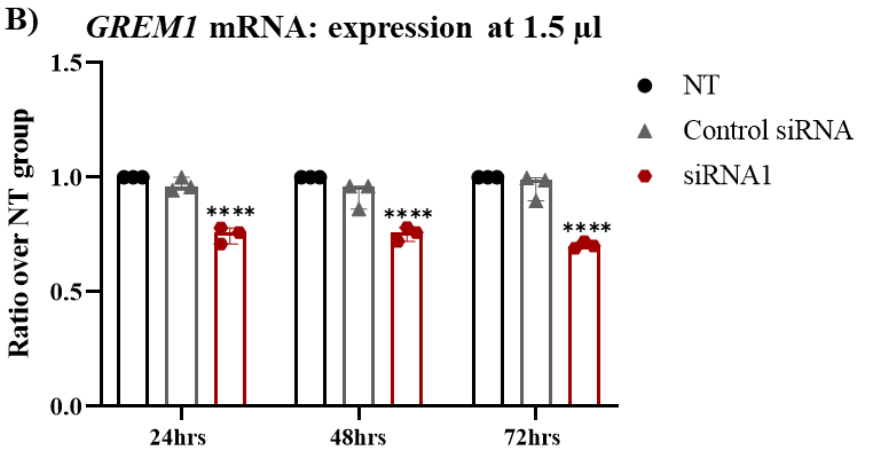
To improve knockdown efficiency of siRNA1, further optimisation experiments to increase the efficiency of GREM1 siRNA1 were performed by addition of varying volumes of transfection reagent and comparing knock down at 3 different time points. CCD-18Co were stimulated with siRNA1 at 60nM with variation in the amount of transfection reagent added (1.5µL, 3µL and 4.5µL) in each well, cells were left in culture at 24, 48 and 72 hrs respectively. The results of GREM1 mRNA expression are shown in figure 5.24 where GREM1 expression was reduced at 60nM of siRNA1. There was a decreased GREM1 expression across all conditions with p values of <0.0001, however, the highest knockdown was seen at 72 hrs where knockdown was up to 70% relative to control siRNA. Multiple comparison analysis using 2way ANOVA was performed to check if siRNA affects the results (control siRNA versus GREM1 target siRNA)

and if time has an effect on knock down. The statistical test comparison showed the effect of siRNA comparison as extremely significant with a p value of <0.0001. The data shows a knock down effect at all time points, however, the highest knock down effect is seen at 72 hrs with transfection reagent used at 4.5µL, the knock down effect was at a ratio of 0.3 (70% knock down effect) relative to control siRNA. This was the highest knock down effect seen throughout the whole optimisation process and was used for further experiments.

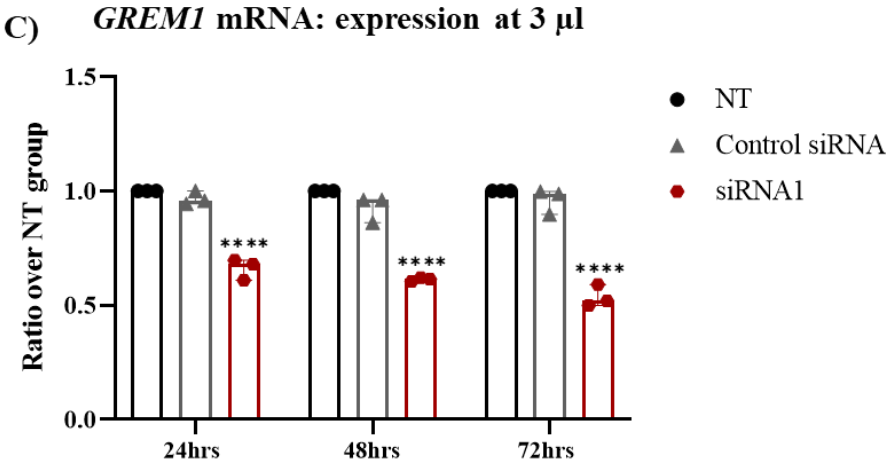
A)

Transfection reagent for 24 well (60nM)	Time course
1.5ul	24hrs
3ul	48hrs
4.5ul	72hrs

B)



C)



D)

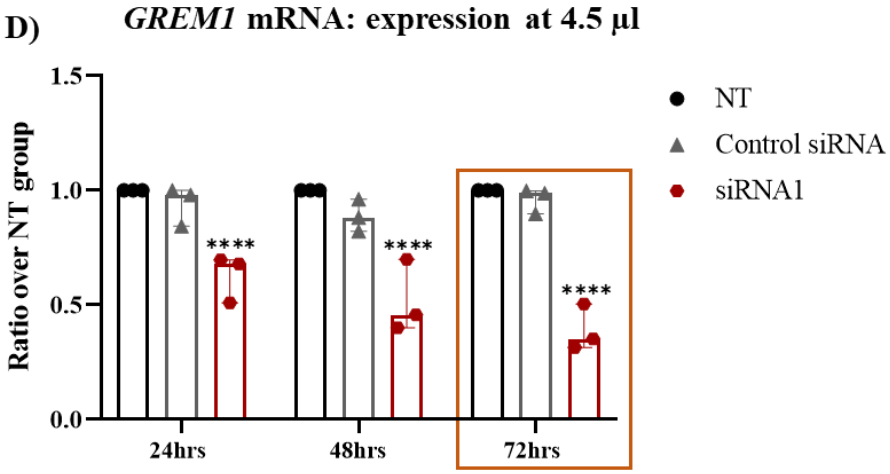


Figure 5.24: *GREM1* mRNA expression was reduced in CCD-18Co at 72 hrs following transfection of 60 nM siRNA. CCD-18Co were seeded at $5X10^4$/cells per well for 12 hrs to adhere. Cells were then transfected with siRNA1 and control siRNA at 60nM, non-transfected cells (NT) were included. HiPerFect transfection reagent (Qiagen) was added at varying volumes 1.5µl, 3µl and 4.5µl made up in EMEM without serum. Cells were left in culture for 24, 48 and 72 hrs respectively. Cells were then lysed, RNA extracted, and RT-qPCR was performed on cDNA to check for *GREM1* expression. **A)** Table showing conditions in which cells were transfected. *GREM1* mRNA expression shown in comparison to non-transfected cells and control siRNA over 24, 48 and 72 hrs at varying dilutions of transfection reagent (1.5µl, 3µl and 4.5µl). **B-D)** Knockdown of *GREM1* expression when using 1.5µl (B), 3.0 µl (C) or 4.5 µl (D) of transfection. Data presented as median with 95% CI, statistics was performed using a two-way ANOVA and multiple comparisons were performed using the Kruskal-Wallis test (*P<0.05, **P<0.005, ***P<0.0005, ****P<0.0001) n=3.

5.3.7.5 GREM1 knock down effect maintained at 48hrs post transfection

Gene knockdown effect has been shown to last up to 5 days post transfection, studies show that 24–96 hrs is the ideal time periods for accessing gene knockdown and investigating functional effects of siRNA knockdown in cell culture (557). To confirm if the knock down effect is maintained during the time frame of the assay window, cells were transfected with siRNA1 at 60nM and transfection reagent of 4.5µL for 72 hrs. After which, media containing the siRNA/transfection reagent complex was removed and replaced with the normal cell culture media. Cells were left in culture for an additional 24 and 48 hrs. Cells were collected at the end of the time points and lysed ready for RNA extraction, and RT-qPCR for GREM1 expression. The data shown in figure 5.25 confirms that the knock down effect was maintained 48 hrs post transfection with comparisons between control and siRNA transfected cells showing statistically significant differences in expression both at 24 hrs (p value=0.0352*) and 48 hrs (p value=0.0002***) post transfection. Interestingly, expression of GREM1 seemed to recover at 48 hrs with a slight increase in expression as compared to 24 hrs with a p value of 0.055, this

showed a trend towards loss of the knock down effect. This was a factor to consider when conducting further mechanistic experiments with longer time points.

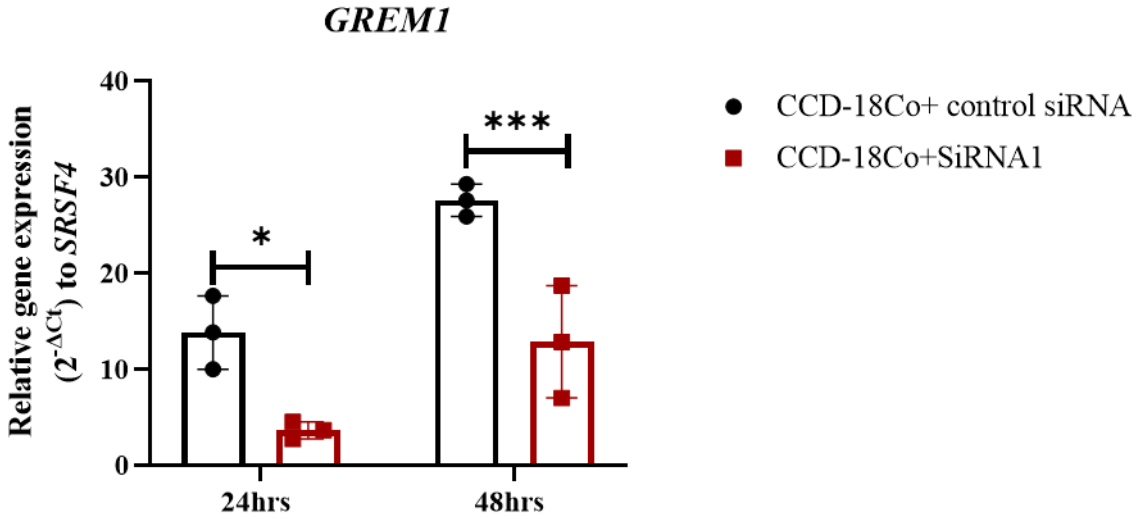


Figure 5.25: *GREM1* mRNA expression maintained 48 hrs post transfection. CCD-18Co were seeded at 5X10^4/cells per well for 12 hrs to adhere. Cells were then transfected with siRNA1 and control siRNA at 60nM, non-transfected cells (NT) were included, HiPerFect transfection reagent (Qiagen) was added at 4.5µl made up in EMEM without serum. Cells were left in culture for 72 hrs. Media containing siRNA was then replaced by complete media (10% FBS+EMEM), and cells were maintained in culture and collected at 24 and 48 hrs post. Cells were lysed, RNA extracted and RT-qPCR was performed on cDNA to check for GREM1 expression. Graph shows expression of GREM1 relative to SRSF4 that was used as the endogenous control. Data presented as median with 95% CI, statistical significance on gene expression with time was determined by performing a one-way ANOVA and multiple comparisons were performed using tukey (*P<0.05, **P<0.005, ***P<0.0005, ****P<0.0001) n=3

5.3.7.6 TGF-β inhibits GREM1 expression in siRNA transfected CCD-18Co 48 hrs post transfection.

TGF-β is known to regulate fibrosis through TGF-β/SMAD signalling axis, the literature suggests cross talks between TGF-β and BMP pathways. To assess the effect of recombinant TGF-β on expression of GREM1 in CCD-18Co, cells were transfected with siRNA1 and control at 60nM concentration with transfection reagent of 4.5µl for 72 hrs. Cells were

stimulated with TGF-β at 50ng/ml and 100 ng/ml respectively. Cells were collected at the 24 hr and 48 hr time points and lysed ready for RNA extraction for RT-qPCR. The data shown in figure 5.26 shows the presence of TGF-β at both concentrations inhibits expression of GREM1 48 hrs post stimulation, no effect on GREM1 expression is seen at 24 hrs. Comparison between unstimulated and stimulated transfected cells (GREM1 knock down) indicate an inhibitory effect that TGF-β stimulation exerts on CCD-18Co expression of GREM1 at 48 hrs with a statistical significance; p value at 50ng/mL (p value 0.014) and 100ng/mL (p value 0.024) respectively. This suggests that TGF-β stimulation does not rescue expression of GREM1 in siRNA transfected CCD-18Co. No effect of TGF-β stimulation was seen at 24 hrs.

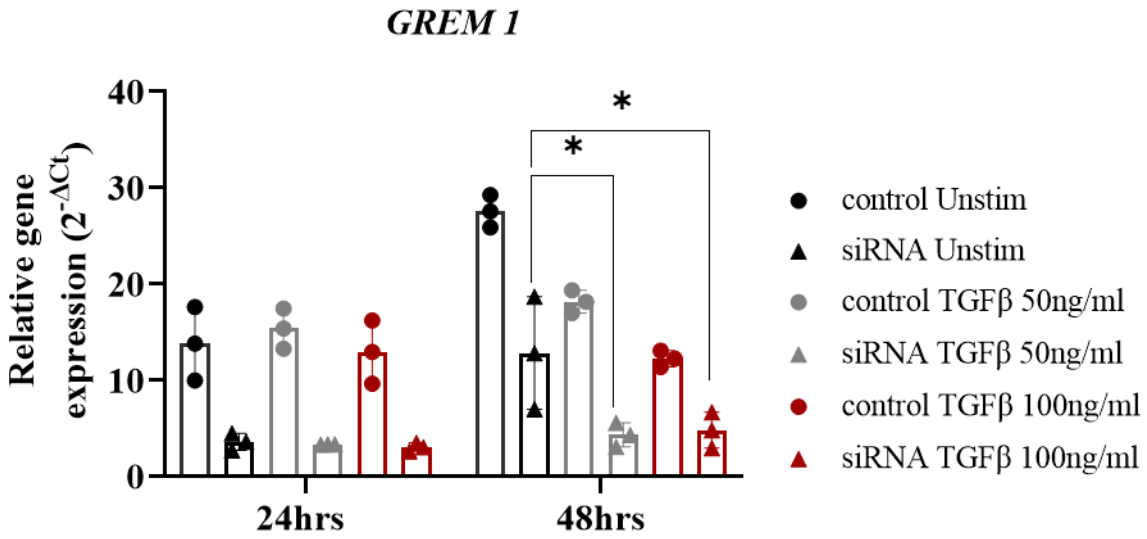


Figure 5.26: TGF-β stimulation does not rescue GREM1 expression in transfected CCD-18Co 48hrs post knockdown. CCD-18Co were cultured and transfected with GREM1 siRNA or control and left in culture for 72hrs at 37^0C and 5% CO_2. Cells were then stimulated with TG-Fβ1 at 50ng/ml and 100ng/ml respectively for 48hrs. Cells were harvested and processed for qPCR to ensure GREM1 gene expression levels were maintained post transfection. Relative gene expression is expressed as $2^{-\Delta Ct}$. Data presented as median with 95% CI, n=3, statistics performed using 2-way ANOVA with multiple comparisons performed.

5.3.7.7 TGF-β inhibits ACTA2 expression in siRNA transfected CCD-18Co 48 hrs post transfection.

To study the effect of TGF-β stimulation on siRNA transfected CCD-18Co, CCD-18Co were transfected with either control siRNA or GREM1 siRNA at 60nM concentration with 4.5µL per well of transfection reagent and cells left in culture for 72 hrs, cells were stimulated with TGF-β at 50ng/ml and 100 ng/ml respectively, unstimulated cells were included as controls. Cells were collected at the 24 hr and 48 hr time points and lysed ready for RNA extraction and RT-qPCR. The readout for this experiment was ECM genes including COL1A, ACTA2, TIMP1 and FN1. The data shown in figure 5.27 shows similar trends in target gene expressions at the 24 and 48 hr time points suggesting that overall, time does not significantly affect gene expression.

However, stimulation with TGF-β has a significant effect on overall gene expression. The data shows that knocking down GREM1 has a significant (p value 0.013) inhibitory effect in expression of ACTA2, a known profibrotic gene as seen in unstimulated cells in figure 5.27B. In addition, stimulation with TGF-β does not reverse the inhibitory effect of GREM1 knock down. On the contrary, knocking down GREM1 slightly increases expression of COL1A1, TIMP1 and FN1 at 24 and 48 hrs respectively. Stimulation of transfected CCD-18Co with TGF-β enhanced the level of expression of COL1A (figure 5.27A), TIMP1(figure 5.27C) and FN1(figure 5.27D). This effect was seen specifically in CCD-18Co in which GREM1 had been knocked down. It suggests a rescue response either due to increased TGF-β signalling or alternative mechanisms that may have been activated due to the knock down effect.

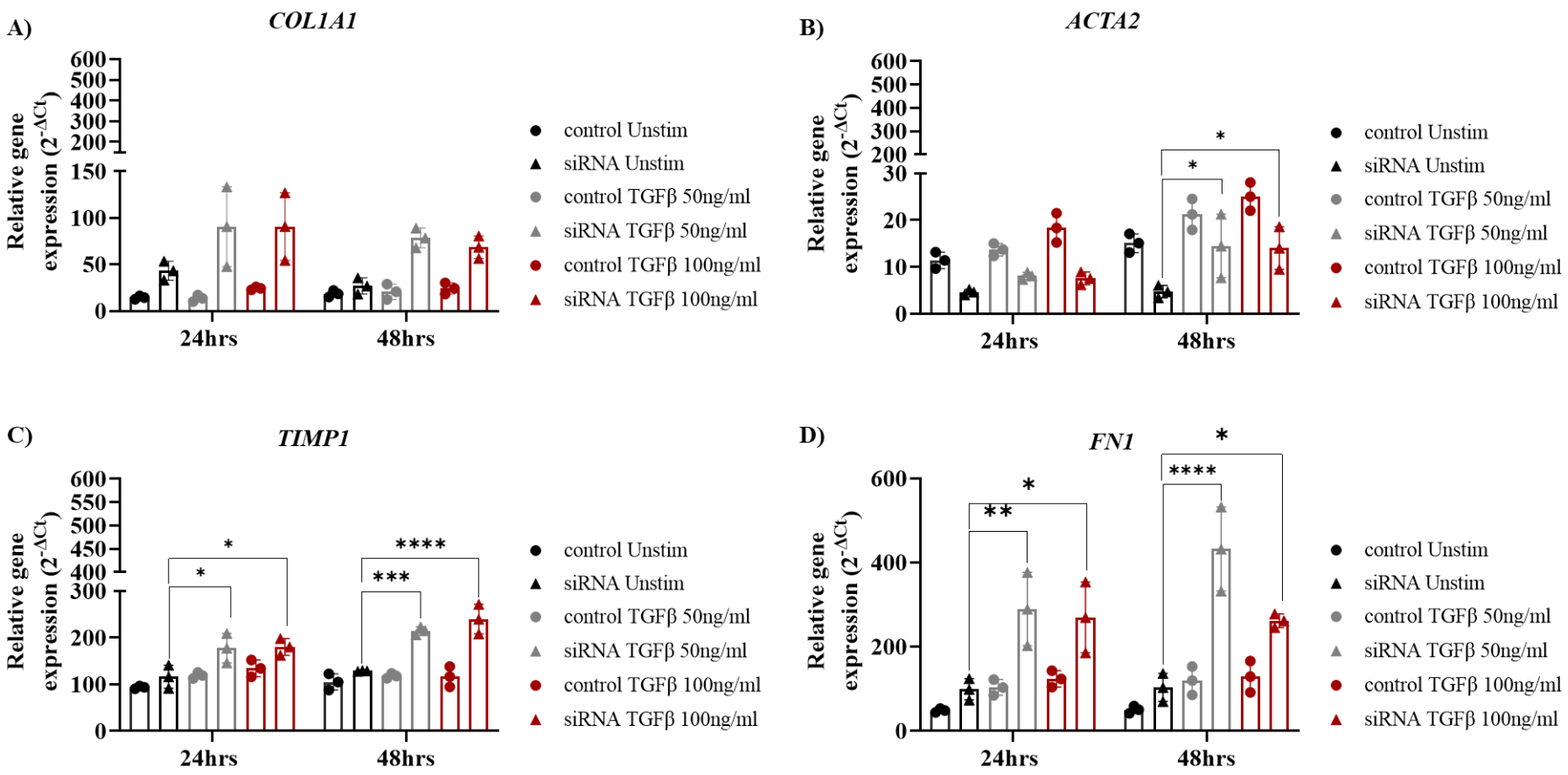

A)
COL1A1
Relative gene expression ($2^{-\Delta Ct}$)
600
500
400
300
200
150
100
50
0
24hrs
48hrs
control Unstim
siRNA Unstim
control TGFβ 50ng/ml
siRNA TGFβ 50ng/ml
control TGFβ 100ng/ml
siRNA TGFβ 100ng/ml
B)
ACTA2
Relative gene expression ($2^{-\Delta Ct}$)
600
500
400
300
200
30
20
10
0
*
*
24hrs
48hrs
control Unstim
siRNA Unstim
control TGFβ 50ng/ml
siRNA TGFβ 50ng/ml
control TGFβ 100ng/ml
siRNA TGFβ 100ng/ml
C)
TIMP1
Relative gene expression ($2^{-\Delta Ct}$)
600
550
500
450
400
300
200
100
0
*
*
****
***
24hrs
48hrs
control Unstim
siRNA Unstim
control TGFβ 50ng/ml
siRNA TGFβ 50ng/ml
control TGFβ 100ng/ml
siRNA TGFβ 100ng/ml
D)
FN1
Relative gene expression ($2^{-\Delta Ct}$)
600
400
200
0
*
**
*
****
24hrs
48hrs
control Unstim
siRNA Unstim
control TGFβ 50ng/ml
siRNA TGFβ 50ng/ml
control TGFβ 100ng/ml
siRNA TGFβ 100ng/ml

Figure 5.27: GREM1 siRNA knockdown decreases ACTA2 expression in both unstimulated and TGF-β stimulated cells. CCD-18Co were seeded at $5X10^4$/cells per well for 12 hrs to adhere. Cells were then transfected with siRNA1 and control siRNA at 60nM, non-transfected cells (NT) were included, HiPerFect transfection reagent (Qiagen) was added at 4.5µl made up in EMEM without serum. Cells were left in culture for 72 hrs respectively. Media containing siRNA was replaced by complete media (10% FBS+EMEM), cells were maintained in culture and collected at 24 and 48 hrs post. Cells were lysed, RNA extracted and RT-qPCR was performed on cDNA to check for ECM related genes **A)** *COL1A1* **B)** *ACTA2* **C)** *TIMP1* and **D)** *FN1* expression. Graph show expression of ECM genes relative to *SRSF4* that was used as the endogenous control. Data presented as median with 95% CI, statistical significance on gene expression with time was determined by performing a one-way ANOVA and multiple comparisons were performed using tukey (*P<0.05, **P<0.005, ***P<0.0005, ****P<0.0001) n=3.

5.3.7.8 Gremlin 1 promotes proliferation of CCD-18Co in vitro.

To investigate the biological effect of gremlin 1 on proliferation, CCD-18Co were transfected with either control siRNA or GREM1 siRNA at 60nM concentration with 4.5µL per well of transfection reagent and cells left in culture for 72 hrs, cells were stimulated with TGF-β at 50ng/ml and 100 ng/ml respectively, unstimulated cells were included as controls. Cells were collected at the 24 hr and 48 hr time points and lysed ready for RNA extraction and RT-qPCR. The readout for this experiment was *MKI67* as a known marker for cellular proliferation. Figure 5.28 showed that GREM1 knockdown resulted in a significant reduction in proliferation of CCD-18Co as seen by reduced expression of MKI67 (p value <0.0001) at 24 hrs. A similar trend was seen at 48 hrs however, this was not statistically significant. Stimulation with TGF-β did not have an effect on CCD-18Co at both time points. The data suggest a role for GREM1 in proliferation of CCD-18Co.

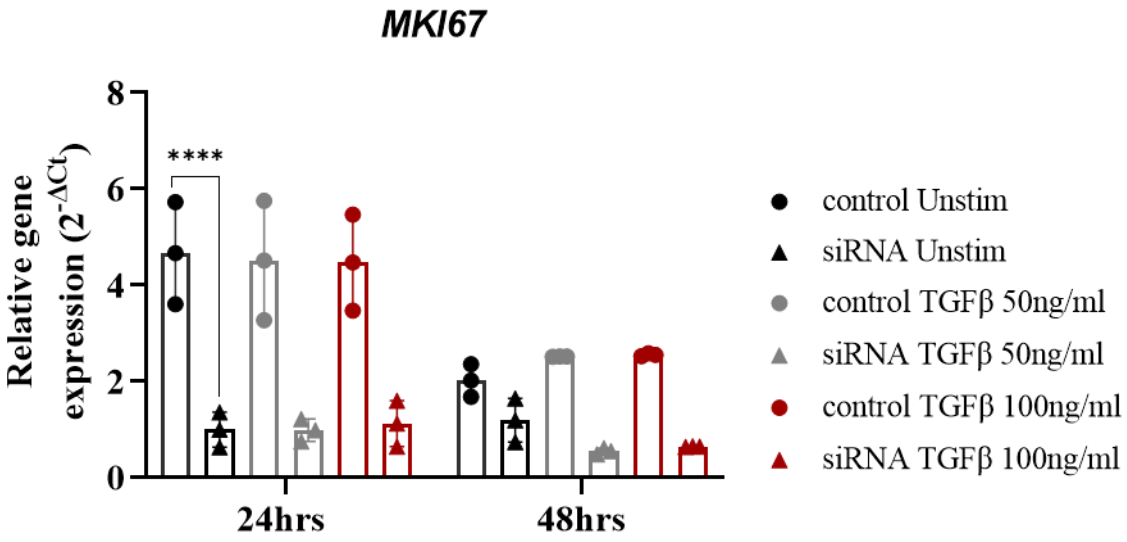


Figure 5.28: GREM1 siRNA knockdown decreases proliferation in CCD-18Co. CCD-18Co were seeded at $5X10^4$/cells per well for 12 hrs to adhere. Cells were then transfected with siRNA1 and control siRNA at 60nM, non-transfected cells (NT) were included, HiPerFect transfection reagent (Qiagen) was added at 4.5µl made up in EMEM without serum. Cells were left in culture for 72 hrs respectively. Media containing siRNA was replaced by complete media (10% FBS+EMEM), cells were maintained in culture and collected at 24 and 48 hrs post. Cells were lysed, RNA extracted and RT-qPCR was performed on cDNA to check for MKI67 proliferation gene. Graph shows expression of MKI67 relative to SRSF4 that was used as the endogenous control. Data presented as median with 95% CI, statistical significance on gene expression with time was determined by performing a one-way ANOVA and multiple comparisons were performed using tukey (*P<0.05, **P<0.005, ***P<0.0005, ****P<0.0001) n=3

5.4 Discussion

Fibroblasts have been reported as the main effector phenotype in organ fibrosis (436). The aim of this section of the study was to understand the role of gremlin 1 in primary intestinal fibroblasts isolated from the proximal, stricture and distal segments of resected gut. These were to be used for in vitro cell cultures to study the role of gremlin 1 in the pathogenesis of stricture formation in CD. From this study, isolation of primary fibroblasts was successfully performed, however, the protocol optimisation process was not straight forward. Multiple protocols were tested including the “explant liver protocol” which is a well-established protocol for isolation

of primary fibroblasts from explant liver at the CLGR, Birmingham. The unsuccess with this protocol when used to process resection samples for fibroblast isolation was associated to sample size as explant liver used for the procedure was usually approximately 20-30g while resected tissue was limited to 1-2g as starting material; the protocol involved multiple wash steps that may have led to loss of fibroblast population which is generally a small population in tissue. Other tested protocols included the "organ culture protocol" which involved outgrowth of fibroblasts from tissue left in culture. This was successfully used by (558). However, in this study, the protocol was tested with no success, its not clear why the protocol in the literature could not be reproduced. The reason may be attributed to the tissue not being cut finely enough to expose fibroblasts to easily migrate out of the tissue and adhere to the culture plate, or that the abundance of fibroblasts was very low which may have been coupled with low cell viability leading to loss in the target cell population.

The next protocol tested was based on flow cytometry data generated for this study that indicated that the frequency of epithelial cells was high within the stromal cell population. The "tissue enzymatic digestion protocol" involved depletion of epithelial cells from the tissue followed by tissue dissociation and digestion, the use of this protocol was also an unsuccessful attempt to isolate primary fibroblasts from resected tissue. Overall, the protocols used up to this point involved multiple wash steps. Fibroblasts are a small population in the tissue that are easily lost during multiple wash steps. The thought was to minimise multiple washes, while ensuring total tissue digestion and maintaining cellular viability with reduced processing time; hence the use of the "hybrid technique". This technique combined mechanical and enzymatic digestion and proved successful in isolation of primary intestinal fibroblasts.

In general, the challenges faced during optimisation of the various protocols was limited access to tissue samples and low frequencies of tissue fibroblasts that are easily lost during the

isolation steps. To resolve this, explant liver tissue was used for optimisation, however, due to the differences in tissue type in the form of size and texture, it was not possible to directly replicate the same protocol to be used for isolation of fibroblasts from resected gut. The hybrid protocol that combined enzymatic and mechanical digestion was successfully used to isolate a total of 6 samples, 3 samples from the proximal and 3 samples from the stricture. Isolated fibroblasts were sub-cultured to a passage of 3, the cells seen in culture were spindle-shaped, a similar morphology to that of fibroblasts. The cells were cryopreserved at a low passage of 3 in 5% DMSO:95% FCS for further downstream experiments. Fibroblasts have been shown to change phenotype with higher passages hence the need to keep the passage as low as possible (559). During the thawing process to perform characterisation and confirmation of phenotype, a large proportion of cell death was seen, the remaining fibroblasts in culture did not proliferate any further. The loss in proliferation capacity may be attributed to the cellular stress during the freezing down process resulting in induction of cellular senescence. The failure in cells to proliferate resulted in insufficient cell numbers for characterisation, phenotyping, and experimental assays to study the role of gremlin 1 in stricture formation in CD.

As an alternative, the use of CCD-18Co was proposed for further downstream experiments. CCD-18Co is a normal human gut derived fibroblast cell line. For this study, CCD-18Co was used as an in vitro model to study mechanisms of stricture formation in intestinal fibrosis. The cell line was characterised to determine if it had a similar phenotype to that of primary fibroblasts. Fibroblasts are a heterogenous cell population, this makes cellular identification and characterisation challenging. In the past decade, multiple single cell studies have identified markers that can be used to identify and characterise fibroblasts. They have been shown to express CD90 (Thy1), PDGFRα, PDGFRβ, and podoplanin and α-SMA (560,561). A single cell study by (375) showed that GREM1 expression was localised to a cell population identified as fibroblasts, this was similar to findings from our sc-RNA-seq data in which GREM1

expression was localised to the stromal cell cluster. To validate the use of CCD-18Co as an in vitro model to study fibrotic mechanisms, characterisation was performed for expression of fibroblast markers. By flow cytometry, CCD-18Co showed a negative expression for CD45 which is an immune cell marker, the cells were also negative for CD31 and EpCAM which are markers expressed on endothelial and epithelial cells respectively. Interestingly, CCD-18Co were negative for CD34 which is a mesenchymal cell marker and has been shown as expressed by fibroblasts (562). This may be attributed to the quality of the antibody that was used to perform the staining. In hindsight, antibody titration was not performed for this antibody which should have been done at the time. CCD-18Co were also negative for PDGFRβ by flow cytometry but positive by immunofluorescence, a technique in which permeabilisation for intracellular staining was performed.

PDGFRβ belongs to a family of type III receptor tyrosine kinases (RTK), they are composed of an extracellular ligand binding domain, a transmembrane segment and an intracellular part containing a split tyrosine kinase domain. Ligand-induced dimerization of these receptors leads to activation of intrinsic kinase activity and subsequent autophosphorylation of tyrosine residues in the intracellular part. The flow cytometry staining performed to characterise CCD-18Co was surface based, no permeabilisation was done to access receptor expression intracellularly. This may have limited access to the intracellular tyrosine kinase domain part of PDGFRβ which could be predominant in CCD-18Co. This argument can be supported by the positive staining for PDGFRβ seen in figure 5.9 in which intracellular staining was performed using immunofluorescence. Another reason for the negative staining for PDGFRβ on CCD-18Co by flow cytometry may be linked to the antibody used for flow cytometry, as the staining protocol was optimised to detect target proteins expressed on the surface and not intracellularly. Overall, CCD-18Co were shown to express vimentin, podoplanin, CD90, PDGFRβ and α-SMA using flow cytometry and immunofluorescence. The phenotype of CCD-18Co based on protein

expression of fibroblast like markers was similar to what has been published in the literature as markers that identify fibroblasts and myofibroblasts (549). This validated the use of CCD-18Co as an in vitro cell line to study mechanisms through which gremlin 1 could modulate stricture formation.

CCD-18Co were stimulated with profibrotic mediators including TGF-β, IL-13 and IL-36, results from this study showed increase in expression of COL1A1, ACTA2, FN1 and TIMP1 by CCD-18Co upon stimulation with TGF-β at 24 hrs. However, there was no dose dependent effect seen with increase in concentration. A two time point experiment was further performed comparing gene expression at 24 and 48 hrs respectively, gene expression was increased at the 48hr time point. These findings are consistent with published data for multiple fibrotic conditions that suggests that development of fibrosis involves ECM driven mechanisms. A study performed by (Truffi et al., 2018) showed that inhibiting fibroblast activation protein resulted in reduced expression of type I collagen and TIMP-1 in ex vivo tissue from CD patients. Restoring the balance between MMPs and TIMPs is an important mechanism to target in anti-fibrotic therapy.

Increased proliferation of fibroblasts is one of the key mechanisms that drive organ fibrosis, studies have shown that in the presence of profibrotic mediators, fibroblasts are activated and there is increase in fibroblast proliferation. For example, mouse fibroblasts were stimulated with TGF-β resulting in their increased activation and proliferation (563). In this study, CCD-18Co were stimulated with IL-1β, PDGF-BB and TGF-β, the effects of these mediators on proliferation of CCD-18Co was monitored. The data showed increase in proliferation of CCD-18Co stimulated with IL-1β, TGF-β and PDGF-BB compared to unstimulated cells at the 48 hr time point. Further validation was performed by stimulating primary fibroblasts from explant PSC liver, PSC primary fibroblasts were stimulated with IL-1β and TGF-β, these showed a

clear increase in proliferation at 48 hrs with both profibrotic factors but more so with the IL-1β stimulation. The data is in agreement with what has been published on TGF-β driven increase in proliferation of dermal and lung derived primary fibroblasts (564).

Having shown CCD-18Co as a cell line that can be used as an in in vitro model to study intestinal fibrosis. Further studies were performed to determine the role of gremlin 1 in intestinal fibrosis. Gremlin 1 was detected in CCD-18Co both at mRNA and in supernatant by ELISA. With the use of the lentiviral vector transduction, GREM1 was overexpressed in CCD-18Co using lentivirus vector plasmids, however the transduced cells did not survive sorting by flow cytometry which precluded any further study. Experiments were performed on cells that were transduced but not sorted, gene read outs of *COL1A1*, *ACTA2*, *FN1* and *TIMP1* showed no differences between the unstimulated and TGF-β stimulated cells.

CCD-18Co were also transfected with siRNA to knock down GREM1 expression; 4 synthetic siRNAs were tested for the optimisation. The optimisation involved testing varying concentrations of siRNA, testing two transfection reagents at different time points. A 70% GREM1 siRNA knockdown was achieved at 60nM of siRNA1 using Hi-PerFect transfection reagent at 4.5µl per well at the 72hr time point. Further studies were carried out to investigate the effect of GREM1 knock down on ECM gene expression and proliferation. The study demonstrated that GREM1-targeting siRNA transfected cells stimulated with or without TGF-β showed an increase in expression of *COL1A1*, *TIMP1* and *FN1*. On the other hand, there was a reduction of ACTA2 in GREM1-targeting siRNA transfected cells. This was an interesting finding and it did fit with the hypothesis, knocking down GREM1 led to increased BMP signalling which in turn regulated TGF-β downstream signalling hence restoring the balance between the two pathways. This translates to a reduction in ECM related gene expression. This is all occurring at a transcript level, however, there is need to establish if this could be translated

to protein. It was also interesting to note that stimulation of CCD-18Co with TGF-β did not cause increase in expression of GREM1. Most of the data presented from the gremlin 1 studies were measured at transcript level because it is challenging to obtain commercially available antibodies for detection of gremlin 1 at a protein level.

It cannot be ruled out that TGF-β might be regulating gremlin 1 expression at a post transcription level. The data generated seems to disagree with our general hypothesis that knocking down gremlin 1 results in increased signalling of BMP signalling pathway as BMP2 and 4 bind their cognate receptors. The two pathways share SMAD4 as a co-factor, with the BMP pathway signalling with no inhibition by gremlin 1, our hypothesis suggests a reduction in expression downstream the TGF-β pathway which is in agreement with lowered expression of *ACTA2*. The hypothesis does not agree with increased expression of other ECM proteins including *COL1A1, TIMP1* and *FN1*. From the data, it was also very clear that gremlin 1 has a major role in proliferation. The expression of *MKI67* was reduced in CCD-18Co transfected cells, this has been shown in cancer associated fibroblasts (565,566).

Overall, the challenges in the study were significant especially access to tissue samples for the isolation of primary fibroblasts for further downstream experiments. The use of CCD-18Co was useful in generation of useful data, although it is not clear if the data generated would be reproduced in primary fibroblasts. The data indicate that gremlin 1 plays a role in intestinal fibrosis with measurable differences in fibrotic gene expression following knockdown of GREM1 expression. The involvement of gremlin 1 in the proliferation of CCD-18Co provides a hint of the role of gremlin 1 in the development of organ fibrosis. Further studies will need to be performed in primary cells and in vivo to understand the mechanisms through which gremlin 1 contributes to the pathogenesis of intestinal strictures in CD.

CHAPTER 6

GENERAL DISCUSSION AND CONCLUDING REMARKS

6.1 Overview

Intestinal fibrosis is a multifactorial process that is influenced by various cellular and molecular mediators. Progress has been made in research to develop effective anti-inflammatory therapies, however, approximately 50% of patients still develop strictures in CD within 10 years of diagnosis (567), suggesting inflammation independent pathways that drive fibrosis. Our understanding of the fibrotic microenvironment has greatly improved, however, there is still a knowledge gap in the specific mechanisms that drive disease pathogenesis in stricture formation in CD. In this study, sequencing was used to understand the molecular and cellular drivers of fibrosis in resected tissue from CD patients. The study provides a novel approach of using matched tissue from each patient consisting of the fibrotic and non-fibrotic regions, this makes it possible to identify mechanisms that are dysregulated in the stricture as compared to the non-stricture segments.

6.2 A unique profibrotic signature in the stricture segment of CD resected samples

To identify transcriptome alterations in the stricture compared to the non-stricture segments, bulk and single cell RNA-seq was performed on resected tissue samples collected from the proximal, stricture and distal segments. The three segments collected were matched from the same patient. Although a small sample size of 25 patients was used for the study and the variation of clinical and demographic characteristics may constrain the power of our analysis, the variations observed in the different segments were considered as derived due to disease biology. This was due to the quality assurance performed during the analysis process and the additional validation assay that were performed by RT-qPCR to confirm findings from the bulk RNA-seq data.

Results from this study showed a distinct gene expression profile between the stricture and non-stricture segments. A summary of genetic biomarkers compiled by Rieder (2014) showed

a group of genes that can be used as early predictors of intestinal fibrosis. Some of these genes including *NOD2*, IL23R, *COL1A1* were identified in our sequencing data. Using unsupervised and supervised PCA on 501 DEGs, the samples from the stricture clustered together from the non-stricture based on gene expression. This suggests a unique gene expression signature in the stricture segment. Gene enrichment analysis was performed on DEGs, figure 3.6 provides a list of genes that were specifically downregulated in the stricture and their corresponding pathways that are important in inflammation and fibrosis. These include; desmoplakin (*DSP*) and Keratin 20 (*KRT20*) are important in wound healing and epithelial barrier function, chromogranin A (CHGA) and Kruppel-like factor 5 (KLF5) are important genes in cell differentiation and tissue remodelling. Their downregulation in the stricture suggests impaired function. On the contrary, a list of genes was shown as upregulated in the stricture compared to the proximal, these were important genes in driving the process of inflammation and fibrosis. These included Hypoxia-inducible factor 1 (*HIF1*) which modulates the HIF-1 signalling pathway that regulates hypoxia, *CCL2* which plays a major role in recruitment of monocytes to the inflamed site, endothelin receptor type B (*EDNRB)* which signals through endothelin 1 and 2 to promote proinflammatory and profibrotic mechanisms (569). Other important genes include *SEPINE1*, *PECAM1*, *LY96* and *CTSK*, these modulate processes such as adipogenesis, cellular metabolism and inflammation. Other upregulated genes in the stricture were linked to EMT, they include *COL1A1*, *COL4A1*, *COL4A2*, *COL9A1* and *FN1*. We also identified biomarkers of interest including *GREM1*, *HDAC1* and *FGF2*, these were upregulated in the stricture and align with what is published in the literature as these have been implicated in the pathogenesis of other fibrotic diseases disease (447,499,507,525).

Single cell RNA seq data was analysed to identify a diverse population of cells to a total of 14 populations consisting of both immune and stromal cells in the proximal and stricture segments. Previous studies have suggested that both T cells, B cells and cells of the innate

immune system including dendritic cells, macrophages, neutrophils are associated with development of strictures (570). The throughput of scRNAseq studies limits sample sizes; therefore, a total of 3 samples matched from the proximal and stricture were processed for sc-RNA-seq. Furthermore, *GREM1* expression was localised to the stromal cell cluster, this further validated the findings seen in the bulk RNA-Seq. In the literature, fibroblasts have been proposed as the main cell phenotype that drive the fibrotic milieu through fibroblast to myofibroblast transition (571). *GREM1* has been shown as upregulated in fibroblasts in other fibrotic conditions (520,553). It was therefore of interest to us to explore the role of *GREM1* in fibroblasts and intestinal strictures.

A key limiting factor for the sc-RNA seq data is that a total of 6,106 cells were used for the analysis as many cells were lost during quality control checks. This impacted further analysis to study sub populations which could have provided in depth information on cellular phenotypes in the stricture and proximal segments.

6.3 Localisation of stromal cells within the proximal, stricture and distal segments

IHC has long been used for identification of cells by protein expression, the use of FFPE sections to study protein expression provides an accurate representation of cellular location as the tissue architecture is preserved when fixed with formalin. The aim was to localise fibroblasts and myofibroblasts, being the main effector cells for tissue fibrosis and the highest producers of ECM proteins. At present, identification of fibroblasts is often based on positive staining with vimentin followed by elimination of other non-fibroblast mesenchymal cells (572). I used vimentin to identify stromal cells, and α-SMA staining to distinguish fibroblasts from myofibroblasts. I used desmin to distinguish myofibroblasts and smooth muscle cells. I was able to identify the location of intestinal stromal cells using markers such as vimentin, desmin and α-SMA. These cells were predominantly located within the lamina propria region

of the mucosa as seen in figures 4.3, 4.3 and 4.5. In addition, I used IHC to assess expression patterns for HDAC1, gremlin 1 and FGF2 in the proximal, stricture and distal segments. HDAC 1 staining was ubiquitous and expressed in multiple cells all over the tissue. This matched the expression pattern that was seen in the sc-RNA seq data. FGF2 staining was more specific to areas containing smooth muscle and around endothelial vessels. Gremlin 1 staining was very low and specific, mostly around the mucosal region. Overall, no conclusions can be drawn from these protein expression studies as no quantification was done. However, the data provided understanding into the expression patterns. In general, IHC is a great technique for use when studying a single marker but not effective for co-localisation studies which would require multiplex staining techniques.

6.4 Role of Gremlin 1 in stricture formation

Gremlin 1 has previously been found as elevated in various fibrotic tissue such as the liver, lung, eye and kidney (573,574). A study by Yang (2021) showed that *GREM1* levels were greatly increased in both human and murine fibrosed colon. Functional experiments revealed that *GREM1* promoted intestinal fibroblast cell proliferation (489). Our bulk RNA-seq showed an upregulation in *GREM1* gene expression in the stricture, these findings were validated using additional samples run by RT-PCR. In addition, sc-RNA seq data showed a localised expression of gremlin 1 in a cell population identified as stromal cells. Building upon these findings, I sought to examine the role of gremlin 1 in modulating fibroblast TGF-β signalling as a key pathway in intestinal fibrosis. Multiple protocols were tested for the isolation of primary fibroblasts, a hybrid protocol in which mechanical and enzymatic digestion was finally optimised for the isolation of fibroblasts. However due to limited cell numbers and limited access to primary tissue, an in vitro cell line CCD-18Co which is a gut derived fibroblast cell line was used to study the role of gremlin 1 in the activation of fibroblasts.

CCD-18Co were characterised by showing expression of stromal cell markers including vimentin, CD90, podoplanin, PDGFRβ and ACTA2 by flow cytometry and immunofluorescence. CCD-18Co were also tested for their response to pro-fibrotic factors such as IL-13, IL-33 and TGF-β, the data showed a response to TGF-β with increased expression of *COL1A1*, *ACTA2*, *FN1* and *TIMP1*. In addition, I showed increase in proliferation of CCD-18Co in the presence of TGF-β. These experiments provided validation in the use of CCD-18Co as an in vitro model due to the similarity in phenotype with primary fibroblasts as well characterised in the literature (575,576). Through the use of a *GREM1* siRNA knockdown model, I optimised the knock down effect of *GREM1* in CCD-18Co to 70%. In this model, I demonstrated a reduction in expression of *ACTA2* in the siRNA treated cells, however, this effect was counteracted in the presence of recombinant TGF-β as seen in figure 5.27C. I also demonstrated that gremlin 1 promotes proliferation as seen in figure 5.28 in which proliferation was reduced in siRNA treated CCD-18Co.

6.5 Conclusions

Data generated from this study suggests a role of gremlin 1 in fibrostenosing CD both at mRNA and protein level, it has been shown to influence fibroblast proliferation and activation. In addition, expression of gremlin 1 has been shown in stromal cells, particularly fibroblasts as seen in other tissues. Fibroblasts are also known as the main phenotype that drives tissue fibrosis; this makes fibroblasts an attractive cellular target. Based on these findings, I hypothesise that upregulated levels of gremlin 1 bind BMP2 and BMP4, hence inhibiting the BMP signalling pathway as seen in figure 6.1B, this causes an imbalance between TGF-β/BMP pathways that share a co-factor SMAD4. Therefore, inhibiting BMP signalling results in upregulation of TGF-β resulting in increased expression of profibrotic target genes. Inhibiting gremlin 1 restores that imbalance, this presents an attractive therapeutic target because;

1. It is an indirect means of modulating TGF-β as previous attempts to target TGF-β directly have resulted in off target effects.
2. Expression of gremlin 1 seems to be localised to fibroblasts.

Despite the promising findings from the study, I did not have adequate time to carry out further mechanistic studies, these are addressed in section 6.7. There is need for future studies to understand the mechanisms through which gremlin 1 modulates and drives a profibrotic phenotype especially in the TGF-β signalling pathway.

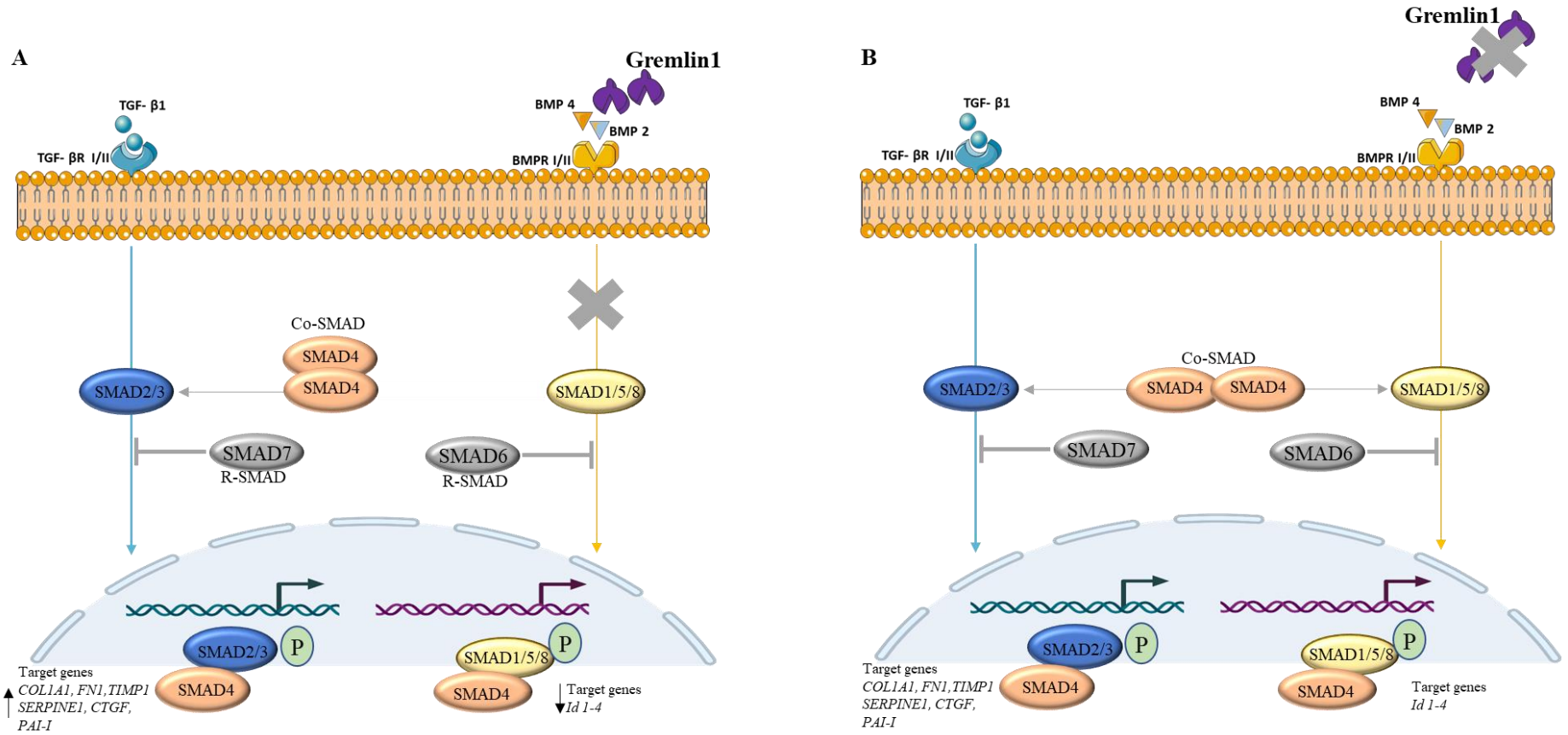


Figure 6.1: Regulation of TGF-β/BMP signalling pathways. **A)** TGF-β and BMP pathways signal through ligands TGF-β1 (TGF-β) and BMP2, BMP4 respectively through their cognate receptors. They share a co-factor SMAD4. Gremlin 1 binds BMP2 and BMP4, inhibiting BMP signalling and upregulating TGF-β signalling with increased availability of SMAD4 and transcription of profibrotic target genes such as *COL1A1*, *ACTA2*, *FN1* and *TIMP1*. Upregulation of gremlin 1 results in an imbalance between the two pathways **B)** Inhibition of gremlin 1 restores the balance between TGF-β and BMP pathways to regular transcription of target genes. *Image obtained from Biorender and modified.*

6.6 Limitations of the study

A total of 25 patient were recruited for the study, the study sample size was small, and this affected the robustness of the dataset that was generated. That was shown by the power calculations that were performed to determine the sample size required for a specific effect size. In addition, the 25 samples that were collected were distributed to perform various assays including optimisation for fibroblasts isolation, bulk and sc-RNA-seq, storage of tissue for RT-qPCR studies and formalin fixed paraffin embedded tissue sections (FFPE) for immunohistochemistry and immunofluorescence staining. As seen in the results, the data from patients' samples was variable, this may have been attributed to the method of sample collection. Tissue supplied to the lab was matched from each patient and supplied as three segments including proximal, stricture and distal; tissue classification was entirely dependent on the judgement of the pathologist based on physical examination at the time of tissue processing post-operative. There is a possibility that the non-fibrotic tissue might have been at an early stage of developing strictures which might be missed from a physical examination hence mis-classified. This may explain the variation seen across the different segments.

Another limitation to the study is the absence of a true healthy control, the proximal and distal segments were used as controls. However, these segments were still obtained from diseased gut and could have been undergoing inflammation or early onset of stricture formation which alters their phenotype similar to that seen in the stricture segment. Another limitation to the study was the lack of commercially available antibodies and reagents for the detection of gremlin 1. In this study, most assays assessed expression of gremlin 1 at a transcript level, assays that studied the translation to protein were limited and this limited in depth studies to be performed on the biomarker. Characterisation of stromal cells was also challenging, fibroblasts were the cell type of interest. However, their phenotype is very similar to that of other stromal cells such as pericytes. The stromal cell phenotyping panel that I used did not include a

definitive marker to distinguish the two cell types. In hindsight, markers such as FSP1-1 expressed on fibroblasts and myofibroblasts and RGS5 expressed on pericytes would have been ideal to include into the stromal cell phenotyping panel.

In chapter 4 of the results section, the IHC studies were qualitative, I was not able to quantify expression of the different markers that I stained for including vimentin, desmin and α-SMA. This was due to limited access to enough FFPE tissue sections due to low sample availability, this was one of the negative effects of the pandemic. In hindsight, the experimental plan for some of the assays carried out needed a lot more input, for example, more robust read out to assess the effects of *GREM1* siRNA knock down could have been included such as the use of expression of markers within BMP signalling (*BMP2*, *BMP4*, *BMPRI*, *BMPRII*, *ID1*, *ID2*).

I also underwent a transition with my supervision during the third year of my PhD, Prof. Subrata Ghosh who was my primary supervisor retired from the University of Birmingham and Dr. Christopher Weston had to step in to guide me through the third year of my PhD which also happened to be during the pandemic with limited access to the lab. There were quite a few changes to the direction of the project with the guidance from Dr Weston to ensure that I generated data that was robust and novel enough. Therefore, I had a lot of time constraint and challenges to get all of this data in time.

6.7 Future work

The challenges of undertaking this PhD during the pandemic limited the time spent in the lab and also impacted the access to resected tissue samples. I would have like to investigate the mechanisms through which gremlin 1 drives fibrosis, the additional work would shed more light on if gremlin can be used as a potential diagnostic marker or could it be a potential target for anti-fibrotic therapy in intestinal fibrosis. Further studies include;

- *GREM1* over expression; we overexpressed *GREM1* in CCD-18Co using the Lentivirus, however, the cells did not proliferate post cell sorting which may have been attributed to stress during the cells sorting process and driven them to early senescence. I would have like to optimise this procedure further to obtain a *GREM1* overexpressing phenotype as this would provide a phenotype to understand the effects of gremlin 1 overexpression in tipping the balance between TGFβ/BMP signalling pathways and how that drives a profibrotic phenotype.
- Utilise the *GREM1* siRNA knock down model for further mechanistic studies; Despite optimising the siRNA knockdown in which *GREM1* was knocked down by 70%, further mechanistic studies were not studied comprehensively. I would have liked to;
    - Study the effect on proliferation using other proliferation markers in addition to Ki-67such as proliferating cell nuclear antigen (PCNA) at both mRNA and protein level.
    - Using the scratch assay to assess cell migration
    - Downstream effect on expression of TGFβ/BMP signalling using a more comprehensive readout such as TGFβRI, TGFβRII, BMPRI, BMPRII, *IDI1* and *IDI2*, SMAD4 (co-factor for both pathways)
- In situ hybridisation (RNA Scope); one of the main challenges in this study was to detect expression of gremlin 1 in tissue due to the poor quality of antibodies in the market. The use of in situ hybridisation would have been a good option of determining rna expression in tissue to understand the cell phenotypes expressing *GREM1*.
- Multiplex immunofluorescence (mIF) and multispectral imaging; this application can be used on FFPE sections with a panel of at least 6 markers in which colocalization studies between target cells and biomarkers of interest are performed.
- Generation of a conditional Grem1 KO mouse model for in vivo studies

BIBLIOGRAPHY

1. Fakhoury M, Negrulj R, Mooranian A, Al-Salami H. Inflammatory bowel disease: Clinical aspects and treatments. Vol. 7, Journal of Inflammation Research. 2014.

2. Wehkamp J, Götz M, Herrlinger K, Steurer W, Stange EF. Inflammatory Bowel Disease: Crohn's disease and ulcerative colitis. Dtsch Arztebl Int [Internet]. 2016 Feb 5 [cited 2022 Feb 19];113(5):72. Available from: /pmc/articles/PMC4782273/

3. Noordzij M, Dekker FW, Zoccali C, Jager KJ. Measures of Disease Frequency: Prevalence and Incidence. Nephron Clin Pract [Internet]. 2010 Apr [cited 2023 Feb 10];115(1):c17–20. Available from: https://www.karger.com/Article/FullText/286345

4. The European Patients' Academy on Therapeutic Innovation (EUPATI). Epidemiologic concepts: incidence and prevalence [Internet]. 2018 [cited 2023 Feb 9]. Available from: https://toolbox.eupati.eu/resources/epidemiologic-concepts-incidence-and-prevalence/

5. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology. 2012;142(1).

6. Hodges P, Kelly P. Inflammatory bowel disease in Africa: What is the current state of knowledge? Int Health. 2020;12(3).

7. Jairath V, Feagan BG. Global burden of inflammatory bowel disease. Vol. 5, The Lancet Gastroenterology and Hepatology. Elsevier Ltd; 2020. p. 2–3.

8. Alatab S, Sepanlou SG, Ikuta K, Vahedi H, Bisignano C, Safiri S, et al. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Gastroenterol Hepatol. 2020;5(1).

9. Park KT, Ehrlich OG, Allen JI, Meadows P, Szigethy EM, Henrichsen K, et al. Erratum: The cost of inflammatory bowel disease: An initiative from the Crohn's & Colitis foundation (Inflammatory Bowel Diseases DOI: 10.1093/ibd/izz104). Vol. 26, Inflammatory Bowel Diseases. 2020.

10. Burisch J, Jess T, Martinato M, Lakatos PL. The burden of inflammatory bowel disease in Europe. Vol. 7, Journal of Crohn's and Colitis. 2013.

11. M'Koma AE. Inflammatory bowel disease: An expanding global health problem. Vol. 6, Clinical Medicine Insights: Gastroenterology. 2013.

12. Stiemsma L, Reynolds L, Turvey S, Finlay B. The hygiene hypothesis: current perspectives and future therapies. Immunotargets Ther. 2015;

13. Fletcher J, Cooper SC, Ghosh S, Hewison M. The role of vitamin D in inflammatory bowel disease: Mechanism to management. Vol. 11, Nutrients. 2019.

14. Kaplan GG, Windsor JW. The four epidemiological stages in the global evolution of inflammatory bowel disease. Vol. 18, Nature Reviews Gastroenterology and Hepatology. 2021.

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

16. Shrestha S, Olén O, Eriksson C, Everhov ÅH, Myrelid P, Visuri I, et al. The use of ICD codes to identify IBD subtypes and phenotypes of the Montreal classification in the Swedish National Patient Register. Scand J Gastroenterol. 2020;55(4).

17. Feakins RM. Ulcerative colitis or crohn's disease? Pitfalls and problems. Vol. 64, Histopathology. 2014.

18. Loddo I, Romano C. Inflammatory bowel disease: Genetics, epigenetics, and pathogenesis. Front Immunol. 2015;6(NOV):6–11.

19. Malik TA. Inflammatory Bowel Disease. Historical Perspective, Epidemiology, and Risk Factors. Vol. 95, Surgical Clinics of North America. 2015.

20. Mak WY, Zhao M, Ng SC, Burisch J. The epidemiology of inflammatory bowel disease: East meets west. Vol. 35, Journal of Gastroenterology and Hepatology (Australia). 2020.

21. Ahluwalia B, Moraes L, Magnusson MK, Öhman L. Immunopathogenesis of inflammatory bowel disease and mechanisms of biological therapies. Vol. 53, Scandinavian Journal of Gastroenterology. 2018.

22. Davies JM, Abreu MT. The innate immune system and inflammatory bowel disease. Vol. 50, Scandinavian Journal of Gastroenterology. 2014.

23. Holleran G, Lopetuso L, Petito V, Graziani C, Ianiro G, McNamara D, et al. The innate and adaptive immune system as targets for biologic therapies in inflammatory bowel disease. Vol. 18, International Journal of Molecular Sciences. MDPI AG; 2017.

24. Geremia A, Biancheri P, Allan P, Corazza GR, di Sabatino A. Innate and adaptive immunity in inflammatory bowel disease. Vol. 13, Autoimmunity Reviews. 2014. p. 3–10.

25. Zhang YZ, Li YY. Inflammatory bowel disease: Pathogenesis. World J Gastroenterol. 2014 Jan 7;20(1):91–9.

26. Onali, Favale, Fantini. The Resolution of Intestinal Inflammation: The Peace-Keeper's Perspective. Cells. 2019;8(4).

27. Momozawa Y, Dmitrieva J, Théâtre E, Deffontaine V, Rahmouni S, Charloteaux B, et al. IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes. Nat Commun. 2018 Dec 1;9(1).

28. Cleynen I, Boucher G, Jostins L, Schumm P, Zeissig S, Ahmad T, et al. Articles Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. The Lancet [Internet]. 2016 [cited 2023 Mar 28];387:156–67. Available from: http://dx.doi.org/10.1016/

29. Brant SR, Okou DT, Simpson CL, Cutler DJ, Haritunians T, Bradfield JP, et al. Genome-Wide Association Study Identifies African-Specific Susceptibility Loci in African Americans With Inflammatory Bowel Disease. Gastroenterology. 2017 Jan 1;152(1):206-217.e2.

30. Uhlig HH. Monogenic diseases associated with intestinal inflammation: Implications for the understanding of inflammatory bowel disease. Gut. 2013;62(12).

31. Greuter T, Vavricka SR. Extraintestinal manifestations in inflammatory bowel disease–epidemiology, genetics, and pathogenesis. Expert Rev Gastroenterol Hepatol. 2019;13(4).

32. Garber A, Regueiro M. Extraintestinal Manifestations of Inflammatory Bowel Disease: Epidemiology, Etiopathogenesis, and Management. Vol. 21, Current Gastroenterology Reports. 2019.

33. Rogler G, Singh A, Kavanaugh A, Rubin DT. Extraintestinal Manifestations of Inflammatory Bowel Disease: Current Concepts, Treatment, and Implications for Disease Management. Vol. 161, Gastroenterology. 2021.

34. Kim JM, Cheon JH. Pathogenesis and clinical perspectives of extraintestinal manifestations in inflammatory bowel diseases. Vol. 18, Intestinal Research. 2020.

35. Vavricka SR, Schoepfer A, Scharl M, Lakatos PL, Navarini A, Rogler G. Extraintestinal manifestations of inflammatory bowel disease. Vol. 21, Inflammatory Bowel Diseases. 2015.

36. Mertz A, Nguyen NA, Katsanos KH, Kwok RM. Primary sclerosing cholangitis and inflammatory bowel disease comorbidity: an update of the evidence. Ann Gastroenterol [Internet]. 2019 Mar 1 [cited 2023 Feb 11];32(2):124. Available from: /pmc/articles/PMC6394256/

37. Hedin CRH, Vavricka SR, Stagg AJ, Schoepfer A, Raine T, Puig L, et al. The Pathogenesis of Extraintestinal Manifestations: Implications for IBD Research, Diagnosis, and Therapy. J Crohns Colitis. 2019 Apr 26;13(5):541–54.

38. Schippers A, Hübel J, Heymann F, Clahsen T, Eswaran S, Schlepütz S, et al. MAdCAM-1/α4β7 Integrin-Mediated Lymphocyte/Endothelium Interactions Exacerbate Acute Immune-Mediated Hepatitis in Mice. CMGH. 2021 Jan 1;11(4):1227-1250.e1.

39. Graham JJ, Mukherjee S, Yuksel M, Sanabria Mateos R, Si T, Huang Z, et al. Aberrant hepatic trafficking of gut-derived T cells is not specific to primary sclerosing cholangitis. Hepatology. 2022 Mar 1;75(3):518–30.

40. Lalor PF, Edwards S, McNab G, Salmi M, Jalkanen S, Adams DH. Vascular adhesion protein-1 mediates adhesion and transmigration of lymphocytes on human hepatic endothelial cells. J Immunol [Internet]. 2002 Jul 15 [cited 2023 Feb 27];169(2):983–92. Available from: https://pubmed.ncbi.nlm.nih.gov/12097405/

41. Kociszewska D, Vlajkovic SM. The Association of Inflammatory Gut Diseases with Neuroinflammatory and Auditory Disorders. Vol. 14, Frontiers in Bioscience - Elite. Bioscience Research Institute; 2022.

42. Halling ML, Kjeldsen J, Knudsen T, Nielsen J, Hansen LK. Patients with inflammatory bowel disease have increased risk of autoimmune and inflammatory diseases. World J Gastroenterol. 2017;23(33).

43. González-Serna D, Ochoa E, López-Isac E, Julià A, Degenhardt F, Ortego-Centeno N, et al. A cross-disease meta-GWAS identifies four new susceptibility loci shared between systemic sclerosis and Crohn's disease. Sci Rep. 2020;10(1).

44. Argollo M, Gilardi D, Peyrin-Biroulet C, Chabot JF, Peyrin-Biroulet L, Danese S. Comorbidities in inflammatory bowel disease: a call for action. Vol. 4, The Lancet Gastroenterology and Hepatology. 2019.

45. Culver EL, Bungay HK, Betts M, Forde C, Buchel O, Manganis C, et al. Prevalence and long-term outcome of sub-clinical primary sclerosing cholangitis in patients with ulcerative colitis. Liver International. 2020;40(11).

46. Lee JM, Lee KM. Endoscopic diagnosis and differentiation of inflammatory bowel disease. Vol. 49, Clinical Endoscopy. 2016.

47. Ha CY, Newberry RD, Stone CD, Ciorba MA. Patients With Late-Adult-Onset Ulcerative Colitis Have Better Outcomes Than Those With Early Onset Disease. Clinical Gastroenterology and Hepatology. 2010;8(8).

48. Betteridge JD, Armbruster SP, Maydonovitch C, Veerappan GR. Inflammatory bowel disease prevalence by age, gender, race, and geographic location in the U.S. military health care population. Inflamm Bowel Dis. 2013;19(7).

49. Benchimol EI, MacK DR, Nguyen GC, Snapper SB, Li W, Mojaverian N, et al. Incidence, outcomes, and health services burden of very early onset inflammatory bowel disease. Gastroenterology. 2014 Oct 1;147(4):803-813.e7.

50. Snapper SB, Ouahed J, Spencer E, Kotlarz D, Shouval DS, Kowalik M, et al. Very Early Onset Inflammatory Bowel Disease: A Clinical Approach With a Focus on the Role of Genetics and Underlying Immune Deficiencies. Inflamm Bowel Dis • [Internet]. 2019 [cited 2023 Feb 28];26(6). Available from: https://academic.oup.com/ibdjournal/article/26/6/820/5675340

51. Conrad MA, Kelsen JR. Genomic and Immunologic Drivers of Very Early-Onset Inflammatory Bowel Disease. https://doi.org/101177/1093526619834807 [Internet]. 2019 Mar 6 [cited 2023 Feb 28];22(3):183–93. Available from: https://journals.sagepub.com/doi/epub/10.1177/1093526619834807

52. Zheng HB, de la Morena MT, Suskind DL. The Growing Need to Understand Very Early Onset Inflammatory Bowel Disease. Vol. 12, Frontiers in Immunology. Frontiers Media S.A.; 2021.

53. Kobayashi T, Siegmund B, le Berre C, Wei SC, Ferrante M, Shen B, et al. Ulcerative colitis. Nat Rev Dis Primers. 2020 Dec 1;6(1).

54. Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative colitis. Vol. 389, The Lancet. 2017.

55. Lamb CA, Kennedy NA, Raine T, Hendy PA, Smith PJ, Limdi JK, et al. British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. Gut [Internet]. 2019 Dec 1 [cited 2023 Feb 10];68(Suppl 3):s1. Available from: /pmc/articles/PMC6872448/

56. Yu YR, Rodriguez JR. Clinical presentation of Crohn's, ulcerative colitis, and indeterminate colitis: Symptoms, extraintestinal manifestations, and disease phenotypes. Semin Pediatr Surg. 2017;26(6).

57. DeRoche TC, Xiao SY, Liu X. Histological evaluation in ulcerative colitis. Vol. 2, Gastroenterology Report. 2014.

58. Conrad K, Roggenbuck D, Laass MW. Diagnosis and classification of ulcerative colitis. Vol. 13, Autoimmunity Reviews. 2014.

59. Dignass A, Eliakim R, Magro F, Maaser C, Chowers Y, Geboes K, et al. Second European evidence-based consensus on the diagnosis and management of ulcerative colitis Part 1: Definitions and diagnosis. Vol. 6, Journal of Crohn's and Colitis. 2012.

60. Rubin DT, Ananthakrishnan AN, Siegel CA, Sauer BG, Long MD. ACG Clinical Guideline: Ulcerative Colitis in Adults. American Journal of Gastroenterology. 2019;114(3).

61. Andrew RE, Messaris E. Update on medical and surgical options for patients with acute severe ulcerative colitis: What is new? World J Gastrointest Surg. 2016;8(9).

62. Sarlos P, Kovesdi E, Magyari L, Banfai Z, Szabo A, Javorhazy A, et al. Genetic update on inflammatory factors in ulcerative colitis: Review of the current literature. World J Gastrointest Pathophysiol. 2014;5(3).

63. Shi W, Zou R, Yang M, Mai L, Ren J, Wen J, et al. Analysis of genes involved in ulcerative colitis activity and tumorigenesis through systematic mining of gene co-expression networks. Front Physiol. 2019;10(MAY).

64. Halme L, Paavola-Sakki P, Turunen U, Lappalainen M, Färkkilä M, Kontula K. Family and twin studies in inflammatory bowel disease. Vol. 12, World Journal of Gastroenterology. 2006.

65. Ben-Horina S, Tamira S, Kopylova U, Katza L, Nadlera M, Langa A, et al. Familial ulcerative colitis in israeli jews: Its prevalence and clinical severity compared to sporadic disease. Ann Gastroenterol. 2011;24(4).

66. Gupta A, Bopanna S, Kedia S, Yadav DP, Goyal S, Jain S, et al. Familial aggregation of inflammatory bowel disease in patients with ulcerative colitis. Intest Res. 2017;15(3).

67. De Lange KM, Moutsianas L, Lee JC, Lamb CA, Luo Y, Kennedy NA, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. Nat Genet. 2017;49(2).

68. Marcin W, Michael S. Genetics and epigenetics of inflammatory bowel disease. Vol. 148, Swiss Medical Weekly. 2018.

69. Ashton JJ, Mossotto E, Stafford IS, Haggarty R, Coelho TAF, Batra A, et al. Genetic Sequencing of Pediatric Patients Identifies Mutations in Monogenic Inflammatory Bowel Disease Genes that Translate to Distinct Clinical Phenotypes. Clin Transl Gastroenterol. 2020;11(2).

70. Khorasani HM, Usefi H, Peña-Castillo L. Detecting ulcerative colitis from colon samples using efficient feature selection and machine learning. Sci Rep. 2020;10(1).

71. Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Vol. 43, Nature Genetics. 2011.

72. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature. 2012;491(7422).

73. Shi W, Zou R, Yang M, Mai L, Ren J, Wen J, et al. Analysis of genes involved in ulcerative colitis activity and tumorigenesis through systematic mining of gene co-expression networks. Front Physiol. 2019;10(MAY).

74. Barrett JC, Lee JC, Lees CW, Prescott NJ, Anderson CA, Phillips A, et al. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. Nat Genet. 2009;41(12).

75. Franke A, Balschun T, Sina C, Ellinghaus D, Häsler R, Mayr G, et al. Genome-wide association study for ulcerative colitis identifies risk loci at 7q22 and 22q13 (IL17REL). Nat Genet. 2010;42(4).

76. Potočnik U, Ferkolj I, Glavač D, Dean M. Polymorphisms in multidrug resistance 1 (MDR1) gene are associated with refractory Crohn disease and ulcerative colitis. Genes Immun. 2004;5(7).

77. Nam SY, Kim N, Kim JS, Lim SH, Jung HC, Song IS. Heat shock protein gene 70-2 polymorphism is differentially associated with the clinical phenotypes of ulcerative colitis and Crohn's disease. Journal of Gastroenterology and Hepatology (Australia). 2007;22(7).

78. Haritunians T, Taylor KD, Targan SR, Dubinsky M, Ippoliti A, Kwon S, et al. Genetic predictors of medically refractory ulcerative colitis. Inflamm Bowel Dis. 2010;16(11).

79. Hu S, Uniken Venema WT, Westra HJ, Vich Vila A, Barbieri R, Voskuil MD, et al. Inflammation status modulates the effect of host genetic variation on intestinal gene expression in inflammatory bowel disease. Nat Commun. 2021;12(1).

80. Cho JH, Brant SR. Recent insights into the genetics of inflammatory bowel disease. Gastroenterology. 2011;140(6).

81. Degenhardt F, Mayr G, Wendorff M, Boucher G, Ellinghaus E, Ellinghaus D, et al. Transethnic analysis of the human leukocyte antigen region for ulcerative colitis reveals not only shared but also ethnicity-specific disease associations. Hum Mol Genet. 2021;30(5).

82. Hartjes L, Ruland J. CARD9 signaling in intestinal immune homeostasis and oncogenesis. Vol. 10, Frontiers in Immunology. 2019.

83. Hsu YMS, Zhang Y, You Y, Wang D, Li H, Duramad O, et al. The adaptor protein CARD9 is required for innate immune responses to intracellular pathogens. Nat Immunol. 2007;8(2).

84. Azzam N, Nounou H, Alharbi O, Aljebreen A, Shalaby M. CARD15/NOD2, CD14 and toll-like 4 receptor gene polymorphisms in saudi patients with crohn's disease. Int J Mol Sci. 2012;13(4).

85. Wang Z, Fan R, Wang L, Zhou J, Zheng S, Hu S, et al. Genetic association between CARD9 variants and inflammatory bowel disease was not replicated in a Chinese Han population. Int J Clin Exp Pathol [Internet]. 2015 [cited 2023 Mar 18];8(10):13465. Available from: /pmc/articles/PMC4680503/

86. Zhernakova A, Festen EM, Franke L, Trynka G, van Diemen CC, Monsuur AJ, et al. Genetic Analysis of Innate Immunity in Crohn's Disease and Ulcerative Colitis Identifies Two Susceptibility Loci Harboring CARD9 and IL18RAP. Am J Hum Genet. 2008;82(5).

87. Yamamoto-Furusho JK, Fonseca-Camarillo G, Furuzawa-Carballeda J, Sarmiento-Aguilar A, Barreto-Zuñiga R, Martínez-Benitez B, et al. Caspase recruitment domain (CARD) family (CARD9, CARD10, CARD11, CARD14 and CARD15) are increased during active inflammation in patients with inflammatory bowel disease. Journal of Inflammation (United Kingdom). 2018;15(1).

88. Danne C, Michaudel C, Skerniskyte J, Planchais J, Magniez A, Agus A, et al. Inflammatory bowel disease CARD9 in neutrophils protects from colitis and controls mitochondrial metabolism and cell survival. Gut [Internet]. 2022 [cited 2023 Mar 18];0:1–12. Available from: http://gut.bmj.com/

89. Luo P, Yang Z, Chen B, Zhong X. The multifaceted role of CARD9 in inflammatory bowel disease. Vol. 24, Journal of Cellular and Molecular Medicine. 2020.

90. Kellermayer R. Epigenetics and the developmental origins of inflammatory bowel diseases. Vol. 26, Canadian Journal of Gastroenterology. 2012.

91. Rajamäki K, Taira A, Katainen R, Välimäki N, Kuosmanen A, Plaketti RM, et al. Genetic and Epigenetic Characteristics of Inflammatory Bowel Disease–Associated Colorectal Cancer. Gastroenterology. 2021;161(2).

92. Yi JM, Kim TO. Epigenetic Alterations in Inflammatory Bowel Disease and Cancer. Intest Res. 2015;13(2).

93. Häsler R, Feng Z, Bäckdahl L, Spehlmann ME, Franke A, Teschendorff A, et al. A functional methylome map of ulcerative colitis. Genome Res. 2012;22(11).

94. McDermott E, Ryan EJ, Tosetto M, Gibson D, Burrage J, Keegan D, et al. DNA methylation profiling in inflammatory bowel disease provides new insights into disease pathogenesis. J Crohns Colitis. 2016;10(1).

95. Kim YS, Ho SB. Intestinal goblet cells and mucins in health and disease: Recent insights and progress. Vol. 12, Current Gastroenterology Reports. 2010.

96. Williams JM, Duckworth CA, Burkitt MD, Watson AJM, Campbell BJ, Pritchard DM. Epithelial Cell Shedding and Barrier Function: A Matter of Life and Death at the Small Intestinal Villus Tip. Vet Pathol. 2015;52(3).

97. Birchenough GMH, Johansson MEV, Gustafsson JK, Bergström JH, Hansson GC. New developments in goblet cell mucus secretion and function. Vol. 8, Mucosal Immunology. 2015.

98. Saito S, Kato J, Hiraoka S, Horii J, Suzuki H, Higashi R, et al. DNA methylation of colon mucosa in ulcerative colitis patients: Correlation with inflammatory status. Inflamm Bowel Dis. 2011;17(9).

99. Meir M, Flemming S, Burkard N, Bergauer L, Metzger M, Germer CT, et al. Glial cell line-derived neurotrophic factor promotes barrier maturation and wound healing in intestinal epithelial cells in vitro. Am J Physiol Gastrointest Liver Physiol. 2015;309(8).

100. Meir M, Burkard N, Ungewiß H, Diefenbacher M, Flemming S, Kannapin F, et al. Neurotrophic factor GDNF regulates intestinal barrier function in inflammatory bowel disease. Journal of Clinical Investigation. 2019;129(7).

101. Schneider MR, Dahlhoff M, Horst D, Hirschi B, Trülzsch K, Müller-Höcker J, et al. A key role for E-cadherin in intestinal homeostasis and paneth cell maturation. PLoS One. 2010;5(12).

102. Ghishan FK, Kiela PR. Epithelial transport in inflammatory bowel diseases. Vol. 20, Inflammatory Bowel Diseases. 2014.

103. Wolfarth AA, Liu X, Darby TM, Boyer DJ, Spizman JB, Owens JA, et al. Proline-Rich Acidic Protein 1 (PRAP1) Protects the Gastrointestinal Epithelium From Irradiation-Induced Apoptosis. CMGH. 2020;10(4).

104. Zhang J, Wong H, Ramanan S, Cheong D, Leong A, Hooi SC. The Proline-Rich Acidic Protein Is Epigenetically Regulated and Inhibits Growth of Cancer Cell Lines. Cancer Res. 2003;63(20).

105. Belshaw NJ, Elliott GO, Foxall RJ, Dainty JR, Pal N, Coupe A, et al. Profiling CpG island field methylation in both morphologically normal and neoplastic human colonic mucosa. Br J Cancer. 2008;99(1).

106. Tahara T, Shibata T, Nakamura M, Yamashita H, Yoshioka D, Okubo M, et al. Effect of MDR1 gene promoter methylation in patients with ulcerative colitis. Int J Mol Med. 2009;23(4).

107. Taman H, Fenton CG, Hensel I V., Anderssen E, Florholmen J, Paulssen RH. Genome-wide DNA methylation intreatment-naïve ulcerative colitis. J Crohns Colitis. 2018;12(11).

108. Taman H, Fenton CG, Anderssen E, Florholmen J, Paulssen RH. DNA hypo-methylation facilitates anti-inflammatory responses in severe ulcerative colitis. PLoS One. 2021;16(4 April).

109. Coskun M, Bjerrum JT, Seidelin JB, Troelsen JT, Olsen J, Nielsen OH. miR-20b, miR-98, miR-125b-1*, and let-7e* as new potential diagnostic biomarkers in ulcerative colitis. World J Gastroenterol. 2013;19(27).

110. Min M, Peng L, Yang Y, Guo M, Wang W, Sun G. Microrna-155 is involved in the pathogenesis of ulcerative colitis by targeting FOXO3a. Inflamm Bowel Dis. 2014;20(4).

111. Snoeks L, Weber CR, Wasland K, Turner JR, Vainder C, Qi W, et al. Tumor suppressor FOXO3 participates in the regulation of intestinal inflammation. Laboratory Investigation. 2009;89(9).

112. Min M, Yang J, Yang YS, Liu Y, Liu LM, Xu Y. Expression of transcription factor FOXO3a is decreased in patients with ulcerative colitis. Chin Med J (Engl). 2015;128(20).

113. Zahm AM, Thayu M, Hand NJ, Horner A, Leonard MB, Friedman JR. Circulating microRNA is a biomarker of pediatric crohn disease. J Pediatr Gastroenterol Nutr. 2011;53(1).

114. Ng SC, Bernstein CN, Vatn MH, Lakatos PL, Loftus E V., Tysk C, et al. Geographical variability and environmental risk factors in inflammatory bowel disease. Gut. 2013;62(4).

115. Lakhan SE, Kirchgessner A. Anti-inflammatory effects of nicotine in obesity and ulcerative colitis. Vol. 9, Journal of Translational Medicine. 2011.

116. Pu W, Su Z, Wazir J, Zhao C, Wei L, Wang R, et al. Protective effect of α7 nicotinic acetylcholine receptor activation on experimental colitis and its mechanism. Molecular Medicine. 2022 Dec 1;28(1).

117. Sheikh SZ, Hegazi RA, Kobayashi T, Onyiah JC, Russo SM, Matsuoka K, et al. An Anti-Inflammatory Role for Carbon Monoxide and Heme Oxygenase-1 in Chronic Th2-Mediated Murine Colitis. The Journal of Immunology. 2011;186(9).

118. Saeed RW, Varma S, Peng-Nemeroff T, Sherry B, Balakhaneh D, Huston J, et al. Cholinergic stimulation blocks endothelial cell activation and leukocyte recruitment during inflammation. Journal of Experimental Medicine. 2005;201(7).

119. Monteleone I, Rizzo A, Sarra M, Sica G, Sileri P, Biancone L, et al. Aryl hydrocarbon receptor-induced signals up-regulate IL-22 production and inhibit inflammation in the gastrointestinal tract. Gastroenterology. 2011;141(1).

120. Bastida G, Beltrán B. Ulcerative colitis in smokers, non-smokers and ex-smokers. World J Gastroenterol. 2011;17(22).

121. Zhai H, Huang W, Liu A, Li Q, Hao Q, Ma L, et al. Current smoking improves ulcerative colitis patients' disease behaviour in the northwest of China. Prz Gastroenterol. 2017;12(4).

122. Piovani D, Danese S, Peyrin-Biroulet L, Nikolopoulos GK, Lytras T, Bonovas S. Environmental Risk Factors for Inflammatory Bowel Diseases: An Umbrella Review of Meta-analyses. Gastroenterology. 2019;157(3).

123. Myrelid P, Landerholm K, Nordenvall C, Pinkney TD, Andersson RE. Appendectomy and the Risk of Colectomy in Ulcerative Colitis: A National Cohort Study. American Journal of Gastroenterology. 2017;112(8).

124. Frisch M, Pedersen B V., Andersson REA. Appendicitis, mesenteric lymphadenitis, and subsequent risk of ulcerative colitis: cohort studies in Sweden and denmark. BMJ (Online). 2009;338(7698).

125. Andersson RE, Olaison G, Tysk C, Ekbom A. Appendectomy and Protection against Ulcerative Colitis. New England Journal of Medicine. 2001;344(11).

126. Sahami S, Kooij IA, Meijer SL, Van Den Brink GR, Buskens CJ, Te Velde AA. The link between the appendix and ulcerative colitis: Clinical relevance and potential immunological mechanisms. American Journal of Gastroenterology. 2016;111(2).

127. Watson Ng WS, Hampartzoumian T, Lloyd AR, Grimm MC. A murine model of appendicitis and the impact of inflammation on appendiceal lymphocyte constituents. Clin Exp Immunol. 2007;150(1).

128. Welsh S, Sam Z, Seenan JP, Nicholson GA. The Role of Appendicectomy in Ulcerative Colitis: Systematic Review and Meta-Analysis. Inflamm Bowel Dis. 2022 Jun 29;

129. Ananthakrishnan AN. Impact of diet on risk of IBD. Vol. 2, Crohn's and Colitis 360. 2020.

130. Abegunde AT, Muhammad BH, Bhatti O, Ali T. Environmental risk factors for inflammatory bowel diseases: Evidence based literature review. Vol. 22, World Journal of Gastroenterology. 2016.

131. Saidel-Odes L, Odes S. Hygiene hypothesis in inflammatory bowel disease. Vol. 27, Annals of Gastroenterology. 2014.

132. Koloski NA, Bret L, Radford-Smith G. Hygiene hypothesis in inflammatory bowel disease: A critical review of the literature. Vol. 14, World Journal of Gastroenterology. 2008.

133. Rook GAW, Bloomfield SF. Microbial exposures that establish immunoregulation are compatible with targeted hygiene. Journal of Allergy and Clinical Immunology. 2021;148(1).

134. Bloomfield SF, Stanwell-Smith R, Crevel RWR, Pickup J. Too clean, or not too clean: The Hygiene Hypothesis and home hygiene. Vol. 36, Clinical and Experimental Allergy. 2006.

135. Sood A, Amre D, Midha V, Sharma S, Sood N, Thara A, et al. Low hygiene and exposure to infections may be associated with increased risk for ulcerative colitis in a North Indian population. Ann Gastroenterol. 2014;27(3).

136. Leong RW, Mitrev N, Ko Y. Hygiene Hypothesis: Is the Evidence the Same All over the World? In: Digestive Diseases. 2016.

137. Ananthakrishnan AN. Epidemiology and risk factors for IBD. Vol. 12, Nature Reviews Gastroenterology and Hepatology. 2015.

138. Pei L yan, Ke Y shi, Zhao H hu, Wang L, Jia C, Liu W zhi, et al. Role of colonic microbiota in the pathogenesis of ulcerative colitis. BMC Gastroenterol. 2019;19(1).

139. Henrick BM, Rodriguez L, Lakshmikanth T, Pou C, Henckel E, Arzoomand A, et al. Bifidobacteria-mediated immune system imprinting early in life. Cell. 2021;184(15).

140. Planer JD, Peng Y, Kau AL, Blanton L v., Ndao IM, Tarr PI, et al. Development of the gut microbiota and mucosal IgA responses in twins and gnotobiotic mice. Nature. 2016;534(7606).

141. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, et al. Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci U S A. 2011;108(SUPPL. 1).

142. di Profio E, Magenes VC, Fiore G, Agostinelli M, la Mendola A, Acunzo M, et al. Special Diets in Infants and Children and Impact on Gut Microbioma. Vol. 14, Nutrients. MDPI; 2022.

143. Biasucci G, Benenati B, Morelli L, Bessi E, Boehm G. Cesarean delivery may affect the early biodiversity of intestinal bacteria. In: Journal of Nutrition. 2008.

144. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, et al. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc Natl Acad Sci U S A. 2010;107(26).

145. Mueller NT, Bakacs E, Combellick J, Grigoryan Z, Dominguez-Bello MG. The infant microbiome development: Mom matters. Vol. 21, Trends in Molecular Medicine. 2015.

146. Khodayar-Pardo P, Mira-Pascual L, Collado MC, Martínez-Costa C. Impact of lactation stage, gestational age and mode of delivery on breast milk microbiota. Journal of Perinatology. 2014;34(8).

147. Pannaraj PS, Li F, Cerini C, Bender JM, Yang S, Rollie A, et al. Association between breast milk bacterial communities and establishment and development of the infant gut microbiome. JAMA Pediatr. 2017;171(7).

148. Dong TS, Gupta A. Influence of Early Life, Diet, and the Environment on the Microbiome. Vol. 17, Clinical Gastroenterology and Hepatology. 2019.

149. Fjalstad JW, Esaiassen E, Juvet LK, van den Anker JN, Klingenberg C. Antibiotic therapy in neonates and impact on gut microbiota and antibiotic resistance development: A systematic review. Vol. 73, Journal of Antimicrobial Chemotherapy. 2018.

150. Yao Y, Cai X, Ye Y, Wang F, Chen F, Zheng C. The Role of Microbiota in Infant Health: From Early Life to Adulthood. Vol. 12, Frontiers in Immunology. 2021.

151. Shen ZH, Zhu CX, Quan YS, Yang ZY, Wu S, Luo WW, et al. Relationship between intestinal microbiota and ulcerative colitis: Mechanisms and clinical application of probiotics and fecal microbiota transplantation. Vol. 24, World Journal of Gastroenterology. 2018.

152. Pickard JM, Zeng MY, Caruso R, Núñez G. Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease. Vol. 279, Immunological Reviews. 2017.

153. Arumugam M, Raes J, Pelletier E, Paslier D le, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. Nature. 2011;473(7346).

154. Machiels K, Joossens M, Sabino J, de Preter V, Arijs I, Eeckhaut V, et al. A decrease of the butyrate-producing species roseburia hominis and faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. Gut. 2014;63(8).

155. Rowland I, Gibson G, Heinken A, Scott K, Swann J, Thiele I, et al. Gut microbiota functions: metabolism of nutrients and other food components. Vol. 57, European Journal of Nutrition. 2018.

156. Sun R, Xu C, Feng B, Gao X, Liu Z. Critical roles of bile acids in regulating intestinal mucosal immune responses. Vol. 14, Therapeutic Advances in Gastroenterology. 2021.

157. Staels B, Fonseca VA. Bile acids and metabolic regulation: mechanisms and clinical responses to bile acid sequestration. Vol. 32 Suppl 2, Diabetes care. 2009.

158. Gasaly N, de Vos P, Hermoso MA. Impact of Bacterial Metabolites on Gut Barrier Function and Host Immunity: A Focus on Bacterial Metabolism and Its Relevance for Intestinal Inflammation. Vol. 12, Frontiers in Immunology. 2021.

159. Yang M, Gu Y, Li L, Liu T, Song X, Sun Y, et al. Bile acid–gut microbiota axis in inflammatory bowel disease: From bench to bedside. Vol. 13, Nutrients. 2021.

160. Kriaa A, Mariaule V, Jablaoui A, Rhimi S, Mkaouar H, Hernandez J, et al. Bile Acids: Key Players in Inflammatory Bowel Diseases? Vol. 11, Cells. 2022.

161. Fiorucci S, Carino A, Baldoni M, Santucci L, Costanzi E, Graziosi L, et al. Bile Acid Signaling in Inflammatory Bowel Diseases. Vol. 66, Digestive Diseases and Sciences. 2021.

162. Sinha SR, Haileselassie Y, Nguyen LP, Tropini C, Wang M, Becker LS, et al. Dysbiosis-Induced Secondary Bile Acid Deficiency Promotes Intestinal Inflammation. Cell Host Microbe. 2020;27(4).

163. Staels B, Fonseca VA. Bile Acids and Metabolic Regulation. Diabetes Care. 2009;32(suppl_2).

164. Fiorucci S, Distrutti E. Bile Acid-Activated Receptors, Intestinal Microbiota, and the Treatment of Metabolic Disorders. Vol. 21, Trends in Molecular Medicine. 2015.

165. Tan J, McKenzie C, Potamitis M, Thorburn AN, Mackay CR, Macia L. The Role of Short-Chain Fatty Acids in Health and Disease. In: Advances in Immunology. 2014.

166. Thangaraju M, Cresci GA, Liu K, Ananth S, Gnanaprakasam JP, Browning DD, et al. GPFM 09A is a G-protein-coupled receptor for the bacterial fermentation product butyrate and functions as a tumor suppressor in colon. Cancer Res. 2009;69(7).

167. Wu J, Zhou Z, Hu Y, Dong S. Butyrate-induced GPR41 Activation Inhibits Histone Acetylation and Cell Growth. Journal of Genetics and Genomics. 2012;39(8).

168. Lee C, Kim BG, Kim JH, Chun J, Im JP, Kim JS. Sodium butyrate inhibits the NF-kappa B signaling pathway and histone deacetylation, and attenuates experimental colitis in an IL-10 independent manner. Int Immunopharmacol. 2017 Oct 1;51:47–56.

169. Zakerska-Banaszak O, Tomczak H, Gabryel M, Baturo A, Wolko L, Michalak M, et al. Dysbiosis of gut microbiota in Polish patients with ulcerative colitis: a pilot study. Sci Rep. 2021;11(1).

170. Varela E, Manichanh C, Gallart M, Torrejón A, Borruel N, Casellas F, et al. Colonisation by Faecalibacterium prausnitzii and maintenance of clinical remission in patients with ulcerative colitis. Aliment Pharmacol Ther. 2013;38(2).

171. Blanchaert C, Strubbe B, Peeters H. Fecal microbiota transplantation in ulcerative colitis. Vol. LXXXII, Acta Gastro-Enterologica Belgica.

172. Vieira ELM, Leonel AJ, Sad AP, Beltrão NRM, Costa TF, Ferreira TMR, et al. Oral administration of sodium butyrate attenuates inflammation and mucosal lesion in experimental acute ulcerative colitis. Journal of Nutritional Biochemistry. 2012 May;23(5):430–6.

173. Yalchin M, Segal JP, Mullish BH, Quraishi MN, Iqbal TH, Marchesi JR, et al. Gaps in knowledge and future directions for the use of faecal microbiota transplant in the treatment of inflammatory bowel disease. Vol. 12, Therapeutic Advances in Gastroenterology. 2019.

174. Landy J, Ronde E, English N, Clark SK, Hart AL, Knight SC, et al. Tight junctions in inflammatory bowel diseases and inflammatory bowel disease associated colorectal cancer. Vol. 22, World Journal of Gastroenterology. 2016.

175. Pastorelli L, Garg RR, Hoang SB, Spina L, Mattioli B, Scarpa M, et al. Epithelial-derived IL-33 and its receptor ST2 are dysregulated in ulcerative colitis and in experimental Th1/Th2 driven enteritis. Proc Natl Acad Sci U S A. 2010;107(17).

176. Brown M, Hughes KR, Moossavi S, Robins A, Mahida YR. Toll-like receptor expression in crypt epithelial cells, putative stem cells and intestinal myofibroblasts isolated from controls and patients with inflammatory bowel disease. Clin Exp Immunol. 2014;178(1).

177. Zhu W, Yu J, Nie Y, Shi X, Liu Y, Li F, et al. Disequilibrium of M1 and M2 macrophages correlates with the development of experimental inflammatory bowel diseases. Immunol Invest. 2014;43(7).

178. Bain CC, Schridde A. Origin, differentiation, and function of intestinal macrophages. Front Immunol. 2018;9(NOV).

179. Dharmasiri S, Garrido-Martin EM, Harris RJ, Bateman AC, Collins JE, Cummings JRF, et al. Human Intestinal Macrophages Are Involved in the Pathology of Both Ulcerative Colitis and Crohn Disease. Inflamm Bowel Dis. 2021;27(10).

180. Wright PB, McDonald E, Bravo-Blas A, Baer HM, Heawood A, Bain CC, et al. The mannose receptor (CD206) identifies a population of colonic macrophages in health and inflammatory bowel disease. Sci Rep. 2021;11(1).

181. Zhang HJ, Zhang YN, Zhou H, Guan L, Li Y, Sun MJ. IL-17A Promotes Initiation and Development of Intestinal Fibrosis Through EMT. Dig Dis Sci. 2018;63(11).

182. Wysoczanski R, Kendall AC, Motwani M, Vega R, Rahman FZ, McCartney S, et al. Ulcerative colitis is characterized by amplified acute inflammation with delayed resolution. bioRxiv. 2019;

183. Nemeth ZH, Bogdanovski DA, Barratt-Stopper P, Paglinco SR, Antonioli L, Rolandelli RH. Crohn's Disease and Ulcerative Colitis Show Unique Cytokine Profiles. Cureus. 2017;

184. Li J, Ueno A, Gasia MF, Luider J, Wang T, Hirota C, et al. Profiles of Lamina Propria T Helper Cell Subsets Discriminate between Ulcerative Colitis and Crohn's Disease. Inflamm Bowel Dis. 2016;22(8).

185. Seidelin JB, Bjerrum JT, Coskun M, Widjaya B, Vainer B, Nielsen OH. IL-33 is upregulated in colonocytes of ulcerative colitis. Immunol Lett. 2010;128(1).

186. Ngo Thi Phuong N, Palmieri V, Adamczyk A, Klopfleisch R, Langhorst J, Hansen W, et al. IL-33 Drives Expansion of Type 2 Innate Lymphoid Cells and Regulatory T Cells and Protects Mice From Severe, Acute Colitis. Front Immunol. 2021;12.

187. Danese S, Rudziński J, Brandt W, Dupas JL, Peyrin-Biroulet L, Bouhnik Y, et al. Tralokinumab for moderate-to-severe UC: A randomised, double-blind, placebo-controlled, phase IIa study. Gut. 2015;64(2).

188. Reinisch W, Panés J, Khurana S, Toth G, Hua F, Comer GM, et al. Anrukinzumab, an anti-interleukin 13 monoclonal antibody, in active UC: Efficacy and safety from a phase IIa randomised multicentre study. Gut. 2015;64(6).

189. Lamb CA, Mansfield JC, Tew GW, Gibbons D, Long AK, Irving P, et al. αEβ7 Integrin Identifies Subsets of Pro-Inflammatory Colonic CD4+ T Lymphocytes in Ulcerative Colitis. J Crohns Colitis. 2017;11(5).

190. Smids C, Horje CSHT, van Wijk F, van Lochem EG. The complexity of alpha E beta 7 blockade in inflammatory bowel diseases. Vol. 11, Journal of Crohn's and Colitis. 2017.

191. Russell SE, Horan RM, Stefanska AM, Carey A, Leon G, Aguilera M, et al. IL-36α expression is elevated in ulcerative colitis and promotes colonic inflammation. Mucosal Immunol. 2016;9(5).

192. Quezada SM, Steinberger EK, Cross RK. Association of age at diagnosis and crohn's disease phenotype. Age Ageing. 2013;42(1).

193. Saad AM, Czul F, Sakuraba A, Rubin DT, Cohen RD. Age of diagnosis is associated with disease presentation and therapeutic complications in patients with Crohn's disease. Inflamm Bowel Dis. 2016;22(5).

194. Torres J, Mehandru S, Colombel JF, Peyrin-Biroulet L. Crohn's disease. Vol. 389, The Lancet. 2017.

195. Li N, Shi RH. Updated review on immune factors in pathogenesis of Crohn?s disease. Vol. 24, World Journal of Gastroenterology. 2018.

196. Hoter A, Naim HY. The functions and therapeutic potential of heat shock proteins in inflammatory bowel disease—an update. Vol. 20, International Journal of Molecular Sciences. 2019.

197. Novak G, Parker CE, Pai RK, Macdonald JK, Feagan BG, Sandborn WJ, et al. Histologic scoring indices for evaluation of disease activity in Crohn's disease. Vol. 2017, Cochrane Database of Systematic Reviews. 2017.

198. Kellermann L, Riis LB. A close view on histopathological changes in inflammatory bowel disease, a narrative review. Dig Med Res. 2021;4.

199. Chang CW, Wong JM, Tung CC, Shih IL, Wang HY, Wei SC. Intestinal Stricture in Crohn's Disease. Intest Res [Internet]. 2015 [cited 2020 Jul 28];13(1):19. Available from: /pmc/articles/PMC4316217/?report=abstract

200. Stenke E, Bourke B, Knaus U. Crohn's Strictures—Moving Away from the Knife. Vol. 5, Frontiers in Pediatrics. 2017.

201. Feagan BG, Rochon J, Fedorak RN, Irvine EJ, Wild G, Sutherland L, et al. Methotrexate for the Treatment of Crohn's Disease. New England Journal of Medicine. 1995;332(5).

202. Colombel JF, Sandborn WJ, Reinisch W, Mantzaris GJ, Kornbluth A, Rachmilewitz D, et al. Infliximab, azathioprine, or combination therapy for Crohn's disease. N Engl J Med [Internet]. 2010 Apr 15 [cited 2023 Mar 9];362(15):1383–95. Available from: https://pubmed.ncbi.nlm.nih.gov/20393175/

203. Peyrin-Biroulet L, Loftus E v., Colombel JF, Sandborn WJ. The natural history of adult crohn's disease in population-based cohorts. Vol. 105, American Journal of Gastroenterology. 2010.

204. McLean LP, Cross RK. Integrin antagonists as potential therapeutic options for the treatment of Crohns disease. Vol. 25, Expert Opinion on Investigational Drugs. 2016.

205. Colombel JF, Sandborn WJ, Reinisch W, Mantzaris GJ, Kornbluth A, Rachmilewitz D, et al. Infliximab, azathioprine, or combination therapy for Crohn's disease. N Engl J Med [Internet]. 2010 Apr 15 [cited 2023 Mar 9];362(15):1383–95. Available from: https://pubmed.ncbi.nlm.nih.gov/20393175/

206. Roda G, Jharap B, Neeraj N, Colombel JF. Loss of Response to Anti-TNFs: Definition, Epidemiology, and Management. Vol. 7, Clinical and Translational Gastroenterology. 2016.

207. Rieder F, Zimmermann EM, Remzi FH, Sandborn WJ. Crohn's disease complicated by strictures: A systematic review. Gut. 2013;62(7):1072–84.

208. Bemelman WA, Warusavitarne J, Sampietro GM, Serclova Z, Zmora O, Luglio G, et al. ECCO-ESCP consensus on surgery for Crohn's disease. J Crohns Colitis. 2018;12(1).

209. Vagianos C, Malgarinos G, Spyropoulos C, Triantafillidis JK. Entero-vesical fistulas in CROHN'S disease: A case series report and review of the literature. Int J Surg Case Rep. 2017;41.

210. Xu X, Ma Y, Yao Z, Zhao Y. Prevalence and risk factors for pressure ulcers in patients with enterocutaneous fistula: A retrospective single-center study in China. Medical Science Monitor. 2019;25.

211. Mamie C, Bruckner RS, Lang S, Shpigel NY, Turina M, Rickenbacher A, et al. MMP9 expression in intestinal fistula from patients with fistulizing CD and from human xenograft mouse model. Tissue Barriers. 2021;

212. Ballard D, Erickson A, Ahuja C, Vea R, Sangster G, D'Agostino H. Percutaneous Management of Enterocutaneous Fistulae and Abscess–Fistula Complexes. Dig Dis Interv. 2018;02(02).

213. Liu T, Han L, Tilley M, Afzelius L, Maciejewski M, Jelinsky S, et al. Distinct clinical phenotypes for Crohn's disease derived from patient surveys. BMC Gastroenterol. 2021;21(1).

214. Halfvarson J, Jess T, Bodin L, Järnerot G, Munkholm P, Binder V, et al. Longitudinal concordance for clinical characteristics in a Swedish-Danish twin population with inflammatory bowel disease. Inflamm Bowel Dis. 2007;13(12).

215. Brant SR. Update on the heritability of inflammatory bowel disease: The importance of twin studies. Vol. 17, Inflammatory Bowel Diseases. 2011.

216. Halfvarson J. Genetics in twins with Crohn's disease: Less pronounced than previously believed? Inflamm Bowel Dis. 2011;17(1).

217. Jung S, Ye BD, Lee HS, Baek J, Kim G, Park D, et al. Identification of Three Novel Susceptibility Loci for Inflammatory Bowel Disease in Koreans in an Extended Genome-Wide Association Study. J Crohns Colitis. 2021;15(11).

218. Wang MH, Picco MF. Crohn's Disease: Genetics Update. Vol. 46, Gastroenterology Clinics of North America. 2017.

219. Sidiq T, Yoshihama S, Downs I, Kobayashi KS. Nod2: A critical regulator of ileal microbiota and Crohn's disease. Vol. 7, Frontiers in Immunology. 2016.

220. Horowitz JE, Warner N, Staples J, Crowley E, Gosalia N, Murchie R, et al. Mutation spectrum of NOD2 reveals recessive inheritance as a main driver of Early Onset Crohn's Disease. Sci Rep. 2021;11(1).

221. Yamamoto S, Ma X. Role of Nod2 in the development of Crohn's disease. Vol. 11, Microbes and Infection. 2009.

222. Brand S, Staudinger T, Schnitzler F, Pfennig S, Hofbauer K, Dambacher J, et al. The role of Toll-like receptor 4 Asp299Gly and Thr399Ile polymorphisms and CARD15/NOD2 mutations in the susceptibility and phenotype of Crohn's disease. Inflamm Bowel Dis. 2005;11(7).

223. Cheng Y, Zhu Y, Huang X, Zhang W, Han Z, Liu S. Association between TLR2 and TLR4 gene polymorphisms and the susceptibility to inflammatory bowel disease: A meta-analysis. PLoS One. 2015;10(5).

224. Tolentino YFM, Elia PP, Fogaça HS, Carneiro AJ V., Zaltman C, Moura-Neto R, et al. Common NOD2/CARD15 and TLR4 Polymorphisms Are Associated with Crohn's Disease Phenotypes in Southeastern Brazilians. Dig Dis Sci. 2016;61(9).

225. Naser SA, Arce M, Khaja A, Fernandez M, Naser N, Elwasila S, et al. Role of ATG16L, NOD2 and IL23R in Crohn's disease pathogenesis. Vol. 18, World Journal of Gastroenterology. 2012.

226. Hradsky O, Dusatkova P, Lenicek M, Bronsky J, Nevoral J, Vitek L, et al. The CTLA4 variants may interact with the IL23R- and NOD2-conferred risk in development of Crohn's disease. BMC Med Genet [Internet]. 2010 Jun 10 [cited 2023 Mar 10];11(1):91. Available from: /pmc/articles/PMC2894789/

227. Festen EAM, Goyette P, Green T, Boucher G, Beauchamp C, Trynka G, et al. A meta-analysis of genome-wide association scans identifies IL18RAP, PTPN2, TAGAP, and PUS10 as shared risk loci for crohn's disease and celiac disease. PLoS Genet. 2011;7(1).

228. Lees CW, Barrett JC, Parkes M, Satsangi J. New IBD genetics: Common pathways with other diseases. Vol. 60, Gut. 2011.

229. Gajendran M, Loganathan P, Catinella AP, Hashash JG. A comprehensive review and update on Crohn's disease. Disease-a-Month. 2018;64(2).

230. Ellinghaus D, Bethune J, Petersen BS, Franke A. The genetics of Crohn's disease and ulcerative colitis-status quo and beyond. Vol. 50, Scandinavian Journal of Gastroenterology. 2014.

231. Gupta S, Bortvin A, Vinogradova S, Shivdasani R, Gimelbrant A. CLONAL EPIGENETIC MOSAICISM: MECHANISM FOR CROHN'S DISEASE HETEROGENEITY. Gastroenterology. 2021;160(3).

232. Moret-Tatay I, Cerrillo E, Sáez-González E, Hervás D, Iborra M, Sandoval J, et al. Identification of Epigenetic Methylation Signatures With Clinical Value in Crohn's Disease. Clin Transl Gastroenterol. 2019;10(10).

233. Greuter T, Piller A, Fournier N, Safroneeva E, Straumann A, Biedermann L, et al. Upper gastrointestinal tract involvement in Crohn's disease: Frequency, risk factors, and disease course. J Crohns Colitis. 2018;12(12).

234. Alemany-Cosme E, Sáez-González E, Moret I, Mateos B, Iborra M, Nos P, et al. Oxidative stress in the pathogenesis of crohn's disease and the interconnection with immunological response, microbiota, external environmental factors, and epigenetics. Vol. 10, Antioxidants. 2021.

235. Kim TO, Park D Il, Han YK, Kang K, Park SG, Park HR, et al. Genome-wide analysis of the DNA methylation profile identifies the fragile histidine triad (FHIT) gene as a new promising biomarker of crohn's disease. J Clin Med. 2020;9(5).

236. Hornschuh M, Wirthgen E, Wolfien M, Singh KP, Wolkenhauer O, Däbritz J. The role of epigenetic modifications for the pathogenesis of Crohn's disease. Vol. 13, Clinical Epigenetics. 2021.

237. Ehrlich M. DNA hypermethylation in disease: mechanisms and clinical relevance. Vol. 14, Epigenetics. 2019.

238. Olaru A V., Selaru FM, Mori Y, Vazquez C, David S, Paun B, et al. Dynamic changes in the expression of MicroRNA-31 during inflammatory bowel disease-associated neoplastic transformation. Inflamm Bowel Dis. 2011;17(1).

239. Keith BP, Barrow JB, Toyonaga T, Kazgan N, O'Connor MH, Shah ND, et al. Colonic epithelial miR-31 associates with the development of Crohn's phenotypes. JCI Insight. 2018;3(19).

240. Ripamonti A, Provasi E, Lorenzo M, De Simone M, Ranzani V, Vangelisti S, et al. Repression of miR-31 by BCL6 stabilizes the helper function of human follicular helper T cells. Proc Natl Acad Sci U S A [Internet]. 2017 Nov 28 [cited 2023 Mar 10];114(48):12797–802. Available from: /pmc/articles/PMC5715737/

241. Schaefer JS, Attumi T, Opekun AR, Abraham B, Hou J, Shelby H, et al. MicroRNA signatures differentiate Crohn's disease from ulcerative colitis. BMC Immunol. 2015;16(1).

242. Chen Y, Wang Y, Shen J. Role of environmental factors in the pathogenesis of Crohn's disease: a critical review. Vol. 34, International Journal of Colorectal Disease. 2019.

243. Georgiou AN, Ntritsos G, Papadimitriou N, Dimou N, Evangelou E. Cigarette Smoking, Coffee Consumption, Alcohol Intake, and Risk of Crohn's Disease and Ulcerative Colitis: A Mendelian Randomization Study. Inflamm Bowel Dis. 2021;27(2).

244. Aldhous MC, Satsangi J. The impact of smoking in Crohn's disease: no smoke without fire. Frontline Gastroenterol. 2010;1(3).

245. Pasvol TJ, Bloom S, Segal AW, Rait G, Horsfall L. Use of contraceptives and risk of inflammatory bowel disease: a nested case–control study. Aliment Pharmacol Ther. 2022;55(3).

246. Langen ML van, Hotte N, Dieleman LA, Albert E, Mulder C, Madsen KL. Estrogen receptor-β signaling modulates epithelial barrier function. Am J Physiol Gastrointest Liver Physiol. 2011;300(4).

247. Braniste V, Jouault A, Gaultier E, Polizzi A, Buisson-Brenac C, Leveque M, et al. Impact of oral bisphenol a at reference doses on intestinal barrier function and sex differences after perinatal exposure in rats. Proc Natl Acad Sci U S A. 2010;107(1).

248. Khalili H, Higuchi LM, Ananthakrishnan AN, Richter JM, Feskanich D, Fuchs CS, et al. Oral contraceptives, reproductive factors and risk of inflammatory bowel disease. Gut. 2013;62(8).

249. Ng SC, Tang W, Leong RW, Chen M, Ko Y, Studd C, et al. Environmental risk factors in inflammatory bowel disease: A population-based case-control study in Asia-Pacific. Gut. 2015;64(7).

250. Vester-Andersen MK, Mirsepasi-Lauridsen HC, Prosberg M V., Mortensen CO, Träger C, Skovsen K, et al. Increased abundance of proteobacteria in aggressive Crohn's disease seven years after diagnosis. Sci Rep. 2019;9(1).

251. Nagpal R, Wang S, Solberg Woods LC, Seshie O, Chung ST, Shively CA, et al. Comparative microbiome signatures and short-chain fatty acids in mouse, rat, non-human primate, and human feces. Front Microbiol. 2018;9(NOV).

252. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. Nature. 2010;464(7285).

253. Forbes JD, Van Domselaar G, Bernstein CN. The gut microbiota in immune-mediated inflammatory diseases. Vol. 7, Frontiers in Microbiology. 2016.

254. Franzin M, Stefančič K, Lucafò M, Decorti G, Stocco G. Microbiota and drug response in inflammatory bowel disease. Vol. 10, Pathogens. 2021.

255. Zhang M, Sun K, Wu Y, Yang Y, Tso P, Wu Z. Interactions between Intestinal microbiota and host immune response in inflammatory bowel disease. Front Immunol. 2017;8(AUG).

256. Strober W, Watanabe T. NOD2, an intracellular innate immune sensor involved in host defense and Crohn's disease. Vol. 4, Mucosal Immunology. 2011.

257. Corridoni D, Chapman T, Ambrose T, Simmons A. Emerging mechanisms of innate immunity and their translational potential in inflammatory bowel disease. Vol. 5, Frontiers in Medicine. 2018.

258. Vignal C, Singer E, Peyrin-Biroulet L, Desreumaux P, Chamaillard M. How NOD2 mutations predispose to Crohn's disease? Vol. 9, Microbes and Infection. 2007.

259. Rescigno M, Iliev ID. Interleukin-23: Linking Mesenteric Lymph Node Dendritic Cells With Th1 Immunity in Crohn's Disease. Vol. 137, Gastroenterology. 2009.

260. Savić Mlakar A, Hojsak I, Jergović M, Čimić S, Bendelja K. Pediatric Crohn disease is characterized by Th1 in the terminal ileum and Th1/Th17 immune response in the colon. Eur J Pediatr. 2018;177(4).

261. Stolarczyk E, Lord GM, Howard JK. The immune cell transcription factor T-bet. Adipocyte. 2014;3(1).

262. Strober W, Zhang F, Kitani A, Fuss I, Fichtner-Feigl S. Proinflammatory cytokines underlying the inflammation of Crohn's disease. Vol. 26, Current Opinion in Gastroenterology. 2010.

263. Murugananthan AU, Bernardo D, Tozer P, Tee CT, Mann E, Hart AL, et al. TH1/TH17 profiles in crohn's disease: a cross sectional single centre study in postoperative crohn's disease. Gut. 2011;60(Suppl 1).

264. Bsat M, Chapuy L, Rubio M, Wassef R, Richard C, Schwenter F, et al. Differential pathogenic Th17 profile in mesenteric lymph nodes of Crohn's disease and ulcerative colitis patients. Front Immunol. 2019;10(MAY).

265. Schmitt H, Neurath MF, Atreya R. Role of the IL23/IL17 Pathway in Crohn's Disease. Vol. 12, Frontiers in Immunology. 2021.

266. Ramesh R, Kozhaya L, McKevitt K, Djuretic IM, Carlson TJ, Quintero MA, et al. Pro-inflammatory human Th17 cells selectively express P-glycoprotein and are refractory to glucocorticoids. Journal of Experimental Medicine. 2014;211(1).

267. Harbour SN, Maynard CL, Zindl CL, Schoeb TR, Weaver CT. Th17 cells give rise to Th1 cells that are required for the pathogenesis of colitis. Proc Natl Acad Sci U S A. 2015;112(22).

268. Veny M, Esteller M, Ricart E, PiquÉ JM, PanÉs J, Salas A. Late Crohn's disease patients present an increase in peripheral Th17 cells and cytokine production compared with early patients. Aliment Pharmacol Ther. 2010;31(5).

269. Feng T, Qin H, Wang L, Benveniste EN, Elson CO, Cong Y. Th17 Cells Induce Colitis and Promote Th1 Cell Responses through IL-17 Induction of Innate IL-12 and IL-23 Production. The Journal of Immunology. 2011;186(11).

270. Moschen AR, Tilg H, Raine T. IL-12, IL-23 and IL-17 in IBD: immunobiology and therapeutic targeting. Nat Rev Gastroenterol Hepatol [Internet]. 2018; Available from: http://www.nature.com/articles/s41575-018-0084-8

271. Khalili A, Ebrahimpour S, Maleki I, Abediankenari S, Afrouzi MM. CD4+CD25+CD127low FoxP3+ regulatory T cells in Crohn's disease. Rom J Intern Med. 2018;56(3).

272. Smids C, Horje CSHT, Drylewicz J, Roosenboom B, Groenen MJM, Koolwijk E van, et al. Intestinal T cell profiling in inflammatory bowel disease: Linking T cell subsets to disease activity and disease course. J Crohns Colitis. 2018;12(4).

273. Gordon IO, Agrawal N, Willis E, Goldblum JR, Lopez R, Allende D, et al. Fibrosis in ulcerative colitis is directly linked to severity and chronicity of mucosal inflammation. Aliment Pharmacol Ther. 2018;47(7).

274. Chan WPW, Mourad F, Leong RWL. Crohn's disease associated strictures. Vol. 33, Journal of Gastroenterology and Hepatology (Australia). 2018.

275. Bonnans C, Chou J, Werb Z. Remodelling the extracellular matrix in development and disease. Vol. 15, Nature Reviews Molecular Cell Biology. 2014.

276. Roulis M, Flavell RA. Fibroblasts and myofibroblasts of the intestinal lamina propria in physiology and disease. Vol. 92, Differentiation. 2016.

277. Bowers SLK, Banerjee I, Baudino TA. The extracellular matrix: At the center of it all. Vol. 48, Journal of Molecular and Cellular Cardiology. 2010.

278. Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. Vol. 123, Journal of Cell Science. 2010.

279. Speca S, Giusti I, Rieder F, Latella G. Cellular and molecular mechanisms of intestinal fibrosis. World J Gastroenterol. 2012;18(28):3635–61.

280. Herrera J, Henke CA, Bitterman PB. Extracellular matrix as a driver of progressive fibrosis. Vol. 128, Journal of Clinical Investigation. American Society for Clinical Investigation; 2018. p. 45–53.

281. Di Carlo SE, Peduto L. The perivascular origin of pathological fibroblasts. Vol. 128, Journal of Clinical Investigation. American Society for Clinical Investigation; 2018. p. 54–63.

282. Breynaert C, de Bruyn M, Arijs I, Cremer J, Martens E, Van Lommel L, et al. Genetic deletion of tissue inhibitor of metalloproteinase-1/TIMP-1 alters inflammation and attenuates fibrosis in dextran sodium sulphate-induced murine models of colitis. J Crohns Colitis. 2016;10(11).

283. Buechler MB, Pradhan RN, Krishnamurty AT, Cox C, Calviello AK, Wang AW, et al. Cross-tissue organization of the fibroblast lineage. Nature. 2021;593(7860).

284. Derkacz A, Olczyk P, Olczyk K, Komosinska-Vassev K. The role of extracellular matrix components in inflammatory bowel diseases. Vol. 10, Journal of Clinical Medicine. 2021.

285. Yoo JH, Holubar S, Rieder F. Fibrostenotic strictures in Crohn's disease. Intest Res. 2020;18(4).

286. Siri S, Zhao Y, Maier F, Pierce DM, Feng B. The macro-and micro-mechanics of the colon and rectum I: Experimental evidence. Vol. 7, Bioengineering. 2020.

287. Santos AJM, Lo YH, Mah AT, Kuo CJ. The Intestinal Stem Cell Niche: Homeostasis and Adaptations. Vol. 28, Trends in Cell Biology. 2018.

288. Pinchuk I V., Mifflin RC, Saada JI, Powell DW. Intestinal mesenchymal cells. Vol. 12, Current Gastroenterology Reports. 2010.

289. Thomson CA, Nibbs RJ, McCoy KD, Mowat AM. Immunological roles of intestinal mesenchymal cells. Vol. 160, Immunology. 2020.

290. Maimets M, Pedersen MT, Guiu J, Dreier J, Thodberg M, Antoku Y, et al. Mesenchymal-epithelial crosstalk shapes intestinal regionalisation via Wnt and Shh signalling. Nat Commun. 2022;13(1).

291. Albert-Bayo M, Paracuellos I, González-Castro AM, Rodríguez-Urrutia A, Rodríguez-Lagunas MJ, Alonso-Cotoner C, et al. Intestinal mucosal mast cells: Key modulators of barrier function and homeostasis. Vol. 8, Cells. 2019.

292. Hosoyamada Y, Sakai T. Structural and mechanical architecture of the intestinal villi and crypts in the rat intestine: Integrative reevaluation from ultrastructural analysis. Anat Embryol (Berl). 2005;210(1).

293. Epple HJ, Friebel J, Moos V, Troeger H, Krug SM, Allers K, et al. Architectural and functional alterations of the small intestinal mucosa in classical Whipple's disease. Mucosal Immunol. 2017;10(6).

294. Allaire JM, Crowley SM, Law HT, Chang SY, Ko HJ, Vallance BA. The Intestinal Epithelium: Central Coordinator of Mucosal Immunity. Vol. 39, Trends in Immunology. 2018.

295. Henderson NC, Rieder F, Wynn TA. Fibrosis: from mechanisms to medicines. Vol. 587, Nature. 2020.

296. Wang J, Lin S, Brown JM, van Wagoner D, Fiocchi C, Rieder F. Novel mechanisms and clinical trial endpoints in intestinal fibrosis*. Vol. 302, Immunological Reviews. 2021.

297. Li C, Kuemmerle JF. Mechanisms that mediate the development of fibrosis in patients with Crohn's disease. Vol. 20, Inflammatory Bowel Diseases. 2014.

298. de Bruyn JR, van den Brink GR, Steenkamer J, Buskens CJ, Bemelman WA, Meisner S, et al. Fibrostenotic phenotype of myofibroblasts in crohn's disease is dependent on tissue stiffness and reversed by LOX inhibition. J Crohns Colitis. 2018;12(7).

299. Truffi M, Sorrentino L, Monieri M, Fociani P, Mazzucchelli S, Bonzini M, et al. Inhibition of Fibroblast Activation Protein Restores a Balanced Extracellular Matrix and Reduces Fibrosis in Crohn's Disease Strictures Ex Vivo. Inflamm Bowel Dis. 2018;24(2).

300. Keihanian S, Moss AC. Crohn's disease stricture evaluation and management. Vol. 18, Techniques in Gastrointestinal Endoscopy. 2016.

301. Bettenworth D, Bokemeyer A, Baker M, Mao R, Parker CE, Nguyen T, et al. Assessment of Crohn's disease-associated small bowel strictures and fibrosis on cross-sectional imaging: A systematic review. Vol. 68, Gut. 2019.

302. Shivaji UN, Evans M, Critchlow T, Gui X, Smith SCL, Pinkney T, et al. Chronic inflammation and other changes are significant components of clinically fibrotic strictures in Crohn's disease: A histological study of resected strictures clinically characterized as noninflamed. Eur J Gastroenterol Hepatol. 2020;

303. Kaushal P, Somwaru AS, Charabaty A, Levy AD. MR enterography of inflammatory bowel disease with endoscopic correlation. Radiographics. 2017;37(1).

304. Medina C, Santos-Martinez MJ, Santana A, Paz-Cabrera MC, Johnston MJ, Mourelle M, et al. Transforming growth factor-beta type 1 receptor (ALK5) and Smad proteins mediate TIMP-1 and collagen synthesis in experimental intestinal fibrosis. Journal of Pathology. 2011;224(4):461–72.

305. Shelley-Fraser G, Borley NR, Warren BF, Shepherd NA. The connective tissue changes of Crohn's disease. Vol. 60, Histopathology. 2012.

306. Latella G, Rogler G, Bamias G, Breynaert C, Florholmen J, Pellino G, et al. Results of the 4th scientific workshop of the ECCO (I): Pathophysiology of intestinal fibrosis in IBD. Vol. 8, Journal of Crohn's and Colitis. 2014.

307. Fiocchi C, Kay Lund P. Themes in fibrosis and gastrointestinal inflammation. Vol. 300, American Journal of Physiology - Gastrointestinal and Liver Physiology. 2011.

308. Alfredsson J, Wick MJ. Mechanism of fibrosis and stricture formation in Crohn's disease. Vol. 92, Scandinavian Journal of Immunology. 2020.

309. Rogler G, Hausmann M. Factors promoting development of fibrosis in crohn's disease. Vol. 4, Frontiers in Medicine. 2017.

310. Lawrance IC, Rogler G, Bamias G, Breynaert C, Florholmen J, Pellino G, et al. Cellular and molecular mediators of intestinal fibrosis. Vol. 11, Journal of Crohn's and Colitis. 2017.

311. Friedrich M, Pohin M, Jackson MA, Korsunsky I, Bullers SJ, Rue-Albrecht K, et al. IL-1-driven stromal–neutrophil interactions define a subset of patients with inflammatory bowel disease that does not respond to therapies. Nat Med. 2021;

312. Thomas MG, Bayliss C, Bond S, Dowling F, Galea J, Jairath V, et al. 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. 2019;9(2).

313. Shouval DS, Biswas A, Kang YH, Griffith AE, Konnikova L, Mascanfroni ID, et al. Interleukin 1β Mediates Intestinal Inflammation in Mice and Patients With Interleukin 10 Receptor Deficiency. Gastroenterology. 2016;151(6).

314. Lee PY, Kumagai Y, Xu Y, Li Y, Barker T, Liu C, et al. IL-1α Modulates Neutrophil Recruitment in Chronic Inflammation Induced by Hydrocarbon Oil. The Journal of Immunology. 2011;186(3).

315. Schroder AL, Chami B, Liu Y, Doyle CM, El Kazzi M, Ahlenstiel G, et al. Neutrophil Extracellular Trap Density Increases With Increasing Histopathological Severity of Crohn's Disease. Inflamm Bowel Dis. 2021;

316. Ulfig A, Leichert LI. The effects of neutrophil-generated hypochlorous acid and other hypohalous acids on host and pathogens. Vol. 78, Cellular and Molecular Life Sciences. 2021.

317. Murray LA, Chen Q, Kramer MS, Hesson DP, Argentieri RL, Peng X, et al. TGF-beta driven lung fibrosis is macrophage dependent and blocked by Serum amyloid P. International Journal of Biochemistry and Cell Biology. 2011;43(1).

318. Isidro RA, Appleyard CB. Colonic macrophage polarization in homeostasis, inflammation, and cancer. Vol. 311, American Journal of Physiology - Gastrointestinal and Liver Physiology. 2016.

319. Lourenssen SR, Blennerhassett MG. M2 Macrophages and Phenotypic Modulation of Intestinal Smooth Muscle Cells Characterize Inflammatory Stricture Formation in Rats. American Journal of Pathology. 2020;190(9).

320. Han X, Ding S, Jiang H, Liu G. Roles of Macrophages in the Development and Treatment of Gut Inflammation. Vol. 9, Frontiers in Cell and Developmental Biology. 2021.

321. Salvador P, Macías-Ceja DC, Gisbert-Ferrándiz L, Hernández C, Bernardo D, Alós R, et al. CD16+ macrophages mediate fibrosis in inflammatory bowel disease. J Crohns Colitis. 2018;12(5).

322. Cheng D, Chai J, Wang H, Fu L, Peng S, Ni X. Hepatic macrophages: Key players in the development and progression of liver fibrosis. Vol. 41, Liver International. 2021.

323. Hou J, Shi J, Chen L, Lv Z, Chen X, Cao H, et al. M2 macrophages promote myofibroblast differentiation of LR-MSCs and are associated with pulmonary fibrogenesis. Cell Communication and Signaling. 2018;16(1).

324. Vijayan V, Pradhan P, Braud L, Fuchs HR, Gueler F, Motterlini R, et al. Human and murine macrophages exhibit differential metabolic responses to lipopolysaccharide - A divergent role for glycolysis. Redox Biol. 2019;22.

325. Mestas J, Hughes CCW. Of Mice and Not Men: Differences between Mouse and Human Immunology. The Journal of Immunology. 2004;172(5).

326. Wynn TA, Vannella KM. Macrophages in Tissue Repair, Regeneration, and Fibrosis. Vol. 44, Immunity. 2016.

327. Curciarello R, Docena GH, MacDonald TT. The role of cytokines in the fibrotic responses in crohn's disease. Vol. 4, Frontiers in Medicine. 2017.

328. Bailey JR, Bland PW, Tarlton JF, Peters I, Moorghen M, Sylvester PA, et al. IL-13 Promotes Collagen Accumulation in Crohn's Disease Fibrosis by Down-Regulation of Fibroblast MMP Synthesis: A Role for Innate Lymphoid Cells? PLoS One. 2012;7(12).

329. Walker JT, McLeod K, Kim S, Conway SJ, Hamilton DW. Periostin as a multifunctional modulator of the wound healing response. Vol. 365, Cell and Tissue Research. 2016.

330. Uchida M, Shiraishi H, Ohta S, Arima K, Taniguchi K, Suzuki S, et al. Periostin, a matricellular protein, plays a role in the induction of chemokines in pulmonary fibrosis. Am J Respir Cell Mol Biol. 2012;46(5).

331. Nanri Y, Nunomura S, Terasaki Y, Yoshihara T, Hirano Y, Yokosaki Y, et al. Cross-talk between transforming growth factor-b and periostin can be targeted for pulmonary fibrosis. Am J Respir Cell Mol Biol. 2020;62(2).

332. Bignold RE, Johnson JR. Matricellular Protein Periostin Promotes Pericyte Migration in Fibrotic Airways. Frontiers in Allergy. 2021;2.

333. Firszt R, Francisco D, Church TD, Thomas JM, Ingram JL, Kraft M. Interleukin-13 induces collagen type-1 expression through matrix metalloproteinase- 2 and transforming growth factor-β1 in airway fibroblasts in asthma. European Respiratory Journal. 2014;43(2).

334. Gour N, Wills-Karp M. IL-4 and IL-13 signaling in allergic airway disease. Vol. 75, Cytokine. 2015.

335. Reiman RM, Thompson RW, Feng CG, Hari D, Knight R, Cheever AW, et al. Interleukin-5 (IL-5) augments the progression of liver fibrosis by regulating IL-13 activity. Infect Immun. 2006;74(3):1471–9.

336. Drygiannakis I, Valatas V, Sfakianaki O, Bourikas L, Manousou P, Kambas K, et al. Proinflammatory cytokines induce crosstalk between colonic epithelial cells and subepithelial myofibroblasts: Implication in intestinal fibrosis. J Crohns Colitis. 2013;7(4).

337. Palomo J, Dietrich D, Martin P, Palmer G, Gabay C. The interleukin (IL)-1 cytokine family - Balance between agonists and antagonists in inflammatory diseases. Vol. 76, Cytokine. 2015.

338. Lawrance IC, Wu F, Leite AZA, Willis J, West GA, Fiocchi C, et al. A Murine Model of Chronic Inflammation-Induced Intestinal Fibrosis Down-Regulated by Antisense NF-κB. Gastroenterology. 2003;125(6).

339. Holvoet T, Devriese S, Castermans K, Boland S, Leysen D, Vandewynckel YP, et al. Treatment of Intestinal Fibrosis in Experimental Inflammatory Bowel Disease by the Pleiotropic Actions of a Local Rho Kinase Inhibitor. Gastroenterology. 2017;153(4).

340. Sedhom MAK, Pichery M, Murdoch JR, Foligné B, Ortega N, Normand S, et al. Neutralisation of the interleukin-33/ST2 pathway ameliorates experimental colitis through enhancement of mucosal healing In mice. Gut. 2013;62(12).

341. Mahapatro M, Foersch S, Hefele M, He GW, Giner-Ventura E, Mchedlidze T, et al. Programming of Intestinal Epithelial Differentiation by IL-33 Derived from Pericryptal Fibroblasts in Response to Systemic Infection. Cell Rep. 2016;15(8).

342. Yang Z, Sun R, Grinchuk V, Blanco JAF, Notari L, Bohl JA, et al. IL-33-induced alterations in murine intestinal function and cytokine responses are MyD88, STAT6, and IL-13 dependent. Am J Physiol Gastrointest Liver Physiol. 2013;304(4).

343. Scheibe K, Kersten C, Schmied A, Vieth M, Primbs T, Carlé B, et al. Inhibiting Interleukin 36 Receptor Signaling Reduces Fibrosis in Mice With Chronic Intestinal Inflammation. Gastroenterology. 2019;156(4).

344. Blumberg H, Dinh H, Trueblood ES, Pretorius J, Kugler D, Weng N, et al. Opposing activities of two novel members of the IL-1 ligand family regulate skin inflammation. Journal of Experimental Medicine. 2007;204(11).

345. Scheibe K, Backert I, Wirtz S, Hueber A, Schett G, Vieth M, et al. IL-36R signalling activates intestinal epithelial cells and fibroblasts and promotes mucosal healing in vivo. Gut. 2016;66(5).

346. Mountford S, Effenberger M, Noll-Puchta H, Griessmair L, Ringleb A, Haas S, et al. Modulation of Liver Inflammation and Fibrosis by Interleukin-37. Front Immunol. 2021;12.

347. Ramani K, Biswas PS. Interleukin-17: Friend or foe in organ fibrosis. Vol. 120, Cytokine. 2019.

348. Wilson MS, Madala SK, Ramalingam TR, Gochuico BR, Rosas IO, Cheever AW, et al. Bleomycin and IL-1β-mediated pulmonary fibrosis is IL-17A dependent. Journal of Experimental Medicine. 2010;207(3).

349. Wang L, Chen S, Xu K. IL-17 expression is correlated with hepatitis B-related liver diseases and fibrosis. Int J Mol Med. 2011;27(3).

350. Van Linthout S, Miteva K, Tschöpe C. Crosstalk between fibroblasts and inflammatory cells. Vol. 102, Cardiovascular Research. 2014.

351. Gaffen SL. Structure and signalling in the IL-17 receptor family. Vol. 9, Nature Reviews Immunology. 2009.

352. Honzawa Y, Nakase H, Matsumura K, Yamamoto S, Uza N, Matsuura M, et al. IL-17 Promotes HSP47 Expression and Intestinal Fibrosis in Crohn's Disease. Gastroenterology. 2011;140(5).

353. Honzawa Y, Nakase H, Shiokawa M, Yoshino T, Imaeda H, Matsuura M, et al. Involvement of interleukin-17A-induced expression of heat shock protein 47 in intestinal fibrosis in Crohn's disease. Gut. 2014;63(12).

354. Mi S, Li Z, Yang HZ, Liu H, Wang JP, Ma YG, et al. Blocking IL-17A Promotes the Resolution of Pulmonary Inflammation and Fibrosis Via TGF-β1–Dependent and –Independent Mechanisms. The Journal of Immunology. 2011;187(6).

355. Bechara R, McGeachy MJ, Gaffen SL. The metabolism-modulating activity of IL-17 signaling in health and disease. Vol. 218, Journal of Experimental Medicine. 2021.

356. Owens BMJ, Simmons A. Intestinal stromal cells in mucosal immunity and homeostasis. Vol. 6, Mucosal Immunology. 2013.

357. Schmelzer E, McKeel DT, Gerlach JC. Characterization of Human Mesenchymal Stem Cells from Different Tissues and Their Membrane Encasement for Prospective Transplantation Therapies. Biomed Res Int. 2019;2019.

358. Wilson AJ, Rand E, Webster AJ, Genever PG. Characterisation of mesenchymal stromal cells in clinical trial reports: analysis of published descriptors. Stem Cell Res Ther. 2021;12(1).

359. Molendijk I, Bonsing BA, Roelofs H, Peeters KCMJ, Wasser MNJM, Dijkstra G, et al. Allogeneic Bone Marrow - Derived Mesenchymal Stromal Cells Promote Healing of Refractory Perianal Fistulas in Patients With Crohn's Disease. Gastroenterology. 2015;149(4).

360. Lian L, Huang Q, Zhang L, Qin H, He X, He X, et al. Anti-fibrogenic Potential of Mesenchymal Stromal Cells in Treating Fibrosis in Crohn's Disease. Dig Dis Sci [Internet]. 2018;63(7):1821–34. Available from: https://doi.org/10.1007/s10620-018-5082-8

361. Dionne S, Restellini S, Koenekoop J, Salvador Escribano P, Piccirillo CA, Charlebois P, et al. Characterization of myofibroblasts isolated from the intestine of patients with inflammatory bowel disease. F1000Res. 2019;8.

362. Lahar N, Lei NY, Wang J, Jabaji Z, Tung SC, Joshi V, et al. Intestinal subepithelial myofibroblasts support in vitro and in vivo growth of human small intestinal epithelium. PLoS One. 2011;6(11).

363. Zidar N, Langner C, Jerala M, Boštjančič E, Drobne D, Tomažič A. Pathology of Fibrosis in Crohn's Disease—Contribution to Understanding Its Pathogenesis. Front Med (Lausanne). 2020;7.

364. Klingberg F, Hinz B, White ES. The myofibroblast matrix: Implications for tissue repair andfibrosis. Vol. 229, Journal of Pathology. 2013.

365. Valatas V, Filidou E, Drygiannakis I, Kolios G. Stromal and immune cells in gut fibrosis: The myofibroblast and the scarface. Vol. 30, Annals of Gastroenterology. 2017.

366. Darby IA, Laverdet B, Bonté F, Desmoulière A. Fibroblasts and myofibroblasts in wound healing. Vol. 7, Clinical, Cosmetic and Investigational Dermatology. 2014.

367. Micallef L, Vedrenne N, Billet F, Coulomb B, Darby IA, Desmoulière A. The myofibroblast, multiple origins for major roles in normal and pathological tissue repair. Fibrogenesis Tissue Repair. 2012;5(SUPPL.1).

368. Brannigan AE, Watson RWG, Beddy D, Hurley H, Fitzpatrick JM, O'Connell PR. Increased adhesion molecule expression in serosal fibroblasts isolated from patients with inflammatory bowel disease is secondary to inflammation. Ann Surg. 2002;235(4).

369. Okuno T, Andoh A, Bamba S, Araki Y, Fujiyama Y, Fujimiya M, et al. Interleukin-1β and tumor necrosis factor-α induce chemokine and matrix metalloproteinase gene expression in human colonic subepithelial myofibroblasts. Scand J Gastroenterol. 2002;37(3).

370. Holt AP, Salmon M, Buckley CD, Adams DH. Immune Interactions in Hepatic Fibrosis. Vol. 12, Clinics in Liver Disease. 2008.

371. Holt AP, Haughton EL, Lalor PF, Filer A, Buckley CD, Adams DH. Liver Myofibroblasts Regulate Infiltration and Positioning of Lymphocytes in Human Liver. Gastroenterology. 2009;136(2).

372. Brenner DA, Kisseleva T, Scholten D, Paik YH, Iwaisako K, Inokuchi S, et al. Origin of myofibroblasts in liver fibrosis. Fibrogenesis Tissue Repair. 2012;5(SUPPL.1).

373. Loubaki L, Hadj-Salem I, Fakhfakh R, Jacques E, Plante S, Boisvert M, et al. Co-culture of human bronchial fibroblasts and CD4+ T cells increases Th17 cytokine signature. PLoS One. 2013;8(12):2–13.

374. Swonger JM, Liu JS, Ivey MJ, Tallquist MD. Genetic tools for identifying and manipulating fibroblasts in the mouse. Vol. 92, Differentiation. 2016.

375. Karpus ON, Westendorp BF, Vermeulen JLM, Meisner S, Koster J, Muncan V, et al. Colonic CD90+ Crypt Fibroblasts Secrete Semaphorins to Support Epithelial Growth. Cell Rep. 2019;26(13).

376. Gomes RN, Manuel F, Nascimento DS. The bright side of fibroblasts: molecular signature and regenerative cues in major organs. Vol. 6, npj Regenerative Medicine. 2021.

377. Speca S, Rousseaux C, Dubuquoy C, Rieder F, Vetuschi A, Sferra R, et al. Novel PPARγ Modulator GED-0507-34 Levo Ameliorates Inflammation-driven Intestinal Fibrosis. Inflamm Bowel Dis. 2016;22(2).

378. D'Urso M, Kurniawan NA. Mechanical and Physical Regulation of Fibroblast–Myofibroblast Transition: From Cellular Mechanoresponse to Tissue Pathology. Vol. 8, Frontiers in Bioengineering and Biotechnology. 2020.

379. Santacroce G, Lenti MV, Di Sabatino A. Therapeutic Targeting of Intestinal Fibrosis in Crohn's Disease. Vol. 11, Cells. 2022.

380. Chen W, Lu C, Hirota C, Iacucci M, Ghosh S, Gui X. Smooth muscle hyperplasia/hypertrophy is the most prominent histological change in Crohn's fibrostenosing bowel strictures: A semiquantitative analysis by using a novel histological grading scheme. J Crohns Colitis. 2017;11(1):92–104.

381. Flynn RS, Murthy KS, Grider JR, Kellum JM, Kuemmerle JF. Endogenous IGF-I and αVβ3 Integrin Ligands Regulate Increased Smooth Muscle Hyperplasia in Stricturing Crohn's Disease. Gastroenterology. 2010;138(1).

382. Stanzel RDP, Lourenssen S, Nair DG, Blennerhassett MG. Mitogenic factors promoting intestinal smooth muscle cell proliferation. Am J Physiol Cell Physiol. 2010;299(4).

383. Ramirez M, Pell N, Mejias M, Fernandez M. Pericytes in the gut. In: Advances in Experimental Medicine and Biology. 2019.

384. Lin SL, Kisseleva T, Brenner DA, Duffield JS. Pericytes and perivascular fibroblasts are the primary source of collagen-producing cells in obstructive fibrosis of the kidney. American Journal of Pathology. 2008;173(6).

385. Humphreys BD. Targeting Pericyte Differentiation as a Strategy to Modulate Kidney Fibrosis in Diabetic Nephropathy. Semin Nephrol. 2012;32(5).

386. Wang N, Deng Y, Liu A, Shen N, Wang W, Du X, et al. Novel Mechanism of the Pericyte-Myofibroblast Transition in Renal Interstitial Fibrosis: Core Fucosylation Regulation. Sci Rep. 2017;7(1).

387. Kuppe C, Ibrahim MM, Kranz J, Zhang X, Ziegler S, Perales-Patón J, et al. Decoding myofibroblast origins in human kidney fibrosis. Nature. 2021;589(7841).

388. Sahebally SM, Burke JP, Chang KH, Kiernan MG, O'Connell PR, Coffey JC. Circulating fibrocytes and Crohn's disease. Vol. 100, British Journal of Surgery. 2013.

389. Grieb G, Bucala R. Fibrocytes in Fibrotic Diseases and Wound Healing. Adv Wound Care (New Rochelle). 2012;1(1).

390. Kuroda N, Masuya M, Tawara I, Tsuboi J, Yoneda M, Nishikawa K, et al. Infiltrating CCR2+ monocytes and their progenies, fibrocytes, contribute to colon fibrosis by inhibiting collagen degradation through the production of TIMP-1. Sci Rep. 2019;9(1).

391. Kuroda N, Masuya M, Tawara I, Tsuboi J, Yoneda M, Nishikawa K, et al. Infiltrating CCR2+ monocytes and their progenies, fibrocytes, contribute to colon fibrosis by inhibiting collagen degradation through the production of TIMP-1. Sci Rep. 2019;9(1).

392. Heukels P, van Hulst JAC, van Nimwegen M, Boorsma CE, Melgert BN, van den Toorn LM, et al. Fibrocytes are increased in lung and peripheral blood of patients with idiopathic pulmonary fibrosis. Respir Res. 2018;19(1).

393. Kany S, Vollrath JT, Relja B. Cytokines in inflammatory disease. Vol. 20, International Journal of Molecular Sciences. 2019.

394. Li DY, Li RF, Sun DX, Pu DD, Zhang YH. Mesenchymal stem cell therapy in pulmonary fibrosis: a meta-analysis of preclinical studies. Stem Cell Res Ther. 2021;12(1).

395. Rescigno M. The intestinal epithelial barrier in the control of homeostasis and immunity. Vol. 32, Trends in Immunology. 2011.

396. Chelakkot C, Ghim J, Ryu SH. Mechanisms regulating intestinal barrier integrity and its pathological implications. Vol. 50, Experimental and Molecular Medicine. 2018.

397. Wang MH, Sun R, Zhou XM, Zhang MY, Lu J Bin, Yang Y, et al. Epithelial cell adhesion molecule overexpression regulates epithelial-mesenchymal transition, stemness and metastasis of nasopharyngeal carcinoma cells via the PTEN/AKT/mTOR pathway. Cell Death Dis. 2018;9(1).

398. Gires O, Pan M, Schinke H, Canis M, Baeuerle PA. Expression and function of epithelial cell adhesion molecule EpCAM: where are we after 40 years? Vol. 39, Cancer and Metastasis Reviews. 2020.

399. Flier SN, Tanjore H, Kokkotou EG, Sugimoto H, Zeisberg M, Kalluri R. Identification of epithelial to mesenchymal transition as a novel source of fibroblasts in intestinal fibrosis. Journal of Biological Chemistry. 2010;285(26).

400. McVicker BL, Bennett RG. Novel anti-fibrotic therapies. Vol. 8, Frontiers in Pharmacology. 2017.

401. Lovisa S, Genovese G, Danese S. Role of Epithelial-to-Mesenchymal Transition in Inflammatory Bowel Disease. Vol. 13, Journal of Crohn's and Colitis. 2019.

402. Kraan J, Van Den Broek P, Verhoef C, Grunhagen DJ, Taal W, Gratama JW, et al. Endothelial CD276 (B7-H3) expression is increased in human malignancies and distinguishes between normal and tumour-derived circulating endothelial cells. Br J Cancer. 2014;111(1).

403. Goncharov N V., Nadeev AD, Jenkins RO, Avdonin P V. Markers and Biomarkers of Endothelium: When Something Is Rotten in the State. Vol. 2017, Oxidative Medicine and Cellular Longevity. 2017.

404. Habtezion A, Nguyen LP, Hadeiba H, Butcher EC. Leukocyte Trafficking to the Small Intestine and Colon. Vol. 150, Gastroenterology. 2016.

405. Rieder F, Kessler SP, West GA, Bhilocha S, De La Motte C, Sadler TM, et al. Inflammation-induced endothelial-to-mesenchymal transition: A novel mechanism of intestinal fibrosis. American Journal of Pathology. 2011;179(5).

406. Cho JG, Lee A, Chang W, Lee MS, Kim J. Endothelial to mesenchymal transition represents a key link in the interaction between inflammation and endothelial dysfunction. Vol. 9, Frontiers in Immunology. 2018.

407. Welch-Reardon KM, Wu N, Hughes CCW. A role for partial endothelial-mesenchymal transitions in angiogenesis? Arterioscler Thromb Vasc Biol. 2015;35(2).

408. Jackson AO, Zhang J, Jiang Z, Yin K. Endothelial-to-mesenchymal transition: A novel therapeutic target for cardiovascular diseases. Vol. 27, Trends in Cardiovascular Medicine. 2017.

409. Latella G, Rogler G, Bamias G, Breynaert C, Florholmen J, Pellino G, et al. Results of the 4th scientific workshop of the ECCO (I): Pathophysiology of intestinal fibrosis in IBD. Vol. 8, Journal of Crohn's and Colitis. 2014.

410. Latella G, Di Gregorio J, Flati V, Rieder F, Lawrance IC. Mechanisms of initiation and progression of intestinal fibrosis in IBD. Vol. 50, Scandinavian Journal of Gastroenterology. 2014.

411. Johnson LA, Luke A, Sauder K, Moons DS, Horowitz JC, Higgins PDR. Intestinal fibrosis is reduced by early elimination of inflammation in a mouse model of IBD:

Impact of a “top-Down” approach to intestinal fibrosis in mice. Inflamm Bowel Dis. 2012;18(3).

412. Yang B, Zhang G, Elias M, Zhu Y, Wang J. The role of cytokine and immune responses in intestinal fibrosis. Vol. 21, Journal of Digestive Diseases. 2020.

413. Di Carmine S, Scott MM, Mclean MH, Mcsorley HJ, Carmine D. The role of interleukin-33 in organ fibrosis. Discovery Immunology [Internet]. 2022 Jan 1 [cited 2023 Mar 18];1(1):1–11. Available from: https://academic.oup.com/discovimmunology/article/1/1/kyac006/6717952

414. Wu YS, Chen SN. Apoptotic cell: Linkage of inflammation and wound healing. Vol. 5 JAN, Frontiers in Pharmacology. 2014.

415. Li J, Mao R, Kurada S, Wang J, Lin S, Chandra J, et al. Pathogenesis of fibrostenosing Crohn’s disease. Translational Research [Internet]. 2019;209:39–54. Available from: https://doi.org/10.1016/j.trsl.2019.03.005

416. di Sabatino A, Jackson CL, Pickard KM, Buckley M, Rovedatti L, Leakey NAB, et al. Transforming growth factor β signalling and matrix metalloproteinases in the mucosa overlying Crohn’s disease strictures. Gut. 2009;58(6):777–89.

417. Troncone E, Marafini I, Stolfi C, Monteleone G. Transforming growth factor-β1/Smad7 in intestinal immunity, inflammation, and cancer. Vol. 9, Frontiers in Immunology. 2018.

418. Stolfi C, Troncone E, Marafini I, Monteleone G. Role of tgf-beta and smad7 in gut inflammation, fibrosis and cancer. Vol. 11, Biomolecules. 2021.

419. Izzo R, Bevivino G, de Simone V, Sedda S, Monteleone I, Marafini I, et al. Knockdown of Smad7 with a Specific Antisense Oligonucleotide Attenuates Colitis and Colitis-Driven Colonic Fibrosis in Mice. Inflamm Bowel Dis. 2018;24(6).

420. Terada Y, Hanada S, Nakao A, Kuwahara M, Sasaki S, Marumo F. Gene transfer of Smad7 using electroporation of adenovirus prevents renal fibrosis in post-obstructed kidney. Kidney Int. 2002;61(1).

421. Di Mola FF, Di Sebastiano P, Gardini A, Innocenti P, Zimmermann A, Büchler M, et al. Differential expression of connective tissue growth factor in inflammatory bowel disease. Digestion. 2004;69(4).

422. Beddy D, Mulsow J, Watson RWG, Fitzpatrick JM, O’Connell PR. Expression and regulation of connective tissue growth factor by transforming growth factor β and tumour necrosis factor α in fibroblasts isolated from strictures in patients with Crohn’s disease. British Journal of Surgery. 2006;93(10).

423. Andoh A, Nishida A. Molecular Basis of Intestinal Fibrosis in Inflammatory Bowel Disease. 2022 [cited 2023 Mar 18]; Available from: www.karger.com/iid

424. Iwayama T, Olson LE. Involvement of PDGF in fibrosis and scleroderma: Recent insights from animal models and potential therapeutic opportunities. Curr Rheumatol Rep. 2013;15(2).

425. Rieder F, Fiocchi C. Intestinal fibrosis in inflammatory bowel disease - Current knowledge and future perspectives. Vol. 2, Journal of Crohn's and Colitis. 2008.

426. Kuhnert F, Davis CR, Wang HT, Chu P, Lee M, Yuan J, et al. Essential requirement for Wnt signaling in proliferation of adult small intestine and colon revealed by adenoviral expression of Dickkopf-1. Proc Natl Acad Sci U S A. 2004;101(1).

427. Lam AP, Gottardi CJ. β-catenin signaling: A novel mediator of fibrosis and potential therapeutic target. Vol. 23, Current Opinion in Rheumatology. 2011.

428. Barry-Hamilton V, Spangler R, Marshall D, McCauley S, Rodriguez HM, Oyasu M, et al. Allosteric inhibition of lysyl oxidase-like-2 impedes the development of a pathologic microenvironment. Nat Med. 2010;16(9).

429. Rieder F, Latella G, Magro F, Yuksel ES, Higgins PDR, Di Sabatino A, et al. European Crohn's and colitis organisation topical review on prediction, diagnosis and management of fibrostenosing Crohn's disease. J Crohns Colitis. 2016;10(8).

430. Luong VH, Chino T, Oyama N, Matsushita T, Sasaki Y, Ogura D, et al. Blockade of TGF-β/Smad signaling by the small compound HPH-15 ameliorates experimental skin fibrosis. Arthritis Res Ther. 2018;20(1).

431. Chung ACK, Dong Y, Yang W, Zhong X, Li R, Lan HY. Smad7 suppresses renal fibrosis via altering expression of TGF-β/Smad3-regulated microRNAs. Molecular Therapy. 2013;21(2).

432. Cui Y, Zhang M, Leng C, Blokzijl T, Jansen BH, Dijkstra G, et al. Pirfenidone Inhibits Cell Proliferation and Collagen I Production of Primary Human Intestinal Fibroblasts. Cells. 2020;9(3).

433. Honig G, Larkin PB, Heller C, Hurtado-Lorenzo A. Research-based product innovation to address critical unmet needs of patients with inflammatory bowel diseases. Inflamm Bowel Dis. 2021;27.

434. Sleiman J, El Ouali S, Qazi T, Cohen B, Steele SR, Baker ME, et al. Prevention and Treatment of Stricturing Crohn's Disease–Perspectives and Challenges. Vol. 15, Expert Review of Gastroenterology and Hepatology. 2021.

435. Rieder F, Kessler S, Sans M, Fiocchi C. Animal models of intestinal fibrosis: New tools for the understanding of pathogenesis and therapy of human disease. Vol. 303, American Journal of Physiology - Gastrointestinal and Liver Physiology. 2012.

436. Kendall RT, Feghali-Bostwick CA. Fibroblasts in fibrosis: Novel roles and mediators. Vol. 5 MAY, Frontiers in Pharmacology. 2014.

437. Verstockt B, Cleynen I. Genetic influences on the development of fibrosis in Crohn's disease [Internet]. Vol. 3, Frontiers in Medicine. Frontiers Media S.A.; 2016 [cited 2020 Jul 28]. p. 1. Available from: www.frontiersin.org

438. Hoffmann P, Lamerz D, Hill P, Kirchner M, Gauss A. Gene polymorphisms of nod2, il23r, ptpn2 and atg16l1 in patients with crohn's disease: On the way to personalized medicine? Genes (Basel). 2021;12(6).

439. Shih DQ, Barrett R, Zhang X, Yeager N, Koon HWK, Phaosawasdi P, et al. Constitutive TL1A (TNFSF15) expression on lymphoid or myeloid cells leads to mild intestinal inflammation and fibrosis. PLoS One. 2011;6(1).

440. Ray S, De Salvo C, Pizarro TT. Central role of IL-17/Th17 immune responses and the gut microbiota in the pathogenesis of intestinal fibrosis. Vol. 30, Current Opinion in Gastroenterology. 2014.

441. Li C, Kuemmerle J. P-147 Epigenetic Mechanisms Inhibit Smad7 and Contribute to Fibrosis in Crohn's Disease. Inflamm Bowel Dis. 2016;22.

442. Zeng Z, Mukherjee A, Zhang H. From genetics to epigenetics, roles of epigenetics in inflammatory bowel disease. Front Genet. 2019;10(OCT).

443. Nimmo ER, Prendergast JG, Aldhous MC, Kennedy NA, Henderson P, Drummond HE, et al. Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway. Inflamm Bowel Dis. 2012;18(5).

444. Ventham NT, Kennedy NA, Adams AT, Kalla R, Heath S, O'Leary KR, et al. Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. Nat Commun. 2016;7.

445. Tsaprouni LG, Ito K, Powell JJ, Adcock IM, Punchard N. Differential patterns of histone acetylation in inflammatory bowel diseases. J Inflamm. 2011;8.

446. Lyu X, Hu M, Peng J, Zhang X, Sanders YY. HDAC inhibitors as antifibrotic drugs in cardiac and pulmonary fibrosis. Vol. 10, Therapeutic Advances in Chronic Disease. 2019.

447. Yoon S, Kang G, Eom GH. HDAC Inhibitors: Therapeutic Potential in Fibrosis-Associated Human Diseases. International Journal of Molecular Sciences 2019, Vol 20, Page 1329 [Internet]. 2019 Mar 16 [cited 2022 May 9];20(6):1329. Available from: https://www.mdpi.com/1422-0067/20/6/1329/htm

448. Nie L, Liu Y, Zhang B, Zhao J. Application of Histone Deacetylase Inhibitors in Renal Interstitial Fibrosis. Kidney Diseases. 2020;6(4).

449. Masterson JC, Capocelli KE, Hosford L, Biette K, McNamee EN, De Zoeten EF, et al. Eosinophils and IL-33 Perpetuate Chronic Inflammation and Fibrosis in a Pediatric Population with Stricturing Crohn's Ileitis. Inflamm Bowel Dis. 2015;21(10).

450. Laukens D. Inflammation-independent mechanisms of intestinal fibrosis: The role of the extracellular matrix. In: Fibrostenotic Inflammatory Bowel Disease. 2018.

451. Rieder F. Toward an antifibrotic therapy for inflammatory bowel disease. United European Gastroenterol J. 2016;4(4).

452. Bettenworth D, Rieder F. Pathogenesis of Intestinal Fibrosis in Inflammatory Bowel Disease and Perspectives for Therapeutic Implication. Digestive Diseases. 2017;35(1–2):25–31.

453. Bettenworth D, Nowacki TM, Cordes F, Buerke B, Lenze F. Assessment of stricturing Crohn's disease: Current clinical practice and future avenues 2016 Inflammatory Bowel Disease: Global view. Vol. 22, World Journal of Gastroenterology. 2016.

454. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods. 2009;6(5).

455. Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. Vol. 50, Experimental and Molecular Medicine. 2018.

456. Gautam A, Donohue D, Hoke A, Miller SA, Srinivasan S, Sowe B, et al. Investigating gene expression profiles of whole blood and peripheral blood mononuclear cells using multiple collection and processing methods. PLoS One. 2019;14(12).

457. Lever J, Krzywinski M, Altman N. Points of Significance: Principal component analysis. Vol. 14, Nature Methods. 2017.

458. Blasco H, Błaszczyński J, Billaut JC, Nadal-Desbarats L, Pradat PF, Devos D, et al. Comparative analysis of targeted metabolomics: Dominance-based rough set approach versus orthogonal partial least square-discriminant analysis. J Biomed Inform. 2015;53.

459. Meng XM, Nikolic-Paterson DJ, Lan HY. TGF-β: The master regulator of fibrosis. Vol. 12, Nature Reviews Nephrology. 2016.

460. Meng XM, Chung ACK, Lan HY. Role of the TGF-β/BMP-7/Smad pathways in renal diseases. Vol. 124, Clinical Science. 2013.

461. Hu HH, Chen DQ, Wang YN, Feng YL, Cao G, Vaziri ND, et al. New insights into TGF-β/Smad signaling in tissue fibrosis. Chem Biol Interact. 2018;292:76–83.

462. Li A, Qin G, Suzuki A, Gajera M, Iwata J, Jia P, et al. Network-based identification of critical regulators as putative drivers of human cleft lip. BMC Med Genomics. 2019;12.

463. Xiong A, Liu Y. Targeting hypoxia inducible factors-1α as a novel therapy in Fibrosis. Vol. 8, Frontiers in Pharmacology. 2017.

464. Xu Z, Castellino FJ, Ploplis VA. Plasminogen activator inhibitor-1 (PAI-1) is cardioprotective in mice by maintaining microvascular integrity and cardiac architecture. Blood. 2010;115(10).

465. Guo W, Shan B, Klingsberg RC, Qin X, Lasky JA. Abrogation of TGF-β1-induced fibroblast-myofibroblast differentiation by histone deacetylase inhibition. Am J Physiol Lung Cell Mol Physiol. 2009;297(5).

466. Zhang Y, Gao F, Tang Y, Xiao J, Li C, Ouyang Y, et al. Valproic acid regulates Ang II-induced pericyte-myofibroblast trans-differentiation via MAPK/ERK pathway. Am J Transl Res [Internet]. 2018 [cited 2023 Mar 21];10(7):1976. Available from: /pmc/articles/PMC6079132/

467. Kolb M, Bonella F, Wollin L. Therapeutic targets in idiopathic pulmonary fibrosis. Vol. 131, Respiratory Medicine. 2017.

468. Yu L, Fernandez S, Brock G. Power analysis for RNA-Seq differential expression studies. BMC Bioinformatics. 2017;18(1).

469. Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation. Vol. 20, Emergency Medicine Journal. 2003.

470. Suresh K, Chandrashekara S. Sample size estimation and power analysis for clinical research studies. Vol. 5, Journal of Human Reproductive Sciences. 2012.

471. Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. Nat Methods. 2014;11(2).

472. Pearce SF, Rebelo-Guiomar P, D'Souza AR, Powell CA, Van Haute L, Minczuk M. Regulation of Mammalian Mitochondrial Gene Expression: Recent Advances. Vol. 42, Trends in Biochemical Sciences. 2017.

473. Osorio D, Cai JJ. Systematic determination of the mitochondrial proportion in human and mice tissues for single-cell RNA sequencing data quality control. bioRxiv [Internet]. 2020 Aug 19 [cited 2022 Apr 4];2020.02.20.958793. Available from: https://www.biorxiv.org/content/10.1101/2020.02.20.958793v2

474. Raj A, van Oudenaarden A. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. Vol. 135, Cell. 2008.

475. Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. Vol. 14, Nature Reviews Genetics. 2013.

476. Zhang X, Lan Y, Xu J, Quan F, Zhao E, Deng C, et al. CellMarker: A manually curated resource of cell markers in human and mouse. Nucleic Acids Res. 2019;47(D1).

477. Silva FAR, Rodrigues BL, Ayrizono MDLS, Leal RF. The Immunological Basis of Inflammatory Bowel Disease. Vol. 2016, Gastroenterology Research and Practice. 2016.

478. Clavarino G, Delouche N, Vettier C, Laurin D, Pernollet M, Raskovalova T, et al. Novel strategy for phenotypic characterization of Human B lymphocytes from precursors to effector cells by flow cytometry. PLoS One. 2016;11(9).

479. Hasegawa A, Nakayama T. Role of CD69 in the pathogenesis of inflammation. Vol. 33, Japanese Journal of Clinical Immunology. 2010.

480. Matsui T, Connolly JE, Michnevitz M, Chaussabel D, Yu CI, Glaser C, et al. CD2 Distinguishes Two Subsets of Human Plasmacytoid Dendritic Cells with Distinct Phenotype and Functions. The Journal of Immunology. 2009;182(11).

481. Acosta-Rodriguez E V., Napolitani G, Lanzavecchia A, Sallusto F. Interleukins 1β and 6 but not transforming growth factor-β are essential for the differentiation of interleukin 17-producing human T helper cells. Nat Immunol. 2007;8(9).

482. Heuberger C, Pott J, Maloy KJ. Why do intestinal epithelial cells express MHC class II? Vol. 162, Immunology. 2021.

483. Wosen JE, Mukhopadhyay D, MacAubas C, Mellins ED. Epithelial MHC class II expression and its role in antigen presentation in the gastrointestinal and respiratory tracts. Vol. 9, Frontiers in Immunology. 2018.

484. Lawrance IC, Maxwell L, Doe W. Altered response of intestinal mucosal fibroblasts to profibrogenic cytokines in inflammatory bowel disease. Inflamm Bowel Dis. 2001;7(3).

485. Roulis M, Flavell RA. Fibroblasts and myofibroblasts of the intestinal lamina propria in physiology and disease. Vol. 92, Differentiation. 2016.

486. Barnhoorn MC, Hakuno SK, Bruckner RS, Rogler G, Hawinkels LJAC, Scharl M. Stromal cells in the pathogenesis of inflammatory bowel disease. Vol. 14, Journal of Crohn's and Colitis. 2020.

487. Angelbello AJ, Chen JL, Childs-Disney JL, Zhang P, Wang ZF, Disney MD. Using Genome Sequence to Enable the Design of Medicines and Chemical Probes. Vol. 118, Chemical Reviews. 2018.

488. Pellino G, Pallante P, Selvaggi F. Novel biomarkers of fibrosis in Crohn's disease. World J Gastrointest Pathophysiol. 2016;7(3).

489. Yang Y, Zeng QS, Zou M, Zeng J, Nie J, Chen DF, et al. Targeting Gremlin 1 Prevents Intestinal Fibrosis Progression by Inhibiting the Fatty Acid Oxidation of Fibroblast Cells. Front Pharmacol. 2021;12.

490. Hazel K, O'Connor A. Emerging treatments for inflammatory bowel disease. Vol. 11, Therapeutic Advances in Chronic Disease. 2020.

491. Revés J, Ungaro RC, Torres J. Unmet needs in inflammatory bowel disease. Vol. 2, Current Research in Pharmacology and Drug Discovery. 2021.

492. Wynn TA. Fibrotic disease and the TH1/TH2 paradigm. Nat Rev Immunol. 2004;4(8):583–94.

493. Wu X, Lin X, Tan J, Liu Z, He J, Hu F, et al. Cellular and Molecular Mechanisms of Intestinal Fibrosis. Gut Liver [Internet]. 2023 Mar 10; Available from: http://gutnliver.org/journal/view.html?doi=10.5009/gnl220045

494. Bochaton-Piallat ML, Gabbiani G, Hinz B. The myofibroblast in wound healing and fibrosis: Answered and unanswered questions. Vol. 5, F1000Research. 2016.

495. Pakshir P, Noskovicova N, Lodyga M, Son DO, Schuster R, Goodwin A, et al. The myofibroblast at a glance. J Cell Sci. 2020;133(13).

496. Barrett RL, Pure E. Cancer-associated fibroblasts and their influence on tumor immunity and immunotherapy. Elife. 2020;9.

497. Marumo T, Hishikawa K, Yoshikawa M, Hirahashi J, Kawachi S, Fujita T. Histone deacetylase modulates the proinflammatory and -fibrotic changes in tubulointerstitial injury. Am J Physiol Renal Physiol. 2010;298(1).

498. Laddha AP, Kulkarni YA. VEGF and FGF-2: Promising targets for the treatment of respiratory disorders. Vol. 156, Respiratory Medicine. 2019.

499. Bordignon P, Bottoni G, Xu X, Popescu AS, Truan Z, Guenova E, et al. Dualism of FGF and TGF-β Signaling in Heterogeneous Cancer-Associated Fibroblast Activation with ETV1 as a Critical Determinant. Cell Rep. 2019;28(9).

500. Song QH, Klepeis VE, Nugent MA, Trinkaus-Randall V. TGF-β1 regulates TGF-β1 and FGF-2 mRNA expression during fibroblast wound healing. Journal of Clinical Pathology - Molecular Pathology. 2002;55(3).

501. Danopoulos S, Schlieve CR, Grikscheit TC, Al Alam D. Fibroblast Growth Factors in the Gastrointestinal Tract: Twists and Turns. Developmental Dynamics. 2017;246(4).

502. Deng M, Yang S, Ji Y, Lu Y, Qiu M, Sheng Y, et al. Overexpression of peptidase inhibitor 16 attenuates angiotensin II–induced cardiac fibrosis via regulating HDAC1 of cardiac fibroblasts. J Cell Mol Med. 2020;24(9).

503. Sehgal M, Jakhete SM, Manekar AG, Sasikumar S. Specific epigenetic regulators serve as potential therapeutic targets in idiopathic pulmonary fibrosis. Heliyon. 2022 Aug 1;8(8):e09773.

504. Pang M, Kothapally J, Mao H, Tolbert E, Ponnusamy M, Chin YE, et al. Inhibition of histone deacetylase activity attenuates renal fibroblast activation and interstitial fibrosis in obstructive nephropathy. Am J Physiol Renal Physiol. 2009;297(4).

505. Wang Z, Chen C, Finger SN, Do Kwajah MM S, Jung M, Schwarz H, et al. Suberoylanilide hydroxamic acid: A potential epigenetic therapeutic agent for lung fibrosis? European Respiratory Journal. 2009;34(1).

506. Turgeon N, Gagné JM, Blais M, Gendron FP, Boudreau F, Asselin C. The acetylome regulators Hdac1 and Hdac2 differently modulate intestinal epithelial cell dependent homeostatic responses in experimental colitis. Am J Physiol Gastrointest Liver Physiol. 2014;306(7).

507. Zhang YQ, Wan LY, He XM, Ni YR, Wang C, Liu CB, et al. Gremlin1 Accelerates Hepatic Stellate Cell Activation Through Upregulation of TGF-Beta Expression. DNA Cell Biol. 2017;36(7).

508. Duffy L, Henderson J, Brown M, Pryzborski S, Fullard N, Summa L, et al. Bone Morphogenetic Protein Antagonist Gremlin-1 Increases Myofibroblast Transition in Dermal Fibroblasts: Implications for Systemic Sclerosis. Front Cell Dev Biol. 2021;9.

509. Dituri F, Cossu C, Mancarella S. The Interactivity between TGF β and BMP Signaling in Organogenesis, Fibrosis, and Cancer. 2019;(I):1–21.

510. Rayego-Mateos S, Campillo S, Rodrigues-Diez RR, Tejera-Muñoz A, Marquez-Exposito L, Goldschmeding R, et al. Interplay between extracellular matrix components and cellular and molecular mechanisms in kidney fibrosis. Vol. 135, Clinical Science. 2021.

511. Abraham C, Cho JH. Inflammatory Bowel Disease. New England Journal of Medicine [Internet]. 2009 Nov 19 [cited 2023 Apr 7];361(21):2066–78. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMra0804647

512. Langner C, Magro F, Driessen A, Ensari A, Mantzaris GJ, Villanacci V, et al. The histopathological approach to inflammatory bowel disease: A practice guide. Vol. 464, Virchows Archiv. 2014.

513. Song X, Dai D, He X, Zhu S, Yao Y, Gao H, et al. Growth Factor FGF2 Cooperates with Interleukin-17 to Repair Intestinal Epithelial Damage. Immunity. 2015;43(3).

514. Cushing MC, Mariner PD, Liao J, Sims EA, Anseth KS. Fibroblast growth factor represses Smad-mediated myofibroblast activation in aortic valvular interstitial cells. The FASEB Journal. 2008;22(6).

515. Tiede S, Ernst N, Bayat A, Paus R, Tronnier V, Zechel C. Basic fibroblast growth factor: A potential new therapeutic tool for the treatment of hypertrophic and keloid scars. Annals of Anatomy. 2009;191(1).

516. Zhao M, Wang L, Wang M, Zhou S, Lu Y, Cui H, et al. Targeting fibrosis, mechanisms and cilinical trials. Vol. 7, Signal Transduction and Targeted Therapy. Springer Nature; 2022.

517. Gatla HR, Muniraj N, Thevkar P, Yavvari S, Sukhavasi S, Makena MR. Regulation of chemokines and cytokines by histone deacetylases and an update on histone decetylase inhibitors in human diseases. Vol. 20, International Journal of Molecular Sciences. 2019.

518. Gonneaud A, Turgeon N, Jones C, Couture C, Lévesque D, Boisvert FM, et al. HDAC1 and HDAC2 independently regulate common and specific intrinsic responses in murine enteroids. Sci Rep. 2019;9(1).

519. Zheng Y, Khan Z, Zanfagnin V, Correa LF, Delaney AA, Daftary GS. Epigenetic modulation of collagen 1A1: Therapeutic implications in fibrosis and endometriosis. Biol Reprod. 2016;94(4).

520. Ren J, Smid M, Iaria J, Salvatori DCF, Van Dam H, Zhu HJ, et al. Cancer-associated fibroblast-derived Gremlin 1 promotes breast cancer progression. Breast Cancer Research. 2019;21(1).

521. Church RH, Krishnakumar A, Urbanek A, Geschwindner S, Meneely J, Bianchi A, et al. Gremlin1 preferentially binds to bone morphogenetic protein-2 (BMP-2) and BMP-4 over BMP-7. Biochemical Journal. 2015;466.

522. Walton KL, Johnson KE, Harrison CA. Targeting TGF-β mediated SMAD signaling for the prevention of fibrosis. Front Pharmacol. 2017 Jul 14;8(JUL).

523. Zeisberg M, Kalluri R. Cellular mechanisms of tissue fibrosis. 1. Common and organ-specific mechanisms associated with tissue fibrosis. Am J Physiol Cell Physiol. 2013;304(3).

524. McAnulty RJ. Fibroblasts and myofibroblasts: Their source, function and role in disease. Vol. 39, International Journal of Biochemistry and Cell Biology. 2007.

525. Church RH, Ali I, Tate M, Lavin D, Krishnakumar A, Kok HM, et al. Gremlin1 plays a key role in kidney development and renal fibrosis. Am J Physiol Renal Physiol. 2017;312(6).

526. Chen G, Xu H, Yao Y, Xu T, Yuan M, Zhang X, et al. BMP Signaling in the Development and Regeneration of Cranium Bones and Maintenance of Calvarial Stem Cells. Vol. 8, Frontiers in Cell and Developmental Biology. 2020.

527. da Silva Madaleno C, Jatzlau J, Knaus P. BMP signalling in a mechanical context – Implications for bone biology. Vol. 137, Bone. 2020.

528. Lavery K, Swain P, Falb D, Alaoui-Ismaili MH. BMP-2/4 and BMP-6/7 differentially utilize cell surface receptors to induce osteoblastic differentiation of human bone marrow-derived mesenchymal stem cells. Journal of Biological Chemistry. 2008;283(30).

529. Kosinski C, Li VSW, Chan ASY, Zhang J, Ho C, Wai YT, et al. Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. Proc Natl Acad Sci U S A. 2007;104(39):15418–23.

530. Miyazono K, Kamiya Y, Morikawa M. Bone morphogenetic protein receptors and signal transduction. Vol. 147, Journal of Biochemistry. 2010.

531. Wynn TA. Integrating mechanisms of pulmonary fibrosis. Journal of Experimental Medicine. 2011;208(7).

532. Chung YH, Huang YH, Chu TH, Chen CL, Lin PR, Huang SC, et al. BMP-2 restoration aids in recovery from liver fibrosis by attenuating TGF-β1 signaling. Laboratory Investigation. 2018;98(8).

533. Guo J, Lin Q, Shao Y, Rong L, Zhang D. BMP-7 suppresses excessive scar formation by activating the BMP-7/Smad1/5/8 signaling pathway. Mol Med Rep. 2017;16(2).

534. Budi EH, Schaub JR, Decaris M, Turner S, Derynck R. TGF-β as a driver of fibrosis: physiological roles and therapeutic opportunities. 2021 [cited 2023 Mar 28]; Available from: www.thejournalofpathology.com

535. Yang YL, Liu YS, Chuang LY, Guh JY, Lee TC, Liao TN, et al. Bone morphogenetic protein-2 antagonizes renal interstitial fibrosis by promoting catabolism of type i transforming growth factor-β receptors. Endocrinology. 2009;150(2).

536. Yang YL, Ju HZ, Liu SF, Lee TC, Shih YW, Chuang LY, et al. BMP-2 suppresses renal interstitial fibrosis by regulating epithelial-mesenchymal transition. J Cell Biochem. 2011;112(9).

537. De Langhe E, Cailotto F, De Vooght V, Aznar-Lopez C, Vanoirbeek JA, Luyten FP, et al. Enhanced endogenous bone morphogenetic protein signaling protects against bleomycin induced pulmonary fibrosis. Respir Res. 2015;16(1).

538. Przekora A, Zarnowski T, Ginalska G. A simple and effective protocol for fast isolation of human Tenon's fibroblasts from a single trabeculectomy biopsy - A comparison of cell behaviour in different culture media. Cell Mol Biol Lett. 2017;22(1).

539. Sawant S, Dongre H, Singh AK, Joshi S, Costea DE, Mahadik S, et al. Establishment of 3D co-culture models from different stages of human tongue tumorigenesis: Utility in understanding neoplastic progression. PLoS One. 2016 Aug 1;11(8).

540. Dolmans MM, Michaux N, Camboni A, Martinez-Madrid B, Van Langendonckt A, Nottola SA, et al. Evaluation of Liberase, a purified enzyme blend, for the isolation of human primordial and primary ovarian follicles. Human Reproduction. 2006;21(2).

541. Nikolić D, Djordjević PB, Srecković VD, Paunović I, Kalezić N, Popović S, et al. The effect of different concentrations of Liberase HI used in a non automated method for human adult pancreatic islet isolation. Arch Biol Sci. 2010;62(3).

542. Saraswati S, Marrow SMW, Watch LA, Young PP. Identification of a pro-angiogenic functional role for FSP1-positive fibroblast subtype in wound healing. Nat Commun. 2019 Dec 1;10(1).

543. LeBleu VS, Neilson EG. Origin and functional heterogeneity of fibroblasts. Vol. 34, FASEB Journal. 2020.

544. Ohyama H, Nishimura F, Meguro M, Takashiba S, Murayama Y, Matsushita S. Counter-antigen presentation: Fibroblasts produce cytokines by signalling through HLA class II molecules without inducing T-cell proliferation. Cytokine. 2002;17(4).

545. Chun Geun Lee, Homer RJ, Zhu Z, Lanone S, Wang X, Koteliansky V, et al. Interleukin-13 induces tissue fibrosis by selectively stimulating and activating transforming growth factor β1. Journal of Experimental Medicine. 2001;194(6).

546. Elias M, Zhao S, Le HT, Wang J, Neurath MF, Neufert C, et al. IL-36 in chronic inflammation and fibrosis - bridging the gap? Vol. 131, Journal of Clinical Investigation. 2021.

547. Calabrese EJ. Hormesis and adult adipose-derived stem cells. Vol. 172, Pharmacological Research. 2021.

548. Calabrese EJ, Dhawan G, Kapoor R, Agathokleous E, Calabrese V. Hormesis: wound healing and fibroblasts. Pharmacol Res. 2022 Oct 1;184.

549. Marconi GD, Fonticoli L, Rajan TS, Lanuti P, Della Rocca Y, Pierdomenico SD, et al. Transforming Growth Factor-Beta1 and Human Gingival Fibroblast-to-Myofibroblast Differentiation: Molecular and Morphological Modifications. Front Physiol. 2021;12.

550. Fessing MY, Atoyan R, Shander B, Mardaryev AN, Botchkarev V V., Poterlowicz K, et al. BMP signaling induces cell-type-specific changes in gene expression programs of human keratinocytes and fibroblasts. Journal of Investigative Dermatology. 2010;130(2).

551. Ghosh AK, Vaughan DE. PAI-1 in tissue fibrosis. Vol. 227, Journal of Cellular Physiology. 2012.

552. Dolivo DM, Larson SA, Dominko T. Fibroblast Growth Factor 2 as an Antifibrotic: Antagonism of Myofibroblast Differentiation and Suppression of Pro-Fibrotic Gene Expression. Cytokine Growth Factor Rev. 2017;38(September):49–58.

553. Koli K, Myllärniemi M, Vuorinen K, Salmenkivi K, Ryynänen MJ, Kinnula VL, et al. Bone morphogenetic protein-4 inhibitor Gremlin is overexpressed in idiopathic pulmonary fibrosis. American Journal of Pathology. 2006;169(1).

554. Westerman KA, Ao Z, Cohen ÉA, Leboulch P. Design of a trans protease lentiviral packaging system that produces high titer virus. Retrovirology. 2007;4.

555. Kim SH, Jang SI, Park CY, You JC. Investigation of requirements for efficient gene delivery using the HIV-1 based lentiviral transduction system. Biochem Biophys Res Commun. 2009;383(2).

556. Pfister G, Toor SM, Sasidharan Nair V, Elkord E. An evaluation of sorter induced cell stress (SICS) on peripheral blood mononuclear cells (PBMCs) after different sort conditions - Are your sorted cells getting SICS? J Immunol Methods. 2020;487.

557. Han H. RNA interference to knock down gene expression. In: Methods in Molecular Biology. 2018.

558. Seluanov A, Vaidya A, Gorbunova V. Establishing primary adult fibroblast cultures from rodents. Journal of Visualized Experiments. 2010;(44).

559. Yang YHK, Ogando CR, Wang See C, Chang TY, Barabino GA. Changes in phenotype and differentiation potential of human mesenchymal stem cells aging in vitro. Stem Cell Res Ther. 2018;9(1).

560. Heinzelmann K, Lehmann M, Gerckens M, Noskovičová N, Frankenberger M, Lindner M, et al. Cell-surface phenotyping identifies CD36 and CD97 as novel markers of fibroblast quiescence in lung fibrosis. Am J Physiol Lung Cell Mol Physiol. 2018;315(5).

561. Eyden B. The myofibroblast: Phenotypic characterization as a prerequisite to understanding its functions in translational medicine: Translational Medicine. J Cell Mol Med. 2008;12(1).

562. Díaz-Flores L, Gutiérrez R, García MP, Sáez FJ, Díaz-Flores L, Valladares F, et al. CD34+ stromal cells/fibroblasts/fibrocytes/telocytes as a tissue reserve and a principal source of mesenchymal cells. Location, morphology, function and role in pathology. Vol. 29, Histology and Histopathology. 2014.

563. Dobaczewski M, Bujak M, Li N, Gonzalez-Quesada C, Mendoza LH, Wang XF, et al. Smad3 signaling critically regulates fibroblast phenotype and function in healing myocardial infarction. Circ Res. 2010;107(3).

564. Mia MM, Boersema M, Bank RA. Interleukin-1β attenuates myofibroblast formation and extracellular matrix production in dermal and lung fibroblasts exposed to transforming growth factor-β1. PLoS One. 2014;9(3).

565. Huang H, Huang H, Li Y, Liu M, Shi Y, Chi Y, et al. Gremlin induces cell proliferation and extra cellular matrix accumulation in mouse mesangial cells exposed to high glucose via the ERK1/2 pathway. BMC Nephrol. 2013;14(1).

566. Sneddon JB, Zhen HH, Montgomery K, Van De Rijn M, Tward AD, West R, et al. Bone morphogenetic protein antagonist gremlin 1 is wideley expressed by cancer-associated stromal cells and can promote tumor cell proliferation. Proc Natl Acad Sci U S A. 2006;103(40).

567. Lin X, Wang Y, Liu Z, Lin S, Tan J, He J, et al. Intestinal strictures in Crohn's disease: a 2021 update. Vol. 15, Therapeutic Advances in Gastroenterology. 2022.

568. Rieder F, de Bruyn JR, Pham BT, Katsanos K, Annese V, Higgins PDR, et al. Results of the 4th Scientific Workshop of the ECCO (Group II): Markers of intestinal fibrosis in inflammatory bowel disease. Vol. 8, Journal of Crohn's and Colitis. 2014.

569. Bensalah M, Muraine L, Boulinguiez A, Giordani L, Albert V, Ythier V, et al. A negative feedback loop between fibroadipogenic progenitors and muscle fibres involving endothelin promotes human muscle fibrosis. J Cachexia Sarcopenia Muscle. 2022;13(3).

570. Rieder F, Fiocchi C. Intestinal fibrosis in IBD - A dynamic, multifactorial process. Vol. 6, Nature Reviews Gastroenterology and Hepatology. 2009.

571. Li C, Kuemmerle JF. The fate of myofibroblasts during the development of fibrosis in Crohn's disease. Vol. 21, Journal of Digestive Diseases. 2020.

572. Chang HY, Chi JT, Dudoit S, Bondre C, Van De Rijn M, Botstein D, et al. Diversity, topographic differentiation, and positional memory in human fibroblasts. Proc Natl Acad Sci U S A. 2002;99(20).

573. Yang T, Chen SL, Lu XJ, Shen CY, Liu Y, Chen YP. Bone morphogenetic protein 7 suppresses the progression of hepatic fibrosis and regulates the expression of gremlin and transforming growth factor β1. Mol Med Rep. 2012;6(1).

574. Wellbrock J, Sheikhzadeh S, Oliveira-Ferrer L, Stamm H, Hillebrand M, Keyser B, et al. Overexpression of gremlin-1 in patients with Loeys-Dietz syndrome: Implications on pathophysiology and early disease detection. PLoS One. 2014;9(8).

575. Parsonage G, Falciani F, Burman A, Filer A, Ross E, Bofill M, et al. Global gene expression profiles in fibroblasts from synovial, skin and lymphoid tissue levels distinct cytokine and chemokine expression patterns. Thromb Haemost. 2003;90(4).

576. Køster D, Egedal JH, Lomholt S, Hvid M, Jakobsen MR, Müller-Ladner U, et al. Phenotypic and functional characterization of synovial fluid-derived fibroblast-like synoviocytes in rheumatoid arthritis. Sci Rep. 2021;11(1).