Characterising bidirectional interactions between synovial fibroblasts and myeloid cells

By

Jason Dale Turner

A thesis submitted to the University of Birmingham for the degree of Doctor of Philosophy in Immunity and Infection

Rheumatology Research group
Institute of Inflammation and Ageing
College of Medical and Dental Sciences
Submitted September 2017

Abstract

Synovial fibroblasts and macrophages are incriminated in rheumatoid arthritis (RA) but the interactions between these cells and the roles of cellular subsets are only partially understood. To address this, fibroblasts isolated from normal, resolving arthritis, very early RA, established RA, and longer duration RA patients were co-cultured *in vitro* with myeloid cells; synovial fibroblast and macrophage subsets were identified, and the transcriptomes of synovial cells were analysed.

Co-culture of fibroblasts and monocytes elicited an increase in IL-6 release and a reduction in CCL2/CCL4 levels. No difference in response was elicited by fibroblasts from different stages of RA.

Fibroblast and macrophage markers were defined in frozen tissue sections. A synovial tissue digestion protocol was then developed and used to isolate synovial cells. Fibroblast and macrophage populations were identified and the proportions of fibroblast subsets correlated with clinical variables. Transcriptional analysis identified differentially expressed genes between fibroblast subsets and one distinct macrophage population.

Analysis of previously generated transcriptome data for fibroblasts from RA stages identified cassettes of genes differentially expressed at each disease stage.

This work expands previous findings on fibroblast function by beginning to assign functions to subsets and demonstrating that fibroblasts from stages of RA have distinct gene expression.

Dedicated to Natasha and Felicity, my everything

Also to those who have been told that they cannot
You can

Acknowledgments

As with all projects of this kind acknowledgments must be made to many people for the various contributions made to the work and towards my development. First of all I must thank the team of supervisors who have guided me through the maze that is studying for a doctorate. Christopher Buckley, Andy Clark, and especially Andrew Filer and Christine Huppertz have provided support, guidance, and comments throughout the degree programme. Without their aid little of this work would have been completed.

I must thank all previous and current members of the Rheumatology Research Group here in Birmingham. Whether you are aware of it or not you have all contributed towards my development and for that I am forever grateful. I give particular thanks to Debbie Hardie; you took me under your wing when I first joined and gave me that critical, formative guidance needed. Karl Welzenbach and Laurent Klein, based at Novartis, also require my thanks for the discussions regarding the pilot work they completed which contributed towards the instigation of this project. Other employees of Novartis, Sabina Pfister, Marc Sultan, and Virginie Petitjean also deserve my gratitude for completing RNA-sequencing work and instructing me in the bioinformatic analysis of the data.

The people who I owe the most to and therefore deserve the most thanks are my family, Natasha and Felicity. Without your unending, unconditional support I would not have made it this far and you were always there to pick me back up when I faltered. For that I cannot thank you enough, you are both saints.

Thank you to the Medical Research Council and Novartis for funding the project.

CONTENTS

List of figures

List of tables

Abbreviations

ACPA = Anti-citrullinated protein antibody
AMP = Accelerating medicines partnership
Anti-CCP = Anti-cyclic citrullinated peptide
APC = antigen-presenting cell
APRIL = A proliferation-inducing ligand
BAFF = B-cell activating factor
BC = Blood cone
bDMARD = Biological disease modifying anti-rheumatic drug
BEACON = Birmingham Early Arthritis Cohort
bp = Base-pair
C3AR1 = Complement component 3a receptor 1
C5 = Complement component 5
CCL14 = C-C motif ligand 14
CCL19 = C-C motif ligand 19
CCL2 = C-C motif ligand 2
CCL3 = C-C motif ligand 3
CCL4 = C-C motif ligand 4
CCL5 = C-C motif ligand 5
CCL8 = C-C motif ligand 8
CCR2 = C-C chemokine receptor type 2
CCR4 = C-C chemokine receptor type 4
CCR7 = C-C chemokine receptor type 7
CDB = Cell dissociation buffer
cDNA = Complementary deoxyribonucleic acid
CI = Confidence interval
CPM = Counts per million
CRP = C-reactive protein
csDMARD = Conventional synthetic disease modifying anti-rheumatic drug
CSF3R = Colony stimulating factor 3 receptor
CSFR1 = Colony stimulating factor receptor 1
CTLA-4 = Cytotoxic T-lymphocyte-associated protein 4
CX3CR1 = CX3C chemokine receptor 1
CXCL10 = C-X-C motif ligand 10
CXCL11 = C-X-C motif ligand 11
CXCL12 = C-X-C motif ligand 12
CXCL13 = C-X-C motif ligand 13
CXCL16 = C-X-C motif ligand 16
CXCL4 = C-X-C motif ligand 4
CXCL7 = C-X-C motif ligand 7
CXCL8 = C-X-C motif ligand 8
CXCR1 = C-X-C chemokine receptor type 1
CXCR2 = C-X-C chemokine receptor type 2
CXCR3 = C-X-C chemokine receptor type 3
CXCR4 = C-X-C chemokine receptor type 4
CyTOF = Cytometry by time of flight
DAMP = Damage-associated molecular patterns
DC = Dendritic cell
DC = Direct co-culture
DE = Differentially expressed
DMARD - Disease modifying anti-rheumatic drug
DNMT = Deoxyribonucleic acid methytransferase
DMSO = Dimethyl sulphoxide
DN = Double negative

DNA = Deoxyribonucleic acid
DNaseI = Deoxyribonuclease I
EDTA = Ethylenediaminetetraacetic acid
ELISA = Enzyme-linked immunosorbent assay
ESR = Erythrocyte sedimentation rate
EstRA = Established rheumatoid arthritis
EULAR = European league against rheumatism
FAP = Fibroblast activation protein alpha
FBS = Foetal bovine serum
$FcRL4^{+}$ = Fc receptor like protein 4
FGF2 = Basic fibroblast growth factor
FLS = Fibroblast like synoviocyte
FRβ = Folate receptor beta
FSC = Forward-scatter
G-CSF = Granulocyte-colony stimulating factor
GM-CSF = Granulocyte-macrophage colony-stimulating factor
GO = Gene ontology
GSEA = Gene set enrichment analysis
HAT = Histone acetylase
HBSS = Hank's balanced salt solution
HD = Healthy donor
HDAC = Histone deacetylase
HLA = Human leukocyte antigen
HMGB1 = High mobility group box 1
HO-1 = Heme oxygenase 1
HOX = Homeobox
HSC = Hematopoietic stem cell
ICAM-1 = Intercellular adhesion molecule 1
IDO = Indolamine-pyrrole 2,3-dioxygenase
IFNβ = Interferon beta
IFNγ = Interferon gamma
IgA = Immunoglobulin A
IgG = Immunoglobulin G
IL-10 = Interleukin-10
IL-11 = Interleukin-11
IL-12p35 = Interleukin-12 p35 subunit
IL-12p70 = Interleukin-12 p70 subunit
IL-13 = Interleukin-13
IL-15 = Interleukin-15
IL-17 = Interleukin-17
IL-1ra = Interleukin-1 receptor agonist
IL-1α = Interleukin-1 alpha
IL-1β = Interleukin-1 beta
IL-2 = Interleukin-2
IL-23 = Interleukin-23
IL-23p19 = Interleukin-23 p19 subunit
IL-3 = Interleukin-3
IL-4 = Interleukin-4
IL-5 = Interleukin-5
IL-6 = Interleukin-6
IL-8 = Interleukin-8
IL-9 = Interleukin-9
Jak = Janus kinase
JRep = Joint replacement rheumatoid arthritis

KEGG = Kyoto encyclopaedia of genes and genomes
LIF = Leukaemia inhibitory factor
LOESS = Locally weighted scatterplot smoothing
Log2CPM = Log 2 counts per million
Log2FC = Log2 fold change
LPS = Lipopolysaccharide
LXRα = Liver X receptor alpha
MCP = Metacarpophalangeal
M-CSF = Macrophage colony-stimulating factor
MDS = Multidimensional scaling
MerTK = MER Proto-Oncogene, Tyrosine Kinase
MFI = Median fluorescent intensity
MHC-II = Major histocompatibility complex 2
miRNA = Micro ribonucleic acid
MMP1 = Matrix metalloproteinase 1
MMP3 = Matrix metalloproteinase 3
MMP9 = Matrix metalloproteinase 9
MRC1 = Mannose receptor C-type 1
mRNA = Messenger ribonucleic acid
NET = Neutrophil extracellular trap
NGS = Next generation sequencing
NHSBT = National Health Service blood and transplant service
NIH = National Institutes of Health
NK cell = Natural killer cell
Norm = Normal
OA = Osteoarthritis
OA-FLS = Osteoarthritis fibroblast-like synoviocytes
OSM = Oncostatin M
PADI4 = Peptidyl arginine deiminase type IV
PAI-1 = Plasminogen activator inhibitor 1
PAMP = Pathogen-associated molecular patterns
PBMC = Peripheral blood mononuclear cells
PBS = Phosphate buffered saline
PCA = Principal component analysis
PCR = Polymerase chain reaction
PD1= Programmed cell death protein 1
PPARG = Peroxisome proliferator-activated receptor gamma
PRR = Pattern recognition receptors
qRT-PCR = Quantitative real-time polymerase chain reaction
RA = Rheumatoid arthritis
RA-FLS = Rheumatoid arthritis fibroblast-like synoviocytes
RANKL = Receptor activator of nuclear factor kappa-B ligand
Res = Resolving arthritis
RF = Rheumatoid factor
RIN = Ribonucleic acid integrity number
RNA = Ribonucleic acid
RPMI 1640 medium = Roswell park memorial institute 1640 medium
SCF = Stem cell factor
SCID = Severe combined immunodeficiency
SEM = Standard error of the mean
Sep = Separate co-culture
SSC = Side-scatter
STAT1 = Signal transducer and activator of transcription 1
TGFβ = Tissue growth factor beta

Th1 = T helper 1 cell
Th17 = T helper 17 cell
TIMP1 = Tissue inhibitor of metalloproteinase 1
TLR1 = Toll-like receptor 1
TLR2 = Toll-like receptor 2
TLR3 = Toll-like receptor 3
TLR4 = Toll-like receptor 4
TLR5 = Toll-like receptor 5
TLR6 = Toll-like receptor 6
TMM = Trimmed mean of M-values
TNF = Tumour necrosis factor
TNFα = Tumour necrosis factor alpha
TRAF1 = TNF receptor-associated factor 1
Treg = Regulatory T cell
tsDMARD = Targeted synthetic disease modifying anti-rheumatic drug
tSNE = t-distributed stochastic neighbour embedding
TW = Transwell co-culture
VCAM-1 = Vascular cell adhesion molecule 1
VE-cadherin = Vascular endothelial cadherin
VEGF = Vascular endothelial growth factor
VeRA = Very early rheumatoid arthritis
VLA-4 = Very late antigen-4
VOOM = Mean-variance modelling at the observational level
XIST = X-inactive specific transcript
αSMA = Alpha smooth muscle actin

1 INTRODUCTION

1.1 Rheumatoid Arthritis

1.1.1 Clinical manifestations

Rheumatoid arthritis (RA) is a chronic inflammatory disease affecting the synovial joints of the body causing damage and destruction of articular cartilage and the underlying bone and manifesting itself as morning stiffness, fatigue, and swelling and tenderness of affected joints. Due to the presence of autoantibodies directed against immune complexes and citrullinated proteins in around two thirds of patients with the disease, RA is frequently classified as an autoimmune disease (Schellekens et al., 1998). RA has a prevalence in the UK population of approximately 1% with the disease being 3 times more likely to affect women than men (Symmons et al., 2002). Patients also suffer from systemic comorbidities with an increased prevalence of cardiovascular disease being one of the most common, occurring in around 6% of people with RA (Dougados et al., 2013). Additionally, due to the immunosuppressive nature of treatments for RA, patients are at an increased risk of infection (Listing.,Gerhold and Zink, 2013).

The progression of RA is heterogeneous with some patients suffering a low-level of inflammation and joint damage while others suffer a highly inflammatory, aggressive disease rapidly leading to joint destruction (Wolfe and Sharp, 1998). This heterogeneity is reflected in the broad range of criteria used to classify inflammatory arthritis as definite RA and in the difficulty in diagnosing newly presenting arthritis (Table 1) (Aletaha et al., 2010; Machold et al., 2002).Indeed it may be more appropriate to view RA as a syndrome with several distinct pathologies as evidenced by the variability of autoantibody status, genetic risk factors, and in the distinct differences in disease manifestations visible in

histological samples of RA synovial tissue (Dennis et al., 2014; Van Der Helm-Van Mil, Annette Hm and Huizinga, 2008; Van Oosterhout et al., 2008).

	Score
Target population (who should be tested?): patients who	
1) have at least one joint with definite clinical synovitis (swelling)	
2) with the synovitis not better explained by another disease	
Classification criteria for RA (score-based algorithm: add score of categories A–D a score of ≥6/10 is needed for classification of a patient as having definite RA)	
A. Joint involvement	
1 large joint	0
2−10 large joints	1
1−3 small joints (with or without involvement of large joints)	2
4−10 small joints (with or without involvement of large joints)	3
>10 joints (at least one small joint)	5
B. Serology (at least 1 test result is needed for classification)	
Negative RF *and* negative ACPA	0
Low-positive RF *or* low-positive ACPA	2
High-positive RF *or* high-positive ACPA	3
C. Acute-phase reactants (at least one test result is needed for classification)	
Normal CRP *and* normal ESR	0
Abnormal CRP *or* abnormal ESR	1
D. Duration of symptoms	
<6 weeks	0
≥6 weeks	1

Table 1: ACR/EULAR 2010 classification criteria for Rheumatoid Arthritis adapted from (Aletaha et al., 2010). RA = rheumatoid arthritis, RF = rheumatoid factor, ACPA = anti-citrullinated protein antibody, CRP = C-reactive protein, ESR = erythrocyte sedimentation rate.

The primary sites in which the signs of RA manifest are the synovial joints of the body such as knees and elbows and also the smaller joints of the hands and feet (Smolen, Aletaha and Mcinnes, 2016). Healthy synovial joints consist of a fluid-filled space into which the two bones of the joint protrude, the termini of which are capped in a protective layer of articular cartilage (Figure 3). The joint space is enclosed within a fibrous capsule lined on the interior by a thin synovial tissue (synovium) which interfaces around the margin of the joint at the junction of the bone and articular cartilage. Under homeostatic conditions the synovium produces and maintains a small volume of synovial fluid rich in lubricants such as

hyaluronic acid and regulates extracellular matrix turnover to facilitate normal joint function (Smith, 2011).

1.2 Treatment

The treatment of RA is complex with no single treatment being suitable for all patients reflecting the heterogeneity of the disease. The current treatment approach recommended by the European League Against Rheumatism (EULAR) is to start treatment as early as possible with the disease modifying anti-rheumatic drugs (DMARDs) methotrexate or leflunomide/sulfasalazine possibly in combination with glucocorticoids (Figure 1 (Smolen et al., 2017)). Early treatment is particularly important for RA patients as several studies have found that early treatment of RA, possibly within 3 months of symptom onset, results in an improvement in disease course with patients showing less joint damage (radiological progression) and taking a longer time to the point of requiring surgical joint replacement (Kyburz et al., 2011; Moura et al., 2015; Nell et al., 2004; Van Aken et al., 2004).

If patients do not respond to treatment with DMARDs then it is recommended that patients move onto one of the biological DMARDs available such as infliximab or entanercept (tumour necrosis factor (TNF)-inhibitors) or rituximab (targeting B cells) (Smolen et al., 2017). Unfortunately, not all patients respond adequately to treatment with biologics and often treatment becomes an approach of trialing different biologics which is an inadequate method by which to treat patients and as such finding biomarkers that predict response to treatment is a current priority of research (Chen, Y. F. et al., 2006; Greenberg et al., 2008; Hyrich et al., 2006;

Smolen et al., 2017).

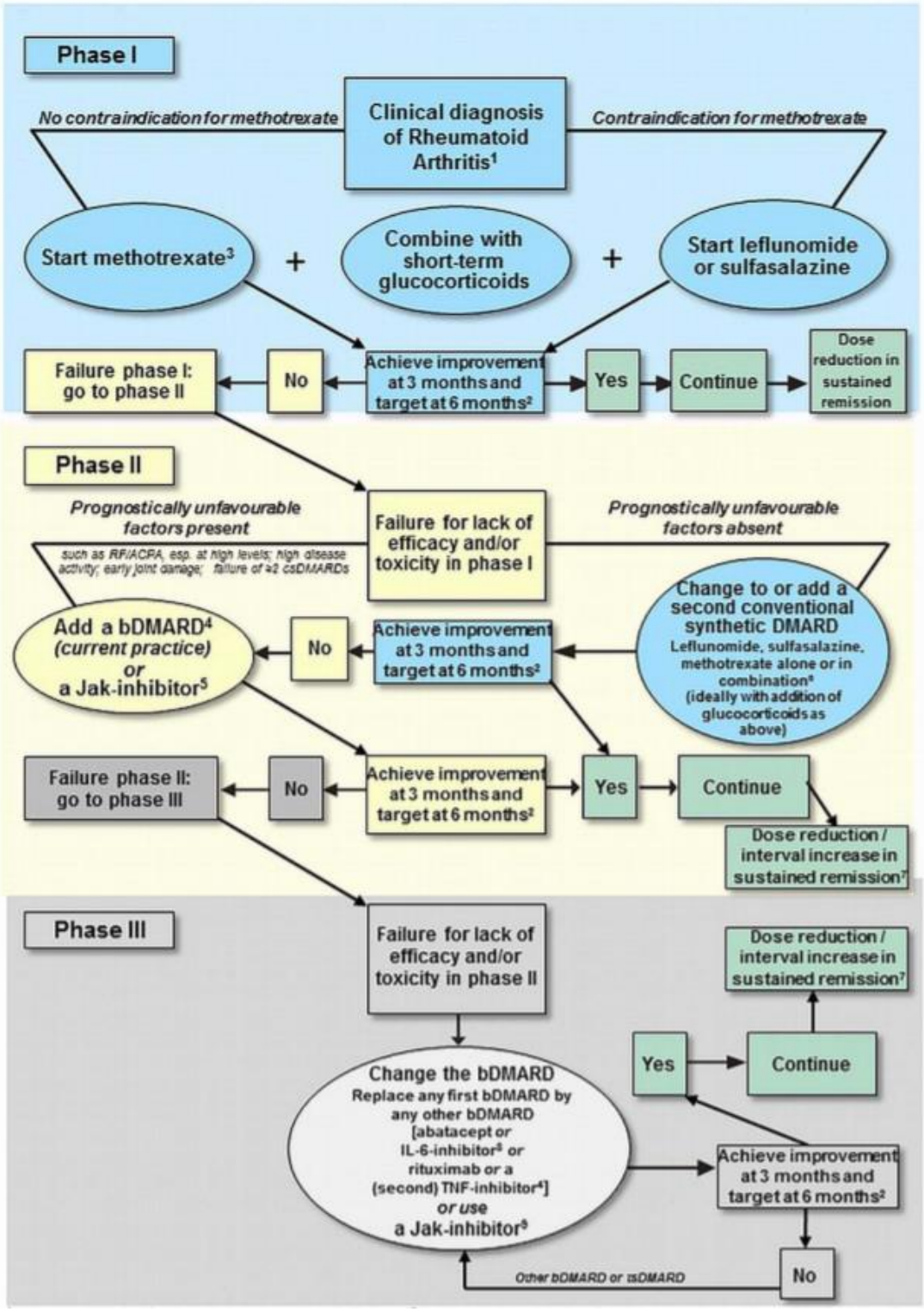


Figure 1: 'Algorithm based on the 2016 European League against Rheumatism (EULAR) recommendations on rheumatoid arthritis (RA) management' taken from Smolen et al. (2017). RF = rheumatoid factor, ACPA = anti-citrullinated protein antibody, DMARD = disease modifying anti-rheumatic drug, csDMARD = conventional synthetic DMARD, bDMARD = biological DMARD, tsDMARD = targeted synthetic DMARD, IL-6 = interleukin-6, TNF = tumour necrosis factor, Jak = Janus kinase.

New treatments targeting epigenetic aspects are being researched as an alternative to the traditional DMARDs and biological DMARDs in an attempt to not just limit inflammation but possibly reverse any epigenetic changes elicited by RA. Treating fibroblasts *in vitro* with Trichostatin A, a histone deacetylase inhibitor, inhibits progression through the cell cycle and also sensitises the fibroblasts to apoptosis induced by TNF-related apoptosis-inducing ligand (Jungel et al., 2006). In the murine collagen-induced arthritis model treatment with Trichostatin A significantly reduced arthritis as measured by a clinical score and synovial inflammation and cartilage damage measured by histology (Nasu et al., 2008). In humans the histone deacetylase inhibitor givinostat was used to treat juvenile idiopathic arthritis patients and successfully reduced swollen and tender joint counts but unfortunately had many adverse effects on patients such as vomiting, nausea, and fatigue (Vojinovic et al., 2011).

1.3 Pathophysiology

1.3.1 Aetiology

The aetiology of RA is unknown however genetic and environmental factors are known to increase predisposition for the disease. For example, tobacco smoking is associated with an increased risk of developing RA with anti-cyclic citrullinated peptide (anti-CCP) antibodies present in the circulation whereas obesity is associated with anti-CCP negative RA (Pedersen et al., 2006; Silman, Newman and Macgregor, 1996). There are also suggestions of a link between infections with organisms such as *Porphyromonas gingivalis* and the development of RA (Mikuls et al., 2014; Totaro et al., 2013).

Anti-CCP antibodies are a subset of anti-citrullinated protein antibodies (ACPA) (Ioan-Facsinay et al., 2011). During tissue damage and immune cell activation there is an increase in citrullination, the conversion of arginine residues to citrulline, of native proteins such as fibrinogen and α-Enolase. In RA patients an autoimmune response occurs targeting these neoepitopes with the production of ACPAs which can in turn drive bone loss through the differentiation of osteoclasts (Harre et al., 2012; Kinloch et al., 2005; Lundberg et al., 2008; Masson-Bessiere et al., 2001). The risk of developing ACPA-positive RA is associated with the presence of human leukocyte antigen (HLA) class II 'shared-epitope' alleles and this is believed to be due to an increased affinity of the MHC II proteins encoded by the shared epitope for citrullinated peptides which leads to the activation of the adaptive arm of the immune system (Gibofsky et al., 1978; Gregersen, Silver and Winchester, 1987; Hill et al., 2003; James et al., 2010; Stastny, 1978; Van Der Helm-Van Mil et al., 2006).

The contribution of genetic factors to RA is estimated to be around 60% with approximately 37% of this genetic contribution being attributable to HLA alleles (Deighton et al., 1989; Macgregor et al., 2000). However the impact of identified risk factors such as the presence of shared epitope alleles is not uniform across all subtypes of RA with the presence of one or more *HLA-DRB1* shared epitope alleles in a patient being significantly associated with an increased risk for anti-CCP positive RA but not anti-CCP negative RA and having no significant association with the presence or absence of Rheumatoid Factor (RF), another autoantibody implicated in RA (Bukhari et al., 2002; Huizinga et al., 2005; Van Der Helm-Van Mil et al., 2006)**.** Another genetic risk factor, the *PTPN22* R620W

single nucleotide polymorphism, is associated with an increased risk of developing all types of RA and has been shown to enhance neutrophil activation in all people regardless disease status as well as potentially altering the signalling threshold required for T lymphocyte activation (Bayley et al., 2014; Begovich et al., 2004; De La Puerta et al., 2013; Källberg et al., 2007; Michou et al., 2007; Taylor et al., 2013; Van Oene et al., 2005; Vang et al., 2012). These alterations could possibly prime the immune system increasing the propensity to develop autoimmune reactions.

There have been several studies providing evidence for the existence of further genetic risk factors associated with RA albeit often with conflicting findings between studies possibly due to the differences in methods and cohorts. Single nucleotide polymorphisms in the genomic region of the *TRAF1* and *C5* genes are significantly associated with an increased risk of developing anti-CCP positive RA in Dutch, North American, and Swedish cohorts and may also promote a more destructive disease course (Kurreeman et al., 2007; Plenge et al., 2007). Interestingly, in the murine collagen-induced arthritis model knockout of the *C5* gene or prophylactic/therapeutic blocking of the function of C5 using an anti-C5 antibody results in both a significant reduction in the incidence and the severity of arthritis to the point that there are no apparent histological differences between the synovium of healthy mice and those in which C5 was knocked out (Grant et al., 2002; Wang et al., 2000; Wang et al., 1995). Several prospective treatments are in development to target C5 or the C5a receptor in inflammatory diseases (Horiuchi et al., 2009).

Given the presence of autoantibodies targeting citrullinated peptides it is perhaps unsurprising that PADI4, an enzyme capable of citrullinating peptides, has been identified as another genetic risk factor for RA. Single nucleotide polymorphisms in the region of the *PADI4* gene are significantly associated with an increased risk of developing RA in several Asian and European cohorts (Hoppe et al., 2006; Kang et al., 2006; Suzuki et al., 2003). Mechanistically, Suzuki et al. (2003) demonstrated that a single nucleotide polymorphism in the *PADI4* gene that is particularly strongly associated with RA increases the stability of *PADI4* messenger ribonucleic acid (mRNA) offering a possible explanation for associations found.

It is clear that RA is a multifactorial disease and that the aetiology is complex. It is also clear that the syndrome of signs and symptoms of RA defines a family of closely related disorders under the umbrella title of Rheumatoid Arthritis; it is important to be aware of this heterogeneity

1.3.2 Physical manifestations

The structure of the synovium can be divided into the intimal or lining layer, which is in juxtaposition with the synovial fluid, and an underlying subintimal or sublining layer (Figure 2). The cellular composition of these two regions within the synovium differ in that the sublining layer consists of loosely associated extracellular matrix, vascular structures, resident tissue macrophages, and synovial fibroblasts, whereas the lining layer forms a cellular membrane (although lacking a classical basement membrane) that consists of type-A synoviocytes (macrophage-like synoviocytes) and type-B synoviocytes (fibroblast-like synoviocytes) (Barland, Novikoff and Hamerman, 1962; Smith, 2011).

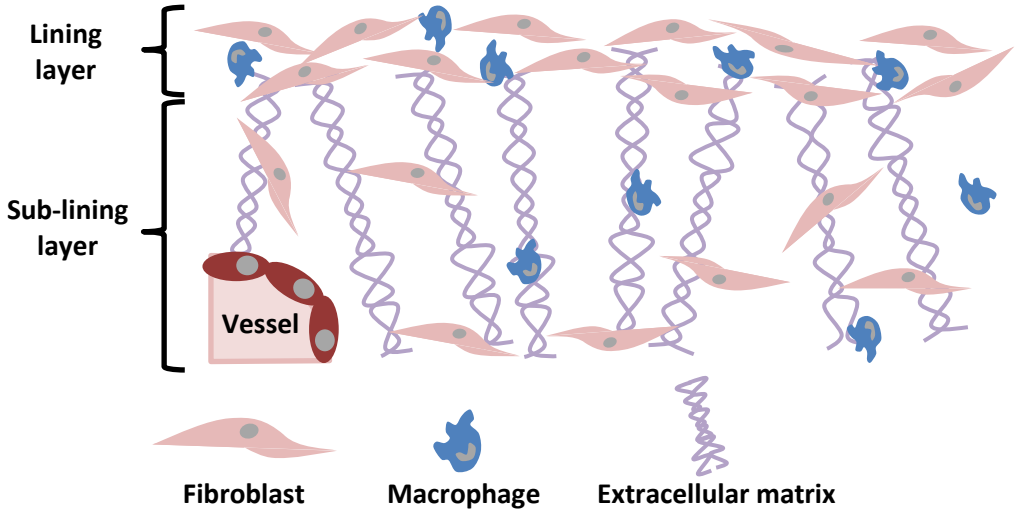


Figure 2: Schematic of the structure of normal synovium. The lining layer consists of tightly assocated fibroblasts and macrophages and is in juxtaposition with the joint space. The sublining layer is a loosely associated tissue consisting of fibroblasts and macrophages but is also penetrated by vessels.

In RA the aforementioned structure of the synovial tissue is disturbed as the tissue enters a chronically inflamed state (Figure 3).The synovial sublining becomes increasingly vascularised with an accompanying increase in inflammatory infiltrate consisting of cells of both the innate and adaptive immune system. Additionally the thickness of the lining layer, normally 2-3 cells, becomes greatly increased leading to the formation of an invasive pannus tissue that damages the articular cartilage and the underlying bone (Barland, Novikoff and Hamerman, 1962; Fassbender and Simmling-Annefeld, 1983; Smith, 2011).

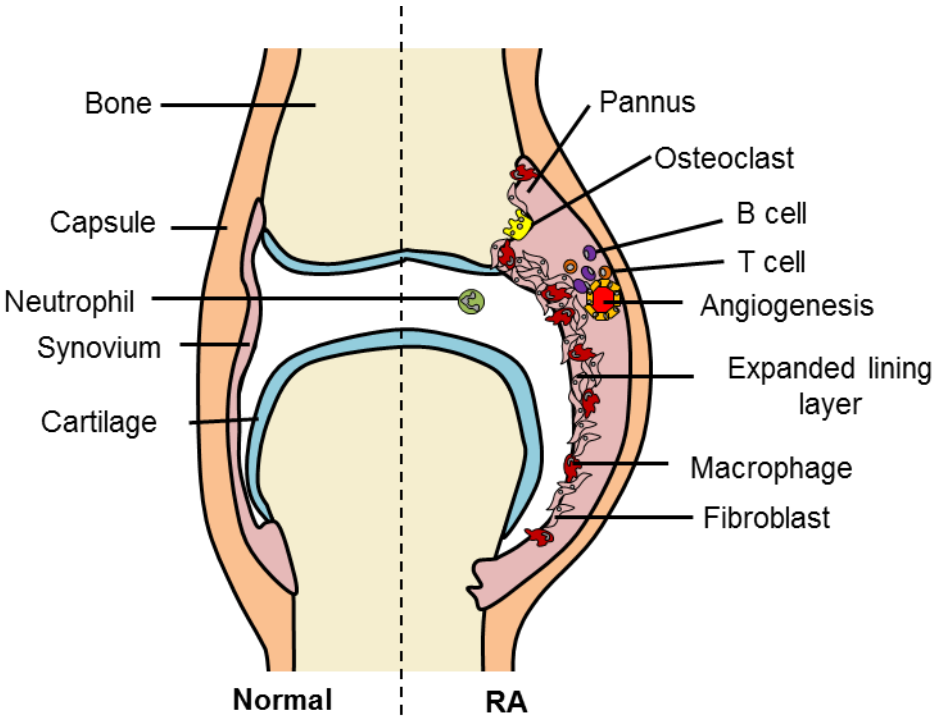


Figure 3: Schematic of a synovial joint comparing normal joint status with an inflamed joint in RA. Adapted from Strand.,Kimberly and Isaacs (2007)

1.3.3 Cellular Innate immunity

Of the members of the innate immune system neutrophils are one of the cell types highly incriminated in the pathogenesis of RA. Neutrophil extracellular traps (NETs) are known to be elicited from neutrophils in the peripheral blood and synovial fluid of RA patients at an increased efficacy than from healthy controls or osteoarthritis patients and induce interleukin-6 (IL-6) and IL-8 expression in synovial fibroblasts (Khandpur et al., 2013; Sur Chowdhury et al., 2014). Interestingly in RA neutrophils higher nuclear levels of PADI4 are found along with increased levels of histone H3 citrullination and NETosis correlates with ACPA levels found in the patients indicating that neutrophils may be tightly involved with auto-antigen generation in RA.

Of the members of the innate immune cell family antigen-presenting cells (APCs) are of particular interest as they bridge the gap between the innate and adaptive immune system. Macrophages (and as macrophage progenitors, monocytes) are important APCs and as they are central to the work in this thesis will be discussed in more detail below. Due to their prominent role as antigen presenting cells and the auto-immune aspect of RA dendritic cells (DCs) are thought to be involved in RA. Both differentiated myeloid- and plasmacytoid-DCs and DC precursors have been found to accumulate in the synovial fluid of RA patients with the concentration of DCs being increased compared to the circulating DC concentration (Jongbloed et al., 2006; Zvaifler et al., 1985). DC progenitors in the synovial fluid can be differentiated into mature DCs with a potent ability to stimulate T helper 1 cell (Th1) responses using a combination of granulocyte-macrophage colony-stimulating factor (GM-CSF), tumour necrosis factor (TNF), stem cell factor (SCF), and IL-4 (Santiago-Schwarz et al., 2001). In addition to the accumulation in synovial fluid DCs also infiltrate the synovium and form perivascular clusters (Pettit et al., 2000). A point of particular interest is that DCs provided with synovial fluid from RA patients present collagen type II and cartilage gp39 and elicit responses from T cells (Tsark et al., 2002).

1.3.4 Adaptive immunity

With regards to the adaptive immune system both T and B lymphocytes have been implicated in RA pathogenesis. T and B cells are involved in the formation and maintenance of tertiary lymphoid organs; structures resembling germinal centres that develop in the synovium of a subset of RA patients and are thought to contribute to the autoimmune response in RA (Hjelmstrom et al., 2000; Wagner

et al., 1998; Young, et al., 1984). B lymphocytes may ultimately play a role in RA through differentiation into plasma cells which produce autoantibodies and B cells are also effective APCs. However a recently discovered subset of Fc receptor like protein 4 (FcRL4) expressing B cells present in the synovial fluid and tissue of RA patients may contribute to bone erosion through expression of the osteoclast differentiating cell surface molecule receptor activator of nuclear factor kappa-B ligand (RANKL) (Massey and Flanagan, 1999; Shipman, Dasoveanu and Lu, 2017; Yeo et al., 2015b; Yeo et al., 2011).

A subset of IL-17 producing T cells termed Th17 cells are thought to contribute to the pathogenesis of RA as the cytokine is produced at higher levels in RA compared to osteoarthritis (OA) synovium and is known to induce IL-6 gene expression and also induce NETosis in neutrophils (Chabaud et al., 1999; Khandpur et al., 2013; Onishi and Gaffen, 2010; Veldhoen et al., 2006; Zrioual et al., 2008). IL-17 also increases the expression of Toll-like receptors (TLRs) 2, 3, and 4 on RA synovial fibroblasts (Lee et al., 2014). In the murine collagen-induced arthritis model genetic knockout of IL-23 prevents the development of arthritis and results in a reduction in the number of Th17 cells, probably through an indirect reduction of circulating IL-6 and tissue growth factor beta (TGFβ) levels (Bettelli et al., 2006; Murphy et al., 2003).

The anti-inflammatory regulatory T cell (Treg) population has also been implicated in RA. Treg cells normally act to restrain inflammation through the production of IL-10 and expression of the immune checkpoint protein cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) which inhibits APC activation

(Sakaguchi, 2004). Treg cells specific to heat shock protein 70 can suppress murine proteoglycan-induced arthritis demonstrating a therapeutic potential (Van Herwijnen et al., 2012). However CTLA-4 is downregulated on RA peripheral blood Treg cells and these cells respond inadequately to stimulation and are poor inhibitors of activated T cells and monocytes (Ehrenstein et al., 2004; Flores-Borja et al., 2008). Anti-TNF treatment restores the regulatory capacity of the Treg cells and increases Treg cell counts in the peripheral blood of RA patients that respond to treatment.

1.4 Monocytes, Macrophages, and Synovial Fibroblasts in Rheumatoid Arthritis

1.4.1 Monocytes and Macrophages

Monocytes and macrophages are cells of myeloid lineage of the innate arm of the immune system. Both are intricately linked to each other as certain subsets of macrophages are known to differentiate from monocytes. However, there are important differences in their phenotypes and differentiation potential that make it clear that they serve different functions in the body.

Monocytes are peripheral blood borne mononuclear cells that patrol the vasculature and to some extent tissues as sentinels aiming to rapidly detect and respond to injury or infection (Meerschaert and Furie, 1994; Priller and Bottcher, 2017; Sumagin et al., 2010). They are capable of producing a range of proinflammatory factors such as IL-1β, TNFα, and C-X-C motif ligand 8 (CXCL8) in response to stimuli inducing toll-like receptor 4 (TLR4) engagement and so are capable of initiating the immune response to invading organisms (De Waal Malefyt et al., 1991; Muller-Alouf et al., 1994). One of the most important functions

of monocytes lies in their capability to differentiate into cells such as macrophages and dendritic cells upon encountering stimuli such as GM-CSF, M-CSF, and IL-4 allowing monocytes to act as a circulating pool of progenitor cells differentiating into other cells of the myeloid lineage as and when needed (Fleetwood et al., 2009; Fleetwood et al., 2007; Sallusto and Lanzavecchia, 1994).

Macrophages are mononuclear phagocytic cells, with key roles as professional antigen presenting cells, found in the tissues of the body. They are often one of the first cell types to encounter and respond to an insult, be it a pathogen or injury. They are capable of rapidly responding to damage and infection, recognising damage-associated molecular patterns (DAMPs) and pathogen-associated molecular pathogens (PAMPs) respectively through pattern recognition receptors (PRRs), triggering immune sentinel responses such as elaboration of cytokines (TNF-α, IL-1β) and chemokines (CXCL8, a hallmark neutrophil chemoattractant in early inflammation), the production of antimicrobial substances such as reactive oxygen species, phagocytosis, and ultimately bridging innate to adaptive immune responses through antigen presentation (Apostolopoulos, Davenport and Tipping, 1996; Ribeiro et al., 1991; Takeuchi and Akira, 2010).

1.4.2 Monocyte and macrophage subsets

Much is known regarding the phenotype and function of monocytes and macrophages but newer molecular biology techniques are revealing the existence of increasing numbers of cellular subsets, often with distinct functions, within both of these populations

1.4.2.1 Monocytes

Monocytes are not a homogeneous population and were thought to consist of three subsets identified by levels of CD14 and CD16 expression, the proportions of which are altered in certain disease states, for example in active RA the intermediate $CD14^{++}CD16^{+}$ subset is expanded and is associated with an increase in the number of IL-17 producing T lymphocytes (Rossol et al., 2012; Wahl et al., 1992; Wong et al., 2011; Ziegler-Heitbrock et al., 2010; Ziegler-Heitbrock and Hofer, 2013). The majority of published studies on human monocytes have used CD14 and CD16 to identify classical ($CD14^{++}CD16^{-}$), intermediate ($CD14^{++}CD16^{+}$), and non-classical ($CD14^{+}CD16^{++}$) subsets of monocytes (Figure 4) (Ziegler-Heitbrock et al., 2010; Ziegler-Heitbrock and Hofer, 2013). In fact, recent single-cell RNA sequencing of human peripheral blood populations by Villani et al. (2017) has identified four distinct monocyte populations and indicates that the intermediate subset consists of two distinct monocyte populations along with classical and non-classical monocytes.

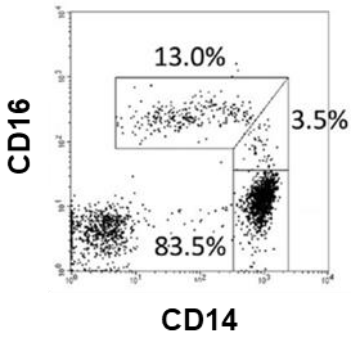


Figure 4: A representative plot demonstrating three major human peripheral blood monocyte populations identified using CD14 and CD16 staining. A $CD14^{++}CD16^{-}$ classical population, $CD14^{++}CD16^{+}$ intermediate population, and $CD14^{+}CD16^{++}$ population can be identified. Adapted from Ziegler-Heitbrock and Hofer (2013).

In homeostatic conditions classical monocytes constitute the majority of the circulating monocyte pool (approximately 92%) (Nockher and Scherberich, 1998). Evidence indicates that the different monocyte subsets may represent progression from a less to a more terminally differentiated state from classical through to non-classical monocytes respectively (Wong et al., 2011). This is supported by a shift in the expression of genes towards less proliferation promoting genes and more pro-apoptotic genes which results in a 3-fold higher level of apoptosis in $CD16^{+}$ monocytes than in the $CD16^{-}$ subset after 4 hours of *in vitro* culture with no additional stimuli or growth factors added (Malavez et al., 2015).

As well as the differences in surface CD14 and CD16 expression and the expression of proliferation and apoptosis associated genes, monocyte subsets also differ in the chemokines and cytokines that they express that hint at the different roles they may play in the immune response. In humans, nonclassical monocytes are thought to be a pro-inflammatory population as highlighted by their increased expression of genes associated with inflammation (*TNF-α*, *IL-1β*) in response to lipopolysaccharide when compared with the other monocyte subpopulations (Wong et al., 2011). Additionally, the $CD14^{++}CD16^{+}$ intermediate and nonclassical populations are found to be expanded in response to bacterial infections in haemodialysis patients and fall to baseline numbers concurrently with a reduction in symptoms (Nockher and Scherberich, 1998). The $CD14^{++}CD16^{+}$ population also has a higher level of expression of MHC-II, higher phagocytic activity, higher TLR2 expression, and higher levels of CD54 and CD86 than the $CD14^{++}CD16^{-}$ population (Nockher and Scherberich, 1998;

Skrzeczynska-Moncznik et al., 2008). Other proteins expressed differentially include colony stimulating factor receptor 1 (CSFR1), complement component 3a receptor 1 (C3AR1), CXCL16, CD43, and CX3C chemokine receptor 1 (CX3CR1), expressed at higher levels on non-classical monocytes, and colony stimulating factor 3 receptor (CSF3R), CD93, CD1d, and CCR2, with greater expressed on the classical subset (Ancuta et al., 2009; Wong et al., 2011). Additionally both subsets express similar levels of mRNA for *TLR-1*, *2*, *4*, and *5*. The intermediate subset expresses levels of surface markers which are in between those expressed by the $CD14^{++}CD16^{-}$ or the $CD14^{+}CD16^{++}$ subsets again hinting at progressive differentiation from the classical to the non-classical subset.

With regards to soluble factor production the classical subpopulation expresses the widest range of cytokines and chemokines including IL-6, CCL5, IL-10, and TNFα whereas the non-classical subset produces the highest levels of TNF-α in response to lipopolysaccharide (Skrzeczynska-Moncznik et al., 2008; Wong et al., 2011). Functional differences of these subpopulations can be seen when genes are grouped according to their involvement in biological processes. Classical monocytes express genes involved in functions ranging from tissue repair and remodelling to host defence and inflammation, the intermediate subpopulation is enriched for genes involved in antigen presentation and costimulatory molecules (i.e. major histocompatibility complex 2 (*MHC-II*) and *CD40*), and nonclassical monocytes show gene expression associated with motility, phagocytosis, and antibacterial responses (Ancuta et al., 2009; Wong et al., 2011).

	Classical	Intermediate	Non-classical
TNFa	+	++	+++
CSF3R (G-CSF-R)	+++	++	+
CD93	+++	++	+
CD1d	+++	++	+
CCR2	+++	++	+
CSF1R (M-CSF-R)	+	++	+++
C3AR1	+	++	+++
CXCL16	+	++	+++
CD43	+	++	+++
CX3CR1	+	++	+++
Biological function	Tissue repair, inflammation	Antigen presentation, costimulation	Phagocytosis, antibacterial responses

Table 2: Summary of the differences in protein expression and biological function between the three monocyte subsets in humans. Expression levels are relative to the other subsets. +++ = relative high expression, ++ = relative intermediate expression, + = relative low expression.

The differentiation of monocytes into dendritic cells or macrophages has been intensively studied, however due to a lack of consistency between publications there are many approaches taken to generate these cell types some of which arguably reflect more of an activation phenotype due to the stimulus used than true differentiation. Additionally the *in vivo* situation is likely to be significantly different to the approach taken *in vitro* as cells typically encounter a milieu of stimuli consisting of many factors derived from differing sources whereas the *in vitro* approach normally consists of the use of one or two factors in combination for several days duration, a situation unlikely to occur during a dynamic immune response. Nethertheless, studies have successfully demonstrated the ability of monocytes to differentiate into various subpopulations and, particularly in the case of macrophages, that the phenotypes of these cells are highly plastic.

One of the most common methods for differentiating both human peripheral blood monocytes and also murine bone marrow cells into two subsets of macrophages

is to culture the cells in the presence of GM-CSF or M-CSF for approximately 6 days (Fleetwood et al., 2007; Jaguin et al., 2013). Treatment with GM-CSF results in higher expression of the IL-12 p35 subunit (*IL-12p35*), *CXCL10*, *CXCL11*, *CCL5*, *CCR7*, indolamine-pyrrole 2,3-dioxygenase (*IDO*), and *TNFα* mRNA which indicates a pro-inflammatory phenotype whereas treatment with M-CSF results in higher levels of mRNA for *TGFβ*, scavenger receptor class B member 1 (*SRB-1*), peroxisome proliferator-activated receptor gamma (*PPARG*), *CCL14*, mannose receptor C-type 1 (*MRC1*), heme oxygenase 1 (*HO-1*), *CD36*, *CCL2*, and *IL-10* indicating an anti-inflammatory scavenging phenotype. The two macrophage populations that can be differentiated in this manner express 87% of genes at a similar level in humans and 80% in mice with 3044 genes differentially expressed in humans at a fold change greater than two (Lacey et al., 2012). Interestingly, GM-CSF treatment of the murine macrophage subset differentiated with M-CSF prior to stimulation with lipopolysaccharide results in an increase in the production of TNFα from these cells above that of those only differentiated with M-CSF, albeit at lower levels than those cells initially differentiated with GM-CSF. Treatment with M-CSF prior to stimulation results in decreased TNFα and IL-12 p70 subunit (IL-12p70) production demonstrating the plasticity inherent in macrophage subsets (Fleetwood et al., 2007).

GM-CSF differentiated	M-CSF differentiated
IL-12p35	TGFB
CXCL10	SRB-1
CXCL11	PPARG
CCL5	CCL14
CCR7	MRC1
IDO	CCL2
TNFa	IL-10

Table 3: Key genes expressed by either GM-CSF differentiated or M-CSF differentiated macrophages generated from peripheral blood monocytes.

In addition to the well characterised differentiation of monocytes to macrophages it has also been demonstrated that monocytes can be terminally differentiated into functional dendritic cells through the combined action of GM-CSF and IL-4 (Chapuis et al., 1997; Sallusto and Lanzavecchia, 1994). This method of differentiation produces competent antigen-presenting cells with the characteristic veiled morphology and dendritic processes associated with dendritic cells. Interestingly, although the dendritic cells differentiated in this manner are incapable of differentiation towards a macrophage phenotype, treatment of M-CSF differentiated macrophages with IL-4 and M-CSF does convert these cells to a dendritic cell phenotype and approximately 50% of the macrophages differentiated with GM-CSF can also be converted by this treatment, demonstrating a difference in plasticity between these cell types. Monocytes are also the source of osteoclasts which are differentiated using a treatment of RANKL, M-CSF, IL-4, and 1,25-dihydroxy-vitamin D3 (Massey and Flanagan, 1999).

1.4.2.2 Macrophages

Macrophages are known to develop from at least two sources. In early foetal life yolk-sac derived macrophages populate the tissues of the embryo and eventually give rise to tissue-resident macrophages such as microglia (Mizutani et al., 2012; Schulz et al., 2012). The early progenitors of the haematopoietic system reside in the foetal liver before migrating to the bone marrow and developing into haematopoietic stem cells (HSCs) (Sasmono et al., 2003; Tavian and Peault, 2005). HSCs are the source of many of the cells of the immune system including monocytes which are known progenitors of macrophages. Upon encountering specific inflammatory chemokines such as CCL2 and CX3CR1, monocytes exit

the blood stream and infiltrate the tissues, where signals such as GM-CSF and M-CSF may drive predetermined pathways in order to fight specific pathogen groups.

Most studies investigating the origin of tissue-resident macrophages are undertaken in mice as fate mapping and knockout studies and can be used to test hypotheses that cannot be tested in humans. A population of macrophages which can develop independent of the transcription factor Myb, which is required for the differentiation of macrophages from peripheral blood monocytes, were found to originate from the yolk sac during embryogenesis and seeded several long lived tissue resident macrophage populations during the ongoing maturation of the embryo, although precisely how long these cells persist in tissues is as yet unclear (Perdiguero et al., 2014). However, intricate work has demonstrated that microglia derived from a precursor shared with yolk sac macrophages and not from haematopoiesis are capable of self-renewal without hematopoietic input (Ginhoux et al., 2010). Using either sublethal irradiation of a mouse before a bone marrow transplant or parabiosis experiments in which the circulatory systems of two mice are surgically joined, these investigators found that although circulating leukocytes consist of roughly 30% donor-derived cells 1 month after transplant or initiation of parabiosis, only 5% of microglia are donor derived and that even 21 months after the procedures the contribution of donor cells to the microglia population is only 10-20%. These findings indicate that tissue-resident microglia are seeded during embryogenesis, are capable of self-renewal, and persist into adulthood. However the contribution of embryologically derived macrophages to other tissue-resident macrophage populations and the longevity of these

populations is yet to be defined but initial work indicates that lung, peritoneal, spleen, and bone marrow macrophage populations may all be long-lived tissue-resident populations capable of self-renewal without hematopoietic input (Hashimoto et al., 2013).

Several studies have demonstrated that tissue-resident macrophages, such as microglia in the central nervous system, cardiac macrophages, or red pulp macrophages, require distinct transcription factors for development and function and yet show shared expression of markers specific for all tissue-resident macrophages; CD64 and MerTK (Gautier et al., 2012). A striking example of the tissue tropism of tissue-resident macrophages can be found in the splenic macrophage populations. Red pulp macrophages are dependent on Spi-C expression, which is specific to this population and not found in other tissue-resident macrophages, monocytes, or dendritic cells, whereas marginal zone macrophages require Liver X receptor alpha (LXRα) expression (A-Gonzalez et al., 2013; Kohyama et al., 2008). As previously mentioned tissue-resident macrophages appear to be capable of self-replenishment and thus maintain population integrity without input from peripheral blood monocytes, however during inflammation cardiac macrophages undergo not only self-replenishment but are aided in reconstituting the resident population by peripheral blood monocytes again highlighting the disparity between resident macrophage populations (Epelman et al., 2014). Regarding the features responsible for driving the differences between resident macrophage populations tissue-specific factors appear to be critical rather than terminal differentiation of the macrophages as transfer of one resident population into the niche of another results in the

transferred macrophage assuming a phenotype similar to that of the tissue-resident macrophages found in the destination tissue (Gosselin et al., 2014; Lavin et al., 2014).

The other more well defined source of macrophages is through differentiation from monocytes. The current paradigm of monocyte to macrophage differentiation defines two broad classes of macrophages with the M1 class regarded as proinflammatory and the M2 class anti-inflammatory (Buechler et al., 2000; Jaguin et al., 2013; Lacey et al., 2012; Mantovani et al., 2004; Mantovani et al., 2002; Strauss-Ayali, Conrad and Mosser, 2007). M1 macrophages are characterised by the expression of CD86 and MHC-II; they produce cytokines involved in driving inflammation such as TNF-α and IL-1β along with monocyte and activated T lymphocyte recruiting chemokines such as CXCL10, CXCL11, and CCL5 (Ambarus et al., 2012a; Angiolillo et al., 1995; Cole et al., 1998; Dufour, Jennifer H. et al., 2002b; Jaguin et al., 2013; Mantovani et al., 2004; Rey-Giraud, Hafner and Ries, 2012; Verreck et al., 2004). The functions of these markers and soluble factors lead to the portrayal M1 macrophages as a proinflammatory subset.

M2 macrophages have been split into M2a, M2b, and M2c subsets depending upon the repertoire of markers and soluble factors being produced (Mantovani et al., 2004). Broadly speaking M2 macrophages are considered to have anti-inflammatory roles and are important in tissue remodelling and homeostasis. This is exemplified in the markers and factors used to identify the M2 phenotype, including the scavenger receptor CD163, the mannose receptor (CD206), and

key anti-inflammatory mediators such as IL-10 (Ambarus et al., 2012a; Mantovani et al., 2004; Mikita et al., 2011; Verreck et al., 2004).

The limitations of such an artificial paradigm have become apparent with discordance between the various methods of *in vitro* differentiation of M1 and M2 macrophage phenotypes. Furthermore the phenotypes of *in vivo* macrophages are rarely as extremely polarised as *in vitro* differentiated populations, suggesting that macrophages exist across a wide functional and phenotypic spectrum. Efforts to clarify studies on *in vitro* differentiation and functions of macrophages have been made with the publication of guidelines on macrophage activation nomenclature (Murray et al., 2014).

1.4.3 Roles in disease

1.4.3.1 Monocytes

As with other immune cell types the subset distribution and function of monocytes can become skewed during the pathogenesis of several diseases. In the cardiovascular field a raised intermediate monocyte count is significantly associated with a higher risk of adverse outcomes from strokes and myocardial infarctions (Rogacev et al., 2012). In patients suffering from carotid artery disease the levels of mRNA for high mobility group box 1 (*HMGB1*), plasminogen activator inhibitor 1 (*PAI-1*), matrix metalloproteinase 9 (*MMP9*), *TNFα*, and *TGFβ* are significantly higher in monocytes and T cells within plaques than in the peripheral blood, suggesting that these cells may aid in the rupture of plaques (Sternberg et al., 2013). Monocytes isolated from patients suffering from systemic sclerosis, a disease in which fibrosis occurs in the skin and vasculature of patients, have higher levels of the transcription factor Fra-2 which has been implicated in pro-

fibrotic pathways and drives increased expression of tissue inhibitor of metalloproteinase 1 (TIMP-1) in these cells (Ciechomska et al., 2016; Reich et al., 2010).

Early work investigating the role played by monocytes in RA found a significantly higher number of monocytes in the synovial fluid of RA patients expressed CD16 on their surface than monocytes in the peripheral blood from the same patients indicating a shift in the balance of monocyte subsets in the inflammatory joint environment (Wahl et al., 1992). Further work showed that the increase in CD16 expression was mediated mostly but not entirely by TGFβ as this effect is greatly reduced but not completely abrogated by blocking antibodies against TGFβ *in vitro*. Subsequently an increase in the number of $CD16^+$ monocytes in the circulation of RA patients has been shown particularly in active disease; the $CD16^+$ cell count positively correlated with c-reactive protein levels (CRP), a measurement of systemic inflammation. This study showed that not only TGFβ, but also M-CSF and IL-10 can increase CD16 expression on monocytes and that M-CSF levels are significantly higher in the peripheral blood of RA patients than controls (Kawanaka et al., 2002). Interestingly the duration of disease also appears to affect the number of $CD16^+$ monocytes present as higher numbers are found in longer duration RA than in early disease. Although $CD16^+$ monocyte numbers decrease with successful therapy, patients who do not respond well to methotrexate or steroids have a higher proportion of intermediate monocytes relative to the other subsets when compared to good responders (Cooper et al., 2012; Kawanaka et al., 2002). More evidence that the intermediate population is involved in RA pathogenesis was provided by Rossol et al who found significantly

expanded intermediate populations in the peripheral blood of RA. *In vitro* the intermediate monocyte subsets induces significantly greater Th17 cell differentiation, a T cell subset implicated in RA, from memory T cells than the other monocyte subsets and in RA the frequency of Th17 cells positively correlates with the intermediate monocyte population (Rossol et al., 2012).

Regarding the activation of monocytes in RA both synovial macrophages and monocytes in RA have significantly higher basal and interferon gamma (IFNγ) or TNFα stimulated levels of neopterin, which can be used as a marker of monocyte or macrophage activation, compared to monocytes isolated from healthy controls (Hahn et al., 1993). The levels of Siglec-1, a scavenger and phagocytosis receptor, mRNA and protein are also higher in peripheral blood monocytes isolated from RA patients than from patients with osteoarthritis or healthy controls; this expression positively correlates with the DAS28 composite-disease activity score, rheumatoid factor levels, erythrocyte sedimentation rate (a measure of global inflammation levels, ESR), and CRP (Crocker and Gordon, 1986; Jones, Virji and Crocker, 2003; Xiong et al., 2014). Siglec-1 expression on monocytes can be upregulated by several pro-inflammatory stimuli such as IFNγ, TNFα, and also collagen-II. Monocytes and T cells isolated from the peripheral blood of RA patients have significantly higher signal transducer and activator of transcription (STAT3) levels than in healthy controls and significantly higher levels of phosphorylated STAT3 in monocytes and $CD4^+$ T cells (Isomäki et al., 2015). Levels of IL-6 and TNFα found in the plasma of patients and controls significantly correlate with the number of monocytes and $CD4^+$ T cells positive for phosphorylated STAT3, while levels of IFNγ are significantly correlated with the

number of phosphorylated STAT3 positive monocytes. Taken together, this indicates that RA monocytes shift their expression and activation of STAT family members depending on the cytokine milieu to which they are exposed.

1.4.3.2 Macrophages

Macrophages have been shown to play significant roles in RA. The count of $CD68^+$ cells (CD68 being a marker of macrophages) in the sublining region of the synovium is highly correlated with DAS28 score. Falls in DAS28 score due to successful treatment is associated with a concurrent decrease in the sublining $CD68^+$ cell count (Haringman et al., 2005; Kunisch et al., 2004; Micklem et al., 1989; Pulford et al., 1990). For example, the successful RA treatment infliximab which targets TNF-α, the signature cytokine of M1 macrophages, causes a decrease in $CD68^+$ lining layer cells. In another study, treatment with the glucocorticoid prednisolone decreased synovial IL-1β and TNF-α levels and the sublining $CD68^+$ cell count (Gerlag et al., 2004; Haringman et al., 2005; Smeets et al., 2003). Furthermore, the number of CD14 positive sublining cells and the thickness of the lining layer both correlate with the radiological progression of joint damage (Mulherin, Fitzgerald and Bresnihan, 1996).

Differences in the number and marker expression of synovial macrophage subsets between different disease states have been investigated. Ambarus et al. (2012b) investigated synovial macrophage populations using immunofluorescent staining of tissue sections and found that although expression of the pan macrophage marker CD68 was similar between RA and spondyloarthritis in both lining and sublining layers of the synovium, expression of CD163 was significantly higher in the lining layer of spondyloarthrtis than RA. In general lining layer

macrophages expressed higher levels of CD163 and CD32 but little or no CD64, CD80, CD200R, CD14, or CD16 whereas sublining macrophages expressed CD163, CD32, and CD64. However, such findings are not consistent in the literature, it has also been shown that there are an increased number of $CD163^+$ and $CD68^+$ cells in the lining layer of spondyloarthritis patients and an increase in $CD163^+$ cell numbers in the sublining of spondyloarthritic synovium (Baeten et al., 2002; Baeten et al., 2004). Interestingly, the number of $CD163^+$ sublining cells was shown to correlate with CRP and ESR values. In addition to the difference in phenotype between lining and sublining macrophages Fonseca et al. (2002) found that $CD163^+$cells were not in proximity with lymphocyte rich clusters whereas $CD68^+$cells not expressing CD163 were found in and around these clusters in RA synovium.

Studies investigating the levels of soluble CD163 in the plasma of RA patients have shown that CD163 levels are significantly higher in patients with longstanding RA than in healthy controls and in patients who have had a symptom duration of less than 6 months, soluble CD163 levels are correlated with CRP, ESR, and DAS28 (Greisen et al., 2011; Greisen et al., 2015). In patients with the shorter symptom duration the levels of soluble CD163 were not significantly higher than healthy controls, however levels dropped significantly during treatment for 9 months with methotrexate, betamethasone, and cyclosporine and post-treatment levels correlated with radiological progression over a 5 year period. Differences between early and later stage RA have also been observed in the cytokine profile in macrophages with CXCL4 and CXCL7 being expressed at higher levels in the synovium of early RA patients compared

to longer duration RA or healthy controls and with the expression of these cytokines co-localising with the macrophage marker CD68 in synovium as identified by immunofluorescent staining (Yeo et al., 2015a).

In addition to the above markers, activated macrophages present in the skin, lung, liver, spleen, kidneys, intestines, lymph nodes, and synovium all express folate receptor beta (FRβ) (Nagayoshi et al., 2005; Van Der Heijden et al., 2009; Xia et al., 2009). Macrophage FRβ expression is only present in sublining macrophages distant from T cell infiltrate in inflamed synovium. Murine $FR\beta^{+}$ macrophages express higher levels of CD80 and CD86 than $FR\beta^{-}$ macrophages, indicating an increased co-stimulatory capability in this macrophage subset and $FR\beta^{+}$ macrophages also have intracellular TNFα pools and higher levels of reactive oxygen species than $FR\beta^{-}$ macrophages. Curiously, FRβ expression on macrophages differentiated from human peripheral blood monocytes using M-CSF was not elicited when these cells were stimulated with a range of typical stimuli used in *in vitro* experimentation, so the mechanism by which FRβ is upregulated on macrophages in inflamed synovium is as yet unclear.

Subsets of monocytes or macrophages may have disparate roles in the progression of RA as highlighted in the K/BxN serum-transfer mouse model of arthritis. In this model resident synovial macrophages and infiltrating macrophages could be identified on the basis of MHC-II absence or presence respectively (Misharin et al., 2014). Although both of these subsets of macrophage could differentiate from non-classical monocytes, resident macrophages dampened the severity of arthritis and were required for resolution

of the inflammation whereas infiltrating macrophages exacerbated the inflammation. This highlights the complexity of the macrophage subset field which often does not conform to the current M1-M2 paradigm of macrophage polarisation.

1.4.4 Summary

Macrophages are derived from at least two sources with tissue-resident macrophages believed to be derived from yolk sac progenitors during embryogenesis and infiltrating populations differentiated from peripheral blood monocytes. Macrophages and monocytes are thought to act as sentinels monitoring tissue and the circulation for signs of injury or infection and responding both directly and indirectly. Monocytes respond to pathogens with the production of proinflammatory cytokines and are capable of differentiating into macrophage and dendritic cell subsets. Macrophages are important both in the initiation/propagation of inflammation through the production of cytokines such as TNFα and in limiting inflammation through IL-10 production and are important in clearing infections through phagocytosis and scavenger functions and the production of reactive oxygen species. Macrophages also bridge the innate and adaptive arms of the immune system as they are capable antigen presenting cells and express costimulatory cell surface molecules. Although it is clear macrophage subsets exist within the synovium and that these subsets play different roles during RA the precise function of each macrophage marker and/or subpopulation is yet to be clarified.

1.5 Fibroblasts

1.5.1 Development

Fibroblasts are stromal cells that give structure and support to the tissues of the body but also are critically involved in the immune response. During adult life, mesenchymal stem cells and peripheral-blood fibrocytes may contribute to fibroblast populations present in tissues of the body however it is certain that during embryogenesis fibroblasts develop from the primary mesenchyme (Parsonage et al., 2005). Fibrocytes are classified on the basis of CD45, collagen-I, and alpha smooth muscle actin (αSMA) expression and on their contractile ability, are derived from circulating $CD14^+$ precursors in a tissue growth factor β1 (TGF-β1) dependent manner, and play a role in wound healing but their contribution to fibroblast populations is uncertain (Abe et al., 2001; Pilling et al., 2009). Multipotent mesenchymal stromal cells are present in many connective tissues in the body and have the ability to differentiate into several stromal lineages such as fibroblasts, adipocytes, and osteoblasts (Lee et al., 2010; Uccelli, Moretta and Pistoia, 2008). Identification of mesenchymal stem cells is complex and relies on the exclusion of cells expressing CD45, CD11b, CD19, and HLA-DR and on the use of a suite of markers expressed on the stem cells such as CD73, CD90, and CD105. Fibroblasts can be differentiated from mesenchymal stem cells through the action of connective tissue growth factor which results in cells expressing fibroblast markers such as vimentin and collagen-I. However, the difference between fibroblasts and MSCs is unclear, partially due to the heterogeneity of fibroblasts, and some authors hypothesise that fibroblasts are actually activated MSCs as opposed to a unique cell type (Kalluri, 2016).

Synovial fibroblasts have traditionally been viewed as merely providing tissue-specific structure and function to the synovium and joint (Lee et al., 2010). Maintaining the synovial tissue extra-cellular matrix and synovial fluid to permit normal mechanistic function of the joint is achieved through the production of hyaluronan, amongst other products, which helps maintain the correct viscosity of the joint fluid, and through the production of a range of matrix-metalloproteinases (MMPs) and tissue-inhibitors of metalloproteinases (TIMPs) that allow the fibroblast to respond to the high levels of tissue remodelling required in a high trauma environment such as the joint (Aicher et al., 2003; Huang et al., 2011; Vuorio et al., 1982). Due to these functions, fibroblasts were viewed as having a passive role in the immune response. However this viewpoint is changing as emerging evidence shows fibroblasts not only as active players in the immune system with an immune sentinel role similar to macrophages but also as key players in the progression of RA.

1.5.2 Role in orchestrating the immune response

It is now clear that fibroblasts do not simply provide the backdrop against which immune and inflammatory responses carry out their play but are critically involved in the immune response and this is apparent in the viewpoints expressed in recent publications (Filer, 2013; Juarez, Filer and Buckley, 2012; Naylor, Filer and Buckley, 2013; Turner and Filer, 2015). Simple evidence that the stroma responds to and supports the inflammatory response can be seen in the ability of fibroblasts to proliferate in response to the key pro-inflammatory cytokines IL-1β and TNFα, however this is not an unremitting response and the same two cytokines trigger the release of prostaglandin E2 from fibroblasts which acts in an

autocrine manner to inhibit proliferation (Gitter et al., 1989). This demonstrates the existence of regulatory networks in fibroblasts that respond to inflammatory cues.

More direct evidence that fibroblasts are key sentinel cells, in a similar manner to macrophages, is apparent in the fact that pattern recognition receptors are expressed by these cells (Kim et al., 2007b; Kim et al., 2009; Ospelt et al., 2008). Toll-like receptors (TLRs) 1, 3, 4, 5, and 6 are all expressed by synovial fibroblasts, with TLR3 and TLR4 at the highest levels allowing fibroblasts to detect and mount a response to pathogen associated molecular patterns (PAMPs) such as bacterial lipoprotein, double-stranded RNA, lipopolysaccharide (LPS), and flagellin. Upon stimulation with poly(I:C) (synthetic double-stranded RNA), LPS, flagellin, or bacterial lipoprotein synovial fibroblasts release higher levels of IL-6 than dermal fibroblasts. Furthermore, synovial fibroblasts isolated from the afflicted joints of RA patients (RA-FLS) basally express mRNA for immune mediators such as *IL-6*, *IL-11*, and oncostatin M (*OSM*) which are increased upon stimulation with TNF-α or IL-1α, highlighting the ability of fibroblasts to respond not only to pathogens but also in response to factors elicited from other immune cells (Okamoto et al., 1997).

In addition to being able to react to PAMPs fibroblasts also produce a range of proinflammatory cytokines and chemokines such as IL-23p19, and IL-6, and also CCL2, CCL3, CCL4, CXCL8, CXCL10, and CXCL12 which promote the recruitment of neutrophils, monocytes, and T lymphocytes (Bradfield et al., 2003; Gitter et al., 1989; Hayashida et al., 2001; Hwang et al., 2004). IL-17, a cytokine

frequently implicated in the pathogenesis of RA elicits an increase in IL-6 and CXCL8 production from RA-FLS and acts in synergy with TNFα to increase GM-CSF production (Al-Saadany et al., 2016; Hwang et al., 2004; Lubberts, 2015; Parsonage et al., 2008). Furthermore, fibroblasts can also respond to IL-1α and IL-1β with an increase in GM-CSF production. Taken together the ability of fibroblasts to produce and respond to many factors involved in the immune response indicates the degree to which they are involved in orchestrating the immune response.

As fibroblasts are involved in the recruitment and activation of cells in the tissue they can also act to retain leukocytes in the tissue. Through the constitutive expression of CXCL12, CCL2, and CXCL8 RA-FLS can promote the ingress of B cells, T cells, monocytes, and neutrophils through signals mediated by CXCR4, CCR2, CCR4, CXCR1, and CXCR2 on the surface of these populations. When stimulated with peptidoglycan, which signals via TLR2, RA-FLS upregulate CXCL8, CCL5, and CCL8 leading to recruitment of neutrophils, monocytes, and $CD4^+$ T cells to sites of inflammation (Bradfield et al., 2003; Burger et al., 2001; Pierer et al., 2004; Scaife et al., 2004). The repertoire and magnitude of cytokine release from RA-FLS is modified compared to control synovial fibroblasts highlighted by higher levels of CCL2 and CXCL8 that RA-FLS secrete leading to greater monocyte migration compared to osteoarthritis synovial fibroblasts (OA-FLS) or dermal fibroblasts (Hayashida et al., 2001). In addition RA-FLS release more CXCL12 when co-cultured with $CD4^+$ T lymphocytes, mediated by CD40/CD40L interactions and IL-17 signalling; furthermore higher levels of CXCL12 can be found in the synovial fluid of RA patients compared to that of

osteoarthritis patients (Kim et al., 2007a).

In vitro pseudoemperipolesis assays have provided evidence that the range of chemokines produced by synovial fibroblasts is capable of retaining leukocytes within an environment which could possibly facilitate an inappropriate chronic immune response. Pseudoemperipolesis is the migration and retention of cells underneath a monolayer of cells and in the case of monolayers of RA-FLS or OA-FLS these cells facilitate the pseudoemperipolesis of B cells, activated T cells, or natural killer (NK) cells isolated from peripheral blood compared to dermal fibroblasts (Bradfield et al., 2003; Burger et al., 2001; Chan et al., 2008). Interestingly both B and T cells require CXCL12-CXCR4 interactions for this process however B cells also require additional VCAM-1-VLA-4 interactions for successful pseudoemperipolesis. It is important to note that the pseudoemperipolesis process is not simply a method of leukocyte retention but the co-culture of fibroblasts with another cell type results in the modulation of factors produced by both cell types, in the case of NK cells IL-6, IL-15, CCL2, CXCL8, and GM-CSF release is increased in the pseudoemperipolesis assay.

To further exacerbate the effect that inappropriate leukocyte retention may have on immune responses, co-culture of fibroblasts with B cells, T cells, and NK cells also results in increased survival of these cell types which could allow prolonged exposure of tissues to damaging and pro-inflammatory factors. RA-FLS mediate the increased survival of B cells and NK cells through the expression of IL-15 and B-cell activating factor and, when stimulated via TLR3, the expression of a proliferation-inducing ligand (APRIL) and increased B-cell activating factor

(BAFF) expression (Benito-Miguel et al., 2012; Bombardieri et al., 2011; Burger et al., 2001; Chan et al., 2008). IL-15 engages the IL-15 receptor on NK and B cells, the expression of which on B cells is increased by BAFF stimulation, to increase survival of these cells over 6 days from 3% viable cells in monoculture to 50% in co-culture with RA-FLS, an increase which is significantly higher than that elicited by OA-FLS or dermal fibroblasts. APRIL and BAFF released during co-culture also promotes B cell class switching as evidenced by an increase in cytidine deaminase and increases in immunoglobulin A (IgA) and IgG expression. Finally, activated T cells also survive for longer in co-culture with synovial fibroblasts which is facilitated in part by an IFNβ dependent mechanism (Filer et al., 2006).

In addition to aiding and abetting an inappropriate immune response, RA-FLS are thought to be indirectly involved in the bone erosions present in the joints of RA patients through manipulation of the osteoblast/osteoclast axis. This is mediated through the expression of the pro-osteoclast differentiation factor RANKL which is expressed in the lining layer at significantly higher levels in RA synovium than in osteoarthritis and can be induced in RA-FLS but not OA-FLS or dermal fibroblasts through TLR2, TLR3, or TLR4 stimulation and subsequent autocrine IL-1β signalling (Kim et al., 2007b; Kim et al., 2009). Stimulation of these TLRs on RA-FLS with peptidoglycan, lipopolysaccharide, or poly(I:C) and co-culture of these cells with monocytes directly drives the differentiation of monocytes into osteoclasts demonstrating that RANKL expression on RA-FLS can interfere with the balance of osteoclasts within tissues.

1.5.3 Direct role in joint damage

There is compelling evidence that RA-FLS play a direct and key role in driving the progression and the chronicity of the disease. Significantly, RA-FLS are a key cellular constituent of the invasive pannus tissue that actively erodes into cartilage and bone in severe disease. Current evidence incriminates lining layer fibroblasts as being the subset involved in forming and driving pannus invasion into cartilage as fibroblasts in the pannus tissue express the lining layer marker cadherin-11 (Kiener et al., 2009; Lee et al., 2007). Additionally, modulating the expression of cadherin-11 affects the invasiveness of fibroblasts into cartilage with lower expression resulting in lower invasion and vice versa.

In vitro cultured RA-FLS are capable of directly invading and damaging *ex vivo* articular cartilage through the production of MMP9, and cathepsins B, D, and L, and retain this aggressive behaviour after transfer into a leukocyte free environment, namely a severe combined immunodeficiency (SCID) mouse, whereas OA-FLS invasion in this model is minimal (Allard, Maini and Muirden, 1988; Muller-Ladner et al., 1996; Pap et al., 2001; Unemori, Hibbs and Amento, 1991; Xue et al., 1997). Additionally, RA-FLS appear to be capable of migrating from cartilage in one site to cartilage at another, offering a possible explanation for the progressive involvement of additional joints in RA (Lefevre et al., 2009). The tumour suppressor gene p53 is also implicated in RA with studies demonstrating the upregulation of the protein in RA-FLS invading cartilage and that the forced upregulation of p53 increases the invasiveness of normal fibroblasts (Inazuka et al., 2000; Muller-Ladner et al., 1996; Pap et al., 2001; Seemayer et al., 2003). These findings have led to the postulation that the

behaviour of RA-FLS is permanently altered from the normal and maintains the disease state, suggesting possible involvement of epigenetic mechanisms.

1.5.4 Epigenetic and genetic regulation of fibroblasts

Accumulated evidence suggests that RA-FLS have an 'imprinted' aggressive phenotype that may be mediated through disturbances in the epigenetic regulation of these cell types. Epigenetic regulation controls the expression of genes by limiting or encouraging the access of deoxyribonucleic acid (DNA) transcription factors to gene promoters and common mechanisms by which access is regulated include covalent modification of DNA-binding histones including acetylation, methylation, and phosphorylation, and the direct methylation of CpG islands of DNA (Portela and Esteller, 2010). These modifications are mediated by regulatory complexes containing enzymes such as histone acetylases (HATs), histone deacetylases (HDACs), and DNA methyltransferases (DNMTs), that are known to be involved the addition and removal of histone and DNA epigenetic modifications (Wu and Zhang, 2010).

There is evidence that global levels of DNA methylation, an epigenetic modification implicated in the silencing of gene expression, are different in RA-FLS when compared to levels in osteoarthritis (Portela and Esteller, 2010). When comparing total methylation levels RA-FLS have significantly lower levels of methylation than OA-FLS, even after a period of *in vitro* culture (Karouzakis et al., 2009). Several attempts have been made to assess the number of genes regulated by methylation with one study investigating RA-FLS and OA-FLS and finding 575 genes associated with hypomethylated sites and 714 genes associated with hypermethylated sites and an alternative study demonstrating

that treatment of healthy synovial fibroblasts with 5-azacytidine, a demtheylating agent, for 3 months results in a greater than 2-fold upregulation in more than 180 genes and that genes implicated in RA are within this upregulated set (De La Rica et al., 2013; Karouzakis et al., 2009).

Other methods of gene regulation have also become of interest as possible methods by which RA-FLS are imprinted with a phenotype. With the relatively recent developments in the field of microRNAs (miRNAs), small non-coding RNA sequences that can regulate the stability and expression of multiple coding mRNA species, investigations into the possible function of these molecules in the RA-FLS phenotype have begun (He and Hannon, 2004). *miR-155*, *miR-146a*, and *miR-22* are all differentially expressed between RA-FLS and OA-FLS with *miR-155* and *miR-146a* being higher in RA-FLS than OA-FLS and *miR-22* lower (Lin et al., 2014; Long et al., 2013; Stanczyk et al., 2008). *miR-155* downregulates the expression of *MMP3* and *MMP1* upon TLR or IL-1β stimulation and in synovial tissue *miR-155* expression is 8-fold higher in RA than osteoarthritis which may reflect an attempt to restrain the chronic inflammatory state whereas *miR-22* inhibits proliferation and IL-6 production by downregulating Cry61 but is expressed at lower levels in RA than osteoarthritis synovial tissue and hence may be aiding the aggressive phenotype of RA-FLS.

1.5.5 Subsets

As with other tissue resident cell types, such as macrophages, fibroblasts differ in the phenotype and genes expressed depending on the tissue from which they were isolated. Microarray analysis of fibroblast populations by Parsonage et al.

(2003) has demonstrated variation in basal and stimulated expression of genes related to inflammatory processes and work by Filer et al demonstrated a hierarchy to factors that influence the gene expression in dermal, synovial, and bone marrow fibroblasts with the largest effect being dependent on anatomical location from which the fibroblasts were isolated, the response to serum levels in culture media second, and the disease status of the donor, RA or osteoarthritis, last (Filer et al., 2015).

Fibroblast gene expression is not limited to inter-tissue differences but actually varies across tissues of the same type that are distributed throughout the body. When investigating the transcriptome of fibroblasts isolated from 43 anatomical locations, most of which were taken from dermal tissue at different sites of the body, differences in gene expression were apparent and fibroblasts clustered on the basis of these differences with fibroblasts from similar regions of the body clustering together (Chang et al., 2002; Rinn et al., 2006). Differentially expressed genes between fibroblasts isolated from different sites included homeobox (HOX) genes which are known to be involved in embryological patterning and development but perhaps the most interesting finding of this work is that the variance between fibroblasts from the same site in two distinct donors is less than that found between two different sites in the same donor (Chang et al., 2002; Frank-Bertoncelj et al., 2017). The evidence for stromal heterogeneity has resulted in the hypothesis that the stroma and endothelial tissue function together to present a stromal address code through chemokine and adhesion molecule expression that regulates the influx of certain leukocytes from the vasculature into tissues and the persistence of these cells within the tissue or the efflux into the

lymphatics (Parsonage et al., 2005). It is postulated that expression of an improper postcode leading to the retention of pro-inflammatory immune cells and the prevention anti-inflammatory/pro-resolution cells from entering the afflicted tissue for example increased CXCL13 and CCL19 expression could lead to an increase in B lymphocyte infiltration and retention (Rupprecht et al., 2009; Sellam et al., 2013).

In addition to the heterogeneity between fibroblasts from different sites differences can also be seen between cells isolated from different geographical areas within the same tissue. As mentioned earlier the synovium can be stratified into the lining or sublining layer and the fibroblasts present in these two regions not only differ in location but also in the markers they express, indicating several subsets of fibroblasts are present in synovial tissue with the potential for varying roles in RA progression (Turner and Filer, 2015). For example lining layer fibroblasts express cadherin-11 which facilitates adhesion of the lining layer fibroblasts to one another in a tissue which lacks a basal lamina and in RA has been shown to promote cartilage invasion (Assefnia et al., 2014; Kiener et al., 2009; Lee et al., 2007; Valencia et al., 2004). A monoclonal antibody against this target is now entering clinical trials.

Lining layer fibroblasts also express GP38, particularly in inflamed RA tissue, the expression of which is upregulated by TNF-α and IL-1β and downregulated by anti-TNF-α therapy. Lining layer cells have been shown to express markers including CD55 and fibroblast activation protein (FAP) which sublining fibroblasts express at lower or negligible levels except in very early RA disease (Bauer et

al., 2006; Choi et al., 2017; Croft et al., 2016; Del Rey et al., 2014; Ekwall et al., 2011). Sublining fibroblasts express a different set of markers that can be used to identify alternative synovial fibroblast subsets which consists of CD90 (THY1) and CD248 (Endosialin) (Bauer et al., 2006; Macfadyen et al., 2005; Maia et al., 2010; Saalbach et al., 1996; Saalbach et al., 1998). Recent single cell RNA sequencing of human synovial fibroblasts has expanded the number of core subsets to three with a seemingly lining layer subset identified as $CD90^{-}GP38^{+}$ and a $CD90^{+}GP38^{+}$ sublining layer subset actually consisting of two subsets that are either $CD34^{+}$ or $CD34^{-}$ the second of which appears to expand concurrent with an increase in systemic and local inflammatory markers such as ESR or $CD45^{+}$ cell influx into the synovium (Mizoguchi et al., 2017).

1.5.6 Summary

In summary, RA-FLS are involved in the recruitment, retention, and activation of many cells of the immune system within the synovium (Figure 5). RA-FLS can drive the differentiation of immune cell subsets which, in the case of monocyte to osteoclast differentiation, can push the balance of bone remodelling towards degradation and erosion. In addition, RA-FLS are capable of directly damaging articular cartilage through the production of matrix remodelling enzymes and do not require continued stimulation from cells of the immune system to maintain this behaviour. The maintenance of this aggressive phenotype even after prolonged *in vitro* culture is believed to be enforced through a combination of genetic and epigenetic factors. However, the exact number and function of fibroblast subsets within the synovium is still unclear.

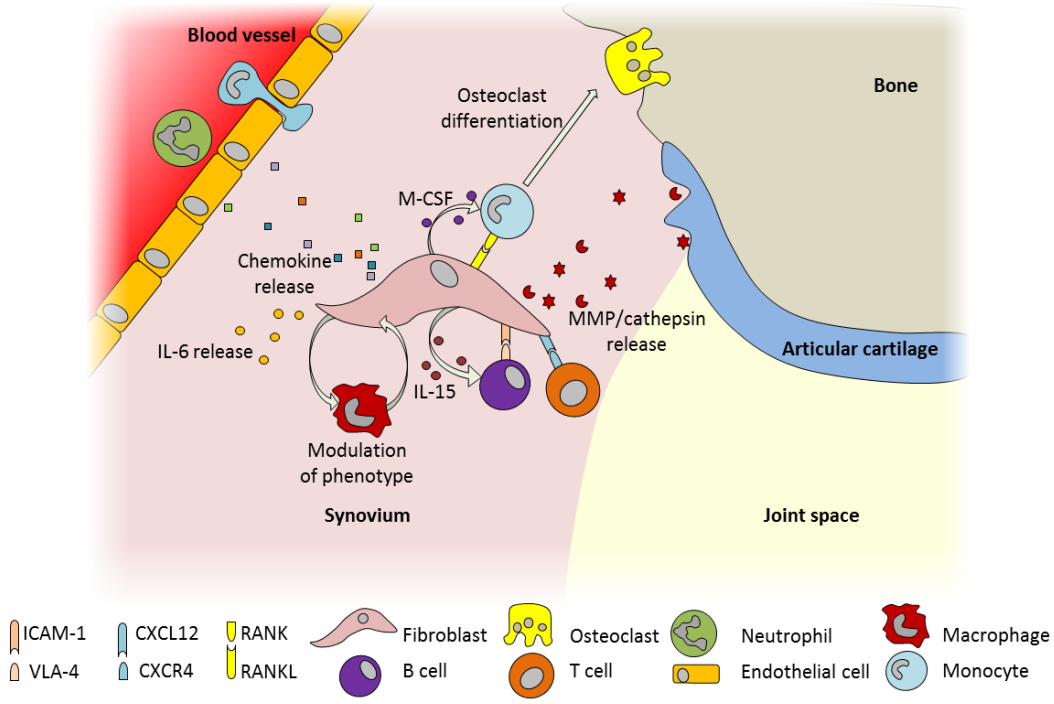


Figure 5: Synovial fibroblasts involvement in inflammation. Chemokines such as CCL2, CCL5, CXCL8, CXCL12, and CXCL13 attract leukocytes into the tissue where fibroblast release of factors such as IL-6 and GM-CSF propagate inflammation. Interactions of B cells and T cells with fibroblasts via CXCL12-CXCR4 and VCAM-1-VLA-4 interactions retains leukocytes in the tissue and provides survival signals reinforced by a positive feedback loop with IL-15 released from fibroblasts. Binding of fibroblast expressed RANK to RANKL on monocytes in combination with M-CSF drives osteoclast differentiation. MMP and cathepsin release damages extracellular matrix and articular cartilage in joints.

1.6 Interactions between macrophages and synovial fibroblasts result in increased invasiveness, and cytokine and growth factor production.

Early studies have highlighted the importance of macrophage/fibroblast interactions in cartilage degradation through the use of *in vitro* assays which found co-culture stimulated the production of cartilage degrading MMPs (Janusz and Hare, 1993; Scott et al., 1997). A subsequent study using *in vitro* fibroblast/macrophage co-cultures highlighted the influence of intercellular adhesion molecule 1 (ICAM-1) and free oxygen radicals on CCL3 production (Steinhauser et al., 1998). Chomarat et al. (1995) demonstrated that co-culture of fibroblasts and macrophages increases IL-6, GM-CSF, leukaemia inhibitory

factor (LIF), and CXCL8 release with the mechanism dependent on IL-1β signalling. During the work undertaken in this thesis it has been published that co-culture of fibroblasts with M-CSF differentiated macrophages modulated gene expression elicited in macrophages by TNFα (Donlin et al., 2014). A set of genes normally increased indirectly in macrophages by TNFα through autocrine IFNβ signalling were suppressed by co-culture with fibroblasts, partially through inhibition of macrophage IFNβ production, however the full mechanism is still unclear. Additionally 34% of the genes normally down-regulated by TNFα in macrophages were upregulated by greater than 2-fold during co-culture. Pathway enrichment analysis of the genes modulated in macrophages by co-culture found enrichment in interferon and MYC related pathways in the downregulated genes and enrichment of growth factor and colony stimulating pathways in the upregulated genes.

Despite these findings the interactions between synovial fibroblasts and macrophages have not been explored in great depth, and the key molecules mediating communication between these cells have not been established. This study aimed to address this issue through the development of a suitable *in vitro* system to allow detailed analysis of the interactions between synovial fibroblasts and macrophages.

An advantage this study has over others that may aim to investigate fibroblast macrophage interactions is access to fibroblasts from samples within the Birmingham Early Arthritis Cohort (BEACON). Samples are donated by patients within this cohort at different stages of disease progression, ranging from early

undifferentiated arthritis through to destructive RA that requires joint replacement, and includes a group of patients with actively inflamed joints whose disease subsequently resolves during follow-up, providing a unique “inflamed control” group. Tissue samples are obtained either at surgery or via minimally invasive, imaging guided procedures. Synovial fibroblasts are routinely isolated and cultured from tissue samples which will allow this study to investigate any differences in interactions between macrophages and fibroblasts that may arise between different disease states, and at different stages of the progression of RA. Additionally, monocytes/macrophages will be readily available throughout the study through the use of peripheral blood apheresis cones commercially available from the NHS Blood and Transplant Service (NHSBT). Peripheral blood monocytes will be used in this study due to the ease of obtaining large numbers of these cells.

1.7 Birmingham Early Arthritis Cohort (BEACON)

BEACON is a well characterised cohort recruiting treatment naïve patients presenting with inflammatory arthralgia or new inflammatory arthritis. The patients are followed longitudinally allowing those presenting with undifferentiated arthritis to be assigned an exact diagnosis and for sequential blood, urine, and tissue samples to be collected during this period along with detailed clinical data. The key outcome groups in the BEACON include treatment naïve new onset very early RA (up to 3 months symptoms, VeRA), treatment naïve resolving arthritis (up to 3 months symptoms, Res), longer duration treatment naïve RA (>three months symptoms, EstRA) and long duration RA disease (from joint replacement surgery, JRep) in addition to control tissue from

patients with musculoskeletal symptoms but no macroscopic or microscopic evidence of joint inflammation (normal, Norm) (Figure 6).

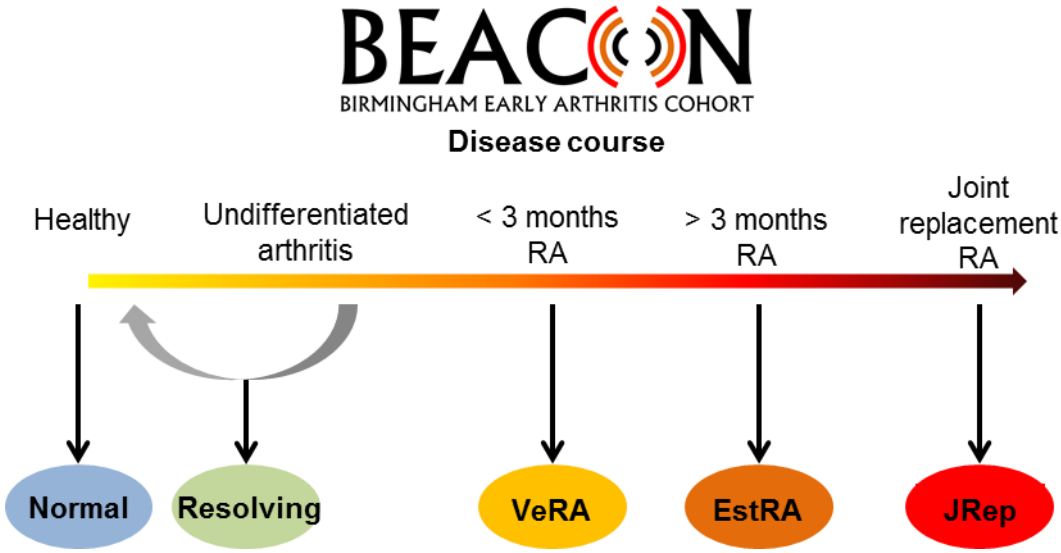


Figure 6: Schematic of patient outcomes included in the Birmingham Early Arthritis Cohort. BEACON recruits patients presenting with undifferentiated arthritis and collects serum, peripheral blood, urine, synovial fluid, and synovial tissue from the patients before following disease progression and assigning patients to disease outcome groups. In addition to BEACON patients synovial tissue is collected from patients undergoing arthroplasty and also from patients with arthralgia in the absence of macro- or microscopic signs of inflammation in the synovium.

1.8 Technologies

Whole transcriptome RNA-sequencing and cytometry by time of flight (CyTOF) are two advanced technologies that will be used to investigate synovial fibroblasts and macrophages within this thesis. RNA-sequencing facilitates the untargeted measurement of the transcriptome of cells as opposed to the probe-based approached offered by microarray and also has a lower limit of detection and higher dynamic range than microarrays (Wang, Gerstein and Snyder, 2009). Briefly, RNA-sequencing requires the conversion of isolated RNA into complementary-DNA (cDNA) libraries either with or without amplification depending on the amount of RNA isolated. Before library preparation RNA not of

interest is removed, for example ribosomal RNA or immature RNA species, and during library synthesis adapters are ligated to the ends of sequences to prepare for next generation sequencing (NGS). Depending on the technology variations in the approach to NGS exist but in most approaches the cDNA sequences act as templates upon which nucleotides are bound and sequences generated by polymerases (Holt and Jones, 2008). In the fluorescence based approach nucleotides labelled with fluorophores to facilitate identification are incorporated into the new sequence generated and the fluorphore, which corresponds with the base added, is released allowing measurement. In the luminescence based approach a solution consisting of one type of nucleotide bound to pyrophosphate is added to the sequences, if the nucleotide is incorporated by polymerases the pyrophosphate is released and used to generate a luminescent signal. The excess solution is then washed away before addition of a solution containing another nucleotide and the process is repeated. In this manner the cDNA library is sequenced and the resulting sequences produced, or reads, are aligned with the genome and quantified.

CyTOF results in similar data to those obtained by traditional flow cytometry but adds the possibility to analyse a much higher number of variables than previously possible. In short, rather than fluorescently labelling antibodies pure metals are conjugated to antibodies and cells are stained using these antibody-metal conjugates (Bandura et al., 2009). The cells are then nebulised, metals ionised, and analysed using time-of-flight mass spectroscopy, all embedded within equipment known as a mass cytometer. CyTOF can simultaneously analyse a higher number of markers than flow cytometry due to the distinct peaks produced

by the metal ions and the lack of spectral overlap seen with fluorophores.

1.9 Hypotheses

This work aims to test two hypotheses:

During the later stages of rheumatoid arthritis synovial fibroblasts develop a chronic inflammatory phenotype which, through the production of soluble and membrane bound factors, drives the pro-inflammatory response/differentiation of monocytes and macrophages to perpetuate disease progression.

Distinct subsets of synovial fibroblasts and macrophages exist within the synovium with a lining subset responsible for driving cartilage and bone destruction in rheumatoid and the regulation of inflammation and leukocyte infiltration driven by a sublining subset.

Although published work has previously investigated differences between RA and healthy controls or osteoarthritis the RA samples used to isolate fibroblasts typically are isolated from longer duration patients who have received multiple different treatments and undergone arthroplasty which immediately limits the findings of these studies to later stage, destructive RA from which extrapolation to the earlier stages of disease may not be appropriate. This work is uniquely positioned due to access to the Birmingham Early Arthritis Cohort which provides a range of samples from different stages of RA and undifferentiated arthritis which allows a more longitudinal approach to investigating the pathogenesis of RA. This

cohort will allow expansion of the current knowledge on fibroblast and macrophage interactions into the interactions during the progression of rheumatoid arthritis (Figure 7).

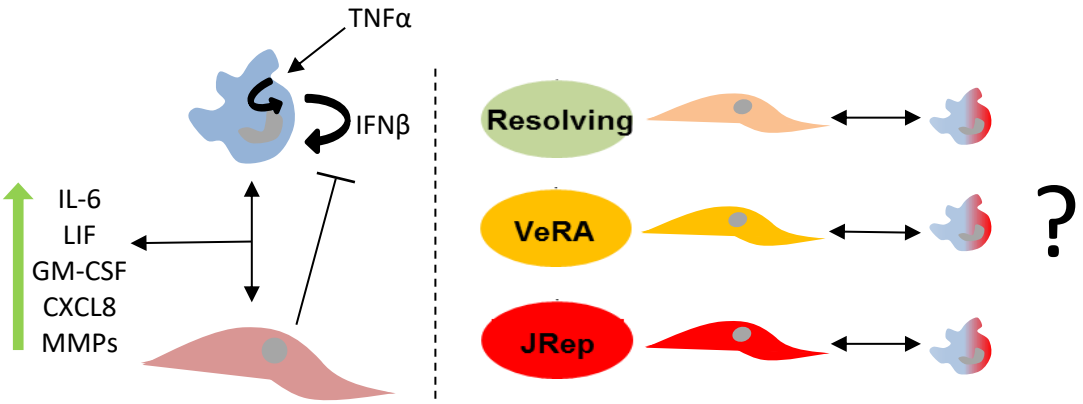


Figure 7: Current knowledge on the interactions between fibroblasts and macrophages. Although information is known regarding the cross-talk between fibroblasts and macrophages (left-hand side of the figure) the influence fibroblasts may have upon monocytes and macrophages during RA progression is unclear (right-hand side of figure).

In addition to addressing the issue of fibroblast-monocyte interactions the BEACON provides a rich source of synovial tissue from the different stages of RA and resolving arthritis which will be used to expand the current model of synovial fibroblast and macrophage subsets in RA and other arthritides by identifying and characterising subpopulations within the tissue (Figure 8).

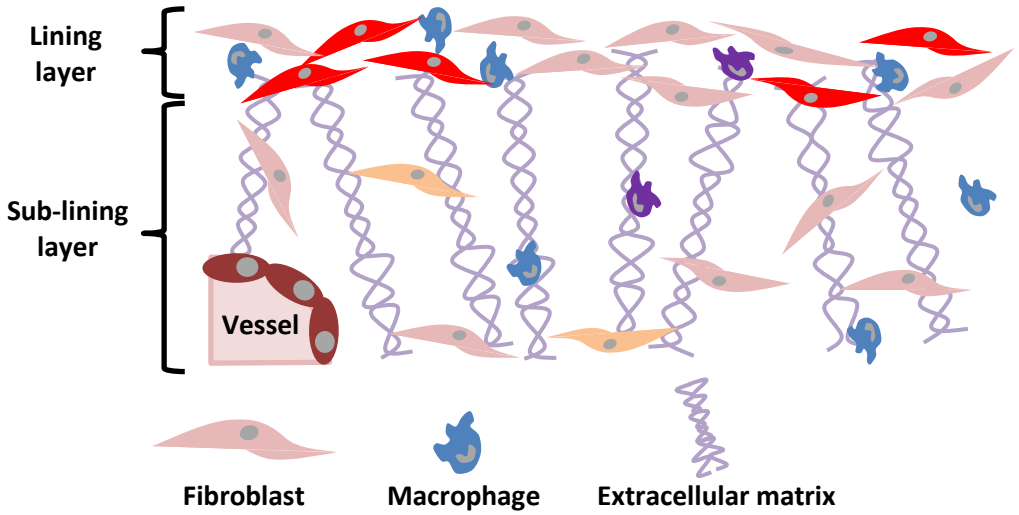


Figure 8: Schematic of the synovium highlighting the current understanding of fibroblast and macrophage populations within the tissue. Although fibroblasts are known to separate into the lining and sublining layers it is not known if the fibroblasts differ in function or if other subsets, represented here by different colours, exist in both regions of the tissue. Additionally, it is not known if different subsets of macrophage are present in the synovium

The three key aims that will be addressed to test the hypotheses are:

Fibroblasts isolated from different stages of RA progression maintain differential gene expression during *in vitro* co-culture and respond to stimuli in different ways.

Fibroblasts isolated from different stages of RA progression differentially modulate monocyte or macrophage phenotype *in vitro.*

Discrete fibroblast and macrophage subsets exist within the synovium with distinct functions.

2 MICROARRAY ANALYSIS OF *IN VITRO* CULTURED SYNOVIAL FIBROBLASTS

2.1 Introduction

To address the first aim a cDNA microarray dataset generated by a previous PhD student was used to analyse gene expression in *in vitro* unselected cultured synovial fibroblasts from the BEACON cohort allowing interrogation of the differences between disease stages (Juárez Pérez, 2014). Previously this dataset was limited to fibroblasts cultured in 10% FBS in the presence or absence of TNFα. Subsequent to Dr Juarez's PhD this dataset was expanded to include samples cultured in 0.1% FBS with no TNFα, an important group given the known impact of serum on the transcriptome of cultured fibroblasts and the possibility that this effect may mask disease specific differences (Filer et al., 2015). Longer duration treatment naïve RA patients were also added to the samples in all treatment groups. In this part of my PhD, this larger dataset was reanalysed, from scratch, including all quality control and data preparation steps, to test the hypothesis that fibroblasts isolated from different stages of RA progression maintain differential gene expression during in vitro co-culture and respond to stimuli in different ways. For this analysis specific comparisons were focused on within each treatment group; (1) the differences between Jrep samples and all other outcomes, (2) the differences between very early RA and resolving samples, and (3) the differences between normal samples and all other outcome groups.

It is known that fibroblasts isolated from late stage RA patients undergoing joint replacment retain an 'imprinted aggressor' phenotype (Bottini and Firestein, 2013). However it is not known at which stage these fibroblasts become imprinted

or if the imprinted phenotype remains the same from early to late RA. This chapter aimed to address this knowledge gap using the aforementioned microarray dataset on cultured synovial fibroblasts from the BEACON. The key questions that were addressed were:

1. Do synovial fibroblasts isolated from RA patients undergoing joint replacement differ in gene expression from those from early RA and non-arthritic fibroblasts?
2. Do synovial fibroblasts from very early RA show differential gene expression compared to fibroblasts from resolving arthritis and hence demonstrate features specific to RA compared to other arthritides?
3. Do synovial fibroblasts from non-arthritic patients differ from all other BEACON groups and what knowledge does this impart regarding important inflammatory pathways in RA synovial fibroblasts?

2.2 Materials and methods

The wet lab work and running of the microarray technology discussed herein was completed as a component of a PhD thesis completed by a previous student, Maria Juárez Pérez (Juárez Pérez, 2014) with the microarray processing and raw data delivery completed as a commercial service by Oxford Gene Technologies.

2.2.1 Culture of synovial fibroblast cell lines

Synovial tissue collected from BEACON patients was used to generate synovial fibroblast cell lines via an outgrowth protocol. From a patient 8 small tissue biopsies were taken, to account for heterogeneity within the synovium, and suspended in 6 ml of complete fibroblast medium (Roswell Park Memorial

Institute 1640 (RPMI 1640, Sigma), 10% v/v foetal bovine serum (FBS, Labtech), [0.87X] minimal essential medium non-essential amino acids (Sigma), 0.87mM sodium pyruvate (Sigma), 1.75mM L-glutamine, 87U/ml penicillin, 87µg/ml streptomycin (added as a mixed L-glutamine-penicillin-streptomycin solution, Sigma)). The suspension was then transferred to a T25 flask and incubated at 37ºC 5% CO_2 for 1 week after which the medium was replaced with fresh medium weekly. Cells begin to grow out from the tissue within 1-2 weeks and reach 80% confluency within a period of 6 months. The weekly feeding cycle was continued with 33% of the medium (conditioned medium) kept in the flask at each feed and 66% replaced with fresh complete fibroblast medium. Upon reaching 80% confluency the cells were passaged by removing the medium, washing in phosphate buffered saline (PBS, Oxoid), and trypsinising by incubating with 10x Trypsin ethylenediaminetetraacetic acid (EDTA) solution (Sigma) diluted to 2x with PBS for 5 minutes at 37ºC. The Trypsin was then inhibited by adding complete fibroblast medium to the suspension before centrifuging at 300g for 6 minutes. The cell suspension was then seeded into two T75 flasks and again cultured until 80% confluent. The cells were then split in a 1:3 ratio with vials of cells frozen in 10% dimethyl sulphoxide (DMSO, Sigma) with 90 % FBS and stored at -140°C when possible. Cells were used for experiments between passages 3 and 7.

2.2.2 Treatment of fibroblasts

The protocol used by Maria Juárez Pérez to generate samples for analysis of gene expression by microarray analysis is as follows: fibroblasts were cultured (section 2.2.1) until an 80% confluent monolayer was reached during passage

three with three flasks of the same fibroblast line. The medium was then removed from the cells before washing with PBS and adding new medium to each flask; one flask with 6ml of complete fibroblast medium (10% FBS v/v), one flask with 6ml of complete fibroblast medium (10% FBS v/v) containing 10ng/ml TNFα, and one flask with 6ml fibroblast medium containing FBS of 0.1% v/v. All flasks were incubated at 37^{o}C 5% CO_2 for 24 hours before another PBS wash, removal of the cells by incubating with 2xTrypsin EDTA for 5 minutes, and producing cell pellets at a concentration of $0.5x10^6$ cells/pellet. Cell pellets were then stored at -80^{o}C until RNA isolations were performed. RNA was subsequently isolated using the RNeasy and QIAshredder kits (Qiagen), and RNA clean and concentrator 25 kit (Cambridge biosciences) all according to the manufacturer's instructions. RNA quantity and quality were assessed using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific). Clinical details of the fibroblast lines used are presented in Table 4.

ID	Group	Disease classification	Site	CCP +/-	RF +/-	Disease duration (Weeks)	Age	Sex	DMARD	Prednisolone	DAS28 ESR baseline	ESR (mm/ hr)	CRP (mg/l)
BX003	VeRA	RA	Knee	+	+	4	50	M	None	None	5.7	31	26
BX004	Res	UA	Knee	-	-	3	33	M	None	None	3.6	9	6
BX005	VeRA	RA	Knee	-	-	5	70	F	None	None	6.0	68	26
BX008	Res	UA	Knee	-	-	6	64	M	None	None	4.5	24	15
BX010	Res	Parvovirus	Knee	-	-	4	40	F	None	None	3.9	5	0
BX011	VeRA	RA	Knee	-	-	2	49	F	None	None	4.7	12	8
BX013	VeRA	RA	Knee	-	-	10	45	F	None	None	3.8	24	12
BX015	VeRA	RA	Knee	-	-	2	48	F	None	None	3.5	4	102
BX016	RA	RA	Knee	+	+	150	46	M	None	None	6.7	34	7
BX020	VeRA	RA	Knee	-	-	6	59	M	None	None	5.0	14	22
BX024	Res	Reactive	Knee	-	-	7	32	M	None	None	2.9	10	10
BX028	Res	RA	Knee	-	-	5	74	M	None	None	4.8	45	13
BX030	Res	RA	Knee	-	-	8	72	M	None	None	3.6	5	0
BX031	VeRA	RA	Knee	-	-	9	43	M	None	None	6.9	58	0
BX032	RA	RA	Ankle	+	-	30	46	F	None	None	5.2	24	32
BX033	Res	Pseudogout	Ankle	-	-	7	81	F	None	None	6.7	60	52
BX038	Res	Parvovirus	Knee	-	-	1	45	F	None	None	4.0	4	0
BX040	VeRA	RA	MCP	-	+	10	56	M	None	None	5.2	5	0
BX042	VeRA	RA	Knee	+	-	4	55	M	None	None	3.5	58	45
BX045	VeRA	RA	Ankle	+	+	2	42	F	None	None	8.3	54	40
BX048	Res	Reactive/post strep	Knee	-	-	2	35	M	None	None	4.1	51	7
BX049	VeRA	RA	Ankle	+	+	3	48	F	None	None	3.9	10	0
BX054	Res	UA	Ankle	-	-	6	55	M	None	None	3.5	2	6
BX055	RA	RA	Knee	-	-	26	64	M	None	None	5.4	13	17
BX063	VeRA	RA	Knee	+	-	9	74	F	None	None	4.4	20	32
BX064	Res	UA	Knee	-	-	1	35	M	None	None	1.6	2	9

BX065	Res	UA	Knee	-	-	7	37	F	None	None	3.3	7	0
BX066	RA	RA	Knee	+	+	38	67	M	None	None	4.2	29	48
BX070	Normal	Normal	Knee	-	-	NA	44	M	None	None	NA	NA	NA
BX071	Res	Parvovirus	Ankle	-	-	4	41	F	None	None	1.7	5	9
BX072	Res	UA	Knee	-	-	10	32	M	None	None	3.6	15	0
BX075	RA	RA	Knee	-	-	52	22	F	None	None	6.4	81	79
BX076	Res	Reactive	Knee	-	-	6	28	M	None	None	4.5	18	8
BX077	RA	RA	Knee	-	-	38	72	M	None	None	7.4	53	43
BX081	Normal	Normal	Knee	-	-	NA	59	F	None	None	NA	NA	NA
BX082	Normal	Normal	Knee	-	-	NA	49	F	None	None	NA	NA	NA
BX083	Normal	Normal	Knee	-	-	NA	42	M	None	None	NA	NA	NA
BX084	VeRA	RA	MCP	+	+	6	49	M	None	None	6.8	25	18
BX085	Normal	Normal	Knee	-	-	NA	38	M	None	None	NA	NA	NA
BX086	RA	RA	Ankle	+	+	38	52	F	None	None	6.5	26	52
BX087	Res	Reactive	Ankle	-	-	4	27	M	None	None	3.8	37	28
BX088	Normal	Normal	Knee	-	-	NA	34	M	None	None	NA	NA	NA
BX089	Normal	Normal	Knee	-	-	NA	38	F	None	None	NA	NA	NA
BX092	VeRA	RA	MCP	+	+	4	48	M	None	None	6.0	63	38
BX093	RA	RA	Knee	+	+	156	65	F	None	None	5.3	72	81
BX094	Normal	Normal	Knee	-	-	NA	46	F	None	None	NA	NA	NA
BX113	RA	RA	Ankle	-	-	53	65	M	None	None	5.0	50	10
BX118	RA	RA	MCP	+	+	114	64	F	None	None	6.1	46	10
RA06SY	Jrep	RA	Knee	NA	+	1560	60	M	SSA	None	6.399599	54	75
RA15SY	Jrep	RA	Knee	NA	-	1144	37	M	MTX, Leflunomide	None	5.055023	46	50
RA20SY	Jrep	RA	Knee	NA	+	1040	72	M	Leflunomide	None	4.83098	18	3
RA23SY	Jrep	RA	Knee	NA	+	572	58	F	MTX	None	6.726424	23	32
RA25SY	Jrep	RA	Knee	NA	+	1560	53	F	HCQ	10mg	3.6	NA	NA
RA28SY	Jrep	RA	Knee	NA	-	1040	41	F	Gold	10mg	3.3	20	26

RA29SY	Jrep	RA	Knee	+	+	364	67	F	Etanercept, MTX	7.5mg	6.6	57	66
RA37SY	Jrep	RA	Knee	+	+	1196	54	F	None	5mg	5.9	45	57

Table 4: Clinical characteristics of the patients from which fibroblast lines were isolated and used for the microarray experiments. NA = Not available, MTX = Methotrexate, DAS28 = Disease activity 28 score, ESR = Erythrocyte sedimentation rate, CRP = C-reactive protein, CCP = Anti-cyclic citrullinated peptide, RF = Rheumatoid factor, MCP = Metacarpophalangeal, UA = Undifferentiated arthritis, HCQ = Hydroxychloroquine, SSA = Sulfasalazine.

2.2.3 Microarray preparation and data generation

Two-colour Sure Print G3 gene expression 8x60k v2 (Agilent) microarrays were completed by Oxford Gene Technologies. The RNA isolates generated in 4.2.1 were converted to cDNA using the Low Input Quick Amp Two-Colour Labelling kit (Agilent) and in parallel Stratagene universal human reference RNA (Agilent) was also converted to cDNA and the sample and reference cDNA were labelled with Cy3 and Cy5 fluorophores respectively. Sample and reference cDNA were mixed in equal ratios and hybridised to the microarrays before reading the slides using an Agilent G2505C scanner and generating data for further analysis using Agilent Feature Extraction version 10.7.3.1. During generation of the data pre-processing steps were completed consisting of locally weighted scatterplot smoothing (LOESS) normalisation to correct for dye-bias and background correction using the subtract method of subtracting background values from the relevant foreground values.

2.2.4 Analysis of microarray data

All quality control and analysis from this point onwards was completed by the author of this thesis. Analysis was completed using R 3.4.1 and RStudio 1.0.153. Feature extraction files generated during imaging of the arrays were loaded into the R session and analysed using the limma package (Ritchie et al., 2015). Due to the functions which generate log2 ratio data requiring the reference sample to be in the green channel and the test sample in the red channel, the data in the columns was swapped before any further analysis. The data was subjected to within-array normalisation using the LOESS method and background correction using the ‘normexp’ method before a final between-array normalisation step using

the ‘G-quantile’ method. Following normalisation, quality checks were completed using the arrayQualityMetrics package which automates the production of several quality metrics for all of the arrays and additionally through the generation of heatmaps for all log2 ratios for the probes in the original layout of the array (allowing the identification of artefacts such as fibres on the arrays) (Kauffmann, Gentleman and Huber, 2009) . Samples showing abnormal MA-plots or that were highlighted as outliers by the arrayQualityMetrics package were excluded from further analysis leaving 148 samples. Replicate probes were then averaged and all probes were non-specifically filtered by calculating the standard deviation of each probe across arrays and keeping the top 60% of probes ranked according to variation. Differential expression analysis was completed using the limma package by fitting a linear model. Due to the lack of an array annotation package for the Agilent arrays with ID AMADID039494 the packages human.db0 and AnnotationForge were used to generate an AMADID039494.db package using the annotation information obtained from the Agilent website.

2.3 Results

2.3.1 Data preparation and exploration

The data used in this chapter were generated from 165 two-colour Agilent microarrays with the green colour of each array corresponding to a universal control hybridised to all arrays and the red colour to an individual sample for each array. Individual samples were divided amongst three treatment groups; TNFα and high-serum stimulated, high-serum stimulated, and low-serum stimulated (see section 2.2.2). The 5 BEACON disease outcome groups were joint replacement (Jrep), established RA (EstRA), very early RA (Early), resolving arthritis (Res), and normal non-inflamed control (Norm) (see section 1.7).

As any large datasets, it was important to explore the structure of the microarray dataset and filter out poor quality samples as the presence of these could lead to spurious results when investigating biological differences. A total of 165 arrays were normalised using a two-step process, first the arrays were corrected for background noise and probes normalised within arrays by fitting locally weighted scatterplot smoothing (LOESS) (Figure 9 **A** and **B**). The arrays were then normalised between each other by aligning the distribution of the green channel intensities and adjusting the red channel accordingly using quantile normalisation. This approach rests on the assumption that the same reference cDNA was used on the green channel of every array, such that the intensity of this channel on each array should have the same distribution (Figure 9 **C**).

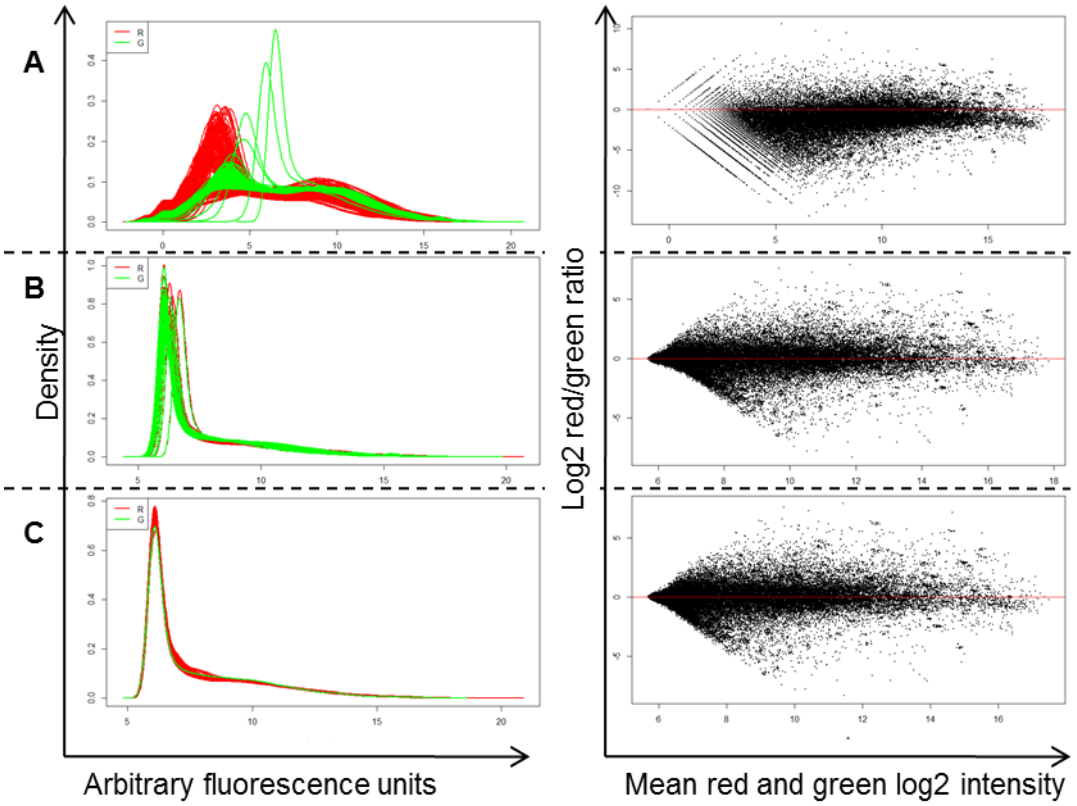


Figure 9: Density plots of the intensity of Red and Green channel fluorescence for all probes on each array (left-hand side) and plots of the log2 ratios of green and red fluorescence against the mean red and green log2 intensity for each probe in the array (known as MA-plots, right-hand side) for sample BX003 grown in high serum with no additional stimuli. Plots A (top row) show the raw non-normalised data, B (middle row) show the LOESS normalised within array data, and C (bottom row) show the within array and G-quantile between array normalisation. The horizontal line on the MA-plots is to aid visualisation of the shape of the plotted data.

Following normalisation, quality control of the array data using the arrayQualityMetrics package was completed. This package calculates distances between arrays, calculated as the mean absolute difference between the arrays when comparing all probes, and flags arrays as outliers when the sum of distances of one array to all others is exceptionally large, compares the distribution of the log2 red and green channel intensity ratios (referred to as log2 ratio from here onwards) to the pooled log2 ratio distribution of all arrays and flags arrays as outliers based on the Kolmogorov-Smirnov statistic of these comparisons, and compares the joint distribution of both the M and A values of

MA-plots for each array using Hoeffding's statistic for outlier detection. In addition to the automatic outlier detection all MA-plots were manually checked for unusual distributions and arrays with concerning MA-plots were flagged. Using this combination of manual and automated outlier detection 17 arrays were detected as outliers and excluded from further analysis leaving 148 arrays.

In a previous analysis of this dataset (Juárez Pérez, 2014) it was found through examination of log2 ratios of sex-specific genes that several samples appeared to have expression levels inconsistent with the biological sex recorded in the clinical data for the patient from which the sample was derived. This analysis was repeated here through comparison of the log2 ratios of *XIST*, a gene which is involved in X-inactivation and expressed at higher levels in females, between arrays (Herzing et al., 1997). As several of the probes on the arrays target *XIST* the mean log2 ratio of these probes was calculated for each array and the results plotted as a frequency histogram (Figure 10). The resulting distribution is bimodal as expected with a higher population representing the female samples and a lower population representing the male samples. In order to compare the sex according to gene expression with that recorded in the clinical data a cutoff of -1.2 for mean *XIST* expression was used as the separation between male and female samples, samples with mean *XIST* log2 ratios below this value were classified as males and those with mean *XIST* log2 ratios above classified as females and the resulting classification compared with that present in the clinical data. Using this approach RA29SY, BX113, BX093, BX084, BX077, and BX072 were identified as having an incorrect sex. As it is unknown if the samples were mislabelled during experiments or the expression of *XIST* does not match the

expected pattern due to an unknown chromosomal abnormality the samples were excluded from further analysis leaving 132 arrays after removal of the 16 abnormal arrays.

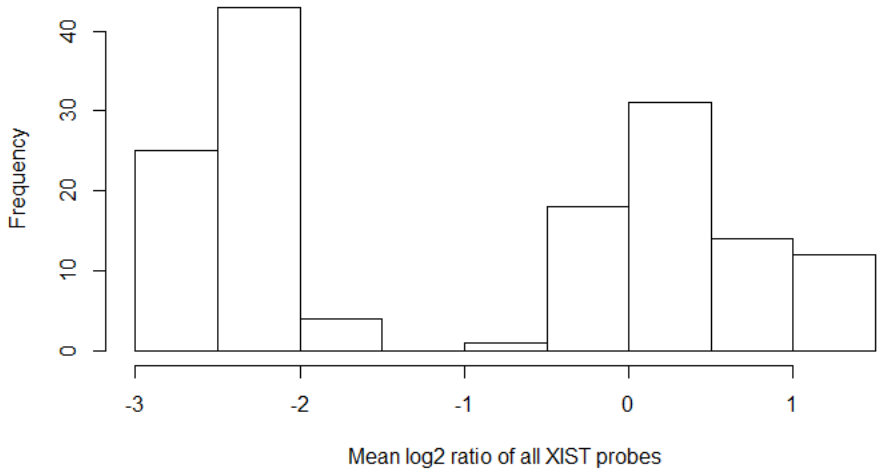


Figure 10: Frequency distribution of the mean log2 ratio of probes targeting *XIST* on each array. The distribution was clearly bimodal as would be expected of a dataset containing male and female samples. The samples with high *XIST* expression are biological females and the samples with low expression are biological males.

As a final quality control step the log2 ratios of each probe on an array were plotted as a heatmap according to the original layout of the probes on the array. This technique allowed visualisation of any gross artefacts that may have been present on the arrays to see if any compromised arrays had passed the quality control steps. Using this approach three arrays were found to have small artefacts affecting the log2 ratios of a small number of probes (an example can be seen in Figure 11). As the number of probes affected on each array appeared to be small it was decided to keep the arrays in the final dataset.

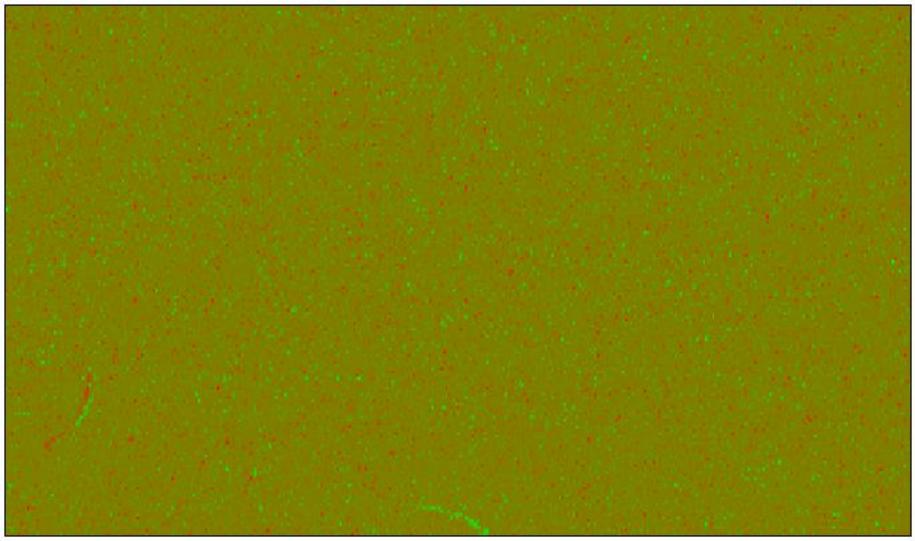


Figure 11: Heatmap of log2 ratios plotted according position of the probes on the array. Green represents a low log2 ratio whereas red represents high. The array shown is for the sample BX031 grown in low serum with no stimuli. Small artefacts could be seen affecting some probes as small red or green 'threads' visible on the array.

As final steps of data preparation replicate probes identified using the Agilent probe ID were averaged (mean) reducing 62,976 probes to 50,739 and these probes were then subjected to non-specific filtering by calculating the standard deviation of each probe across the arrays and keeping the top 60% of the probes when ranked from largest to smallest standard deviation. All these quality control and data preparation steps resulted in a final probe count of 30,443 with the number of samples remaining in each biological group of interest shown in Table 5.

	TNF	High serum	Low serum
Jrep	6	7	6
EstRA	6	6	5
Early	12	12	9
Res	14	15	12
Norm	7	8	7

Table 5: Number of samples remaining in each biological group of interest after quality control and data preparation

Principal component analysis (PCA) of the curated microarray dataset found the largest sources of variation to be the treatments the synovial fibroblasts were subjected to prior to the microarray procedure (Figure 12). The first two principal components accounted for 40% of the variation in the dataset and clearly separated the data into three clusters consisting of the TNF/serum treated samples as would be expected. As hypothesised the diagnostic groups of the patient from which the fibroblast lines were isolated also contributed to the variation in the dataset albeit across several different principal components (Figure 13). The normal samples were separated from other samples along principal components four and five whereas the joint replacement samples separated along components eight and nine. Interestingly, three or four rough clusters could be visualised in principal components 5 and 6 consisting of the Norm samples, the Early and Res samples, and the Jrep and EstRA samples respectively.

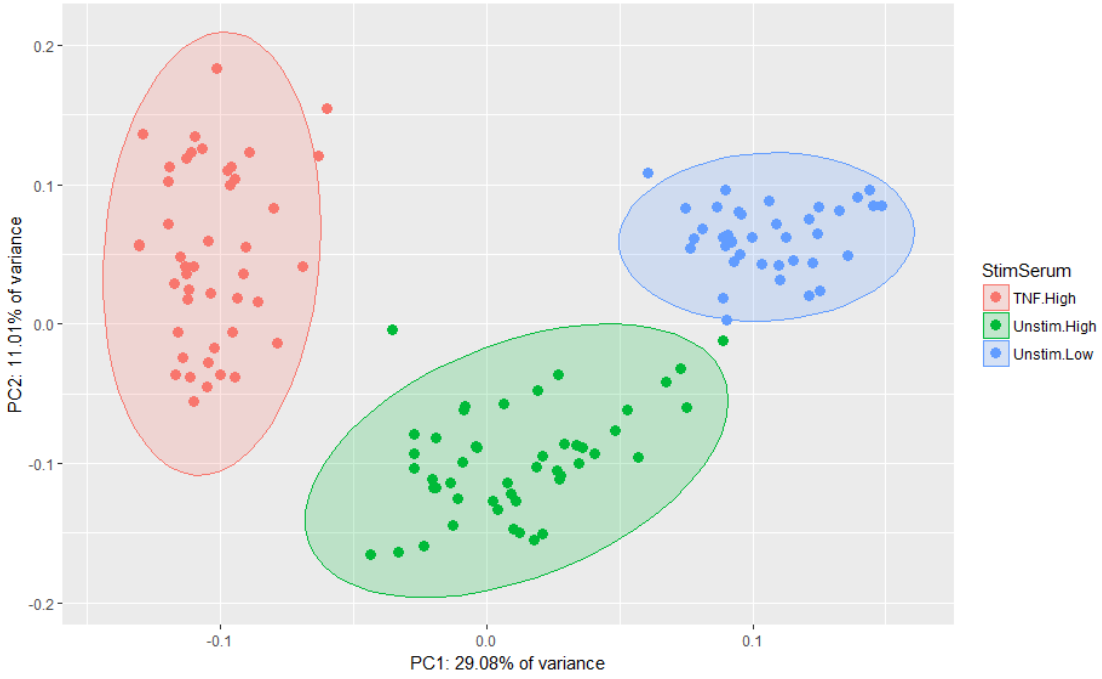


Figure 12: Principal component analysis of the prepared synovial fibroblast microarray dataset. Principal components one and two clearly separated the treatment groups with approximately 40% of variance in the dataset accounted for in these two components. Three clear clusters could be seen which accounted for the TNF stimulated samples (red), the samples grown in high serum (10% FBS) with no additional stimuli (green), and the samples grown in low serum (0.1% FBS) with no additional stimuli (blue). Ellipses are 95% confidence intervals for each cluster. TNF.High = TNF stimulated high-serum samples, Unstim.High = Unstimulated high-serum samples, Unstim.Low = Unstimulated low-serum samples.

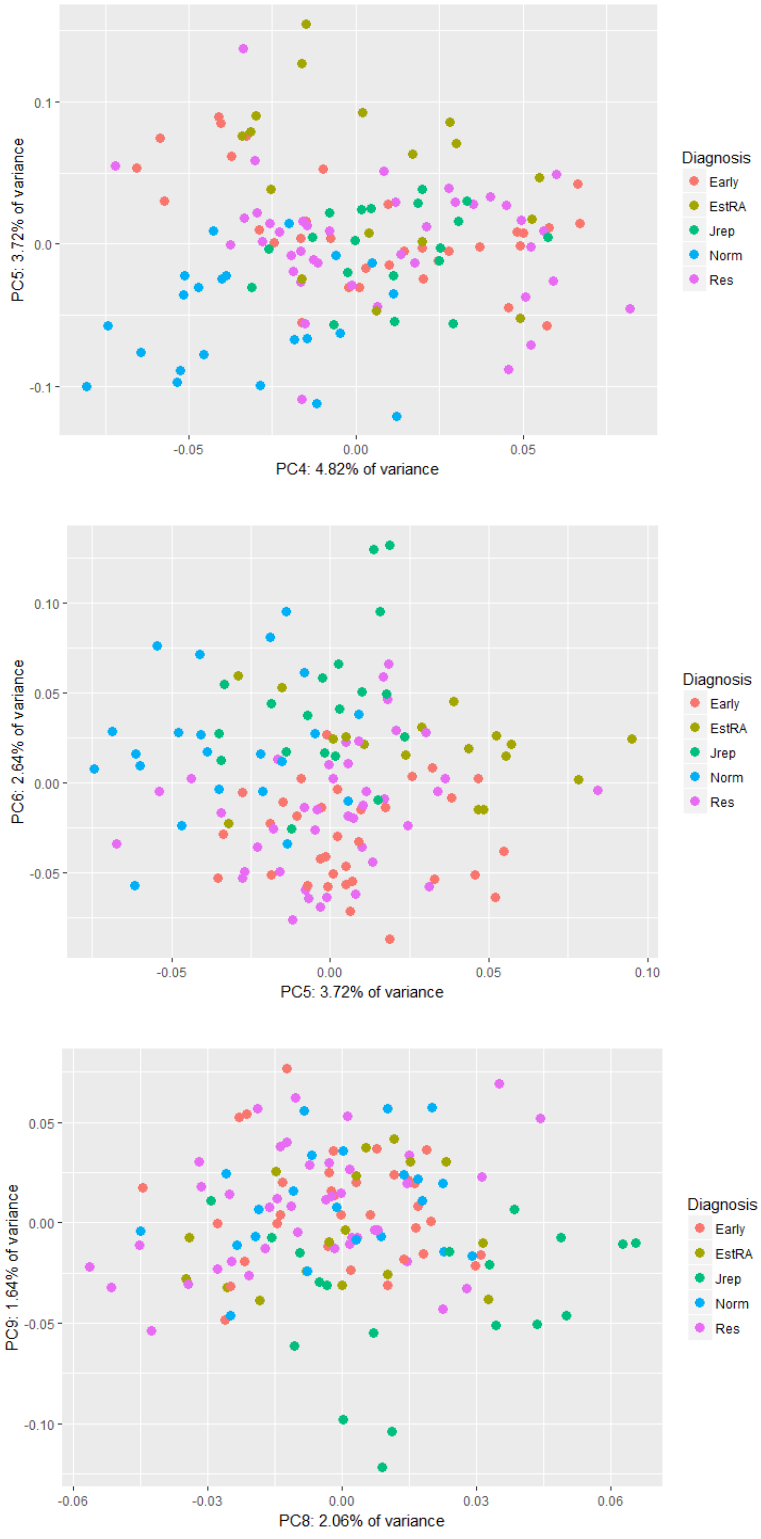


Figure 13: Principal component analysis of the prepared synovial fibroblast microarray dataset. Separation of the dataset on the basis of diagnosis was not clear, however the normal and Jrep samples appeared to separate on principal components four and five or eight and nine respectively. Principal components five and six appeared to slightly separate three broad clusters consisting of the Norm samples, the Early and Res samples, and the Jrep and EstRA samples respectively. Early = very early RA, EstRA = established RA, Jrep = joint replacement RA, Norm = normal controls, Res = resolving arthritis.

Investigating other relevant and potentially confounding biological variables found variation dependent on the sex of the patient from which the samples were generated and the joint from which the biopsy was collected. The sex of the patient from which the fibroblasts were isolated contributed to the variation in the dataset with subtle separation of samples visible in principal components seven, eight, and nine (Figure 14). However, the amount of variation accounted for by sex was less than that between the treatments, with approximately 6% of variation being contributed by this variable; this was similar to that due to the disease grouping of the samples. Site dependent differences between samples (principal component seven, 2.32% of variance) separated the ankle from the knee and metacarpophalangeal (MCP) samples (Figure 15). Some separation could be seen between the MCP and the other samples in other principal components however the low number of MCP samples made it difficult to ascertain if a separate cluster was truly present.

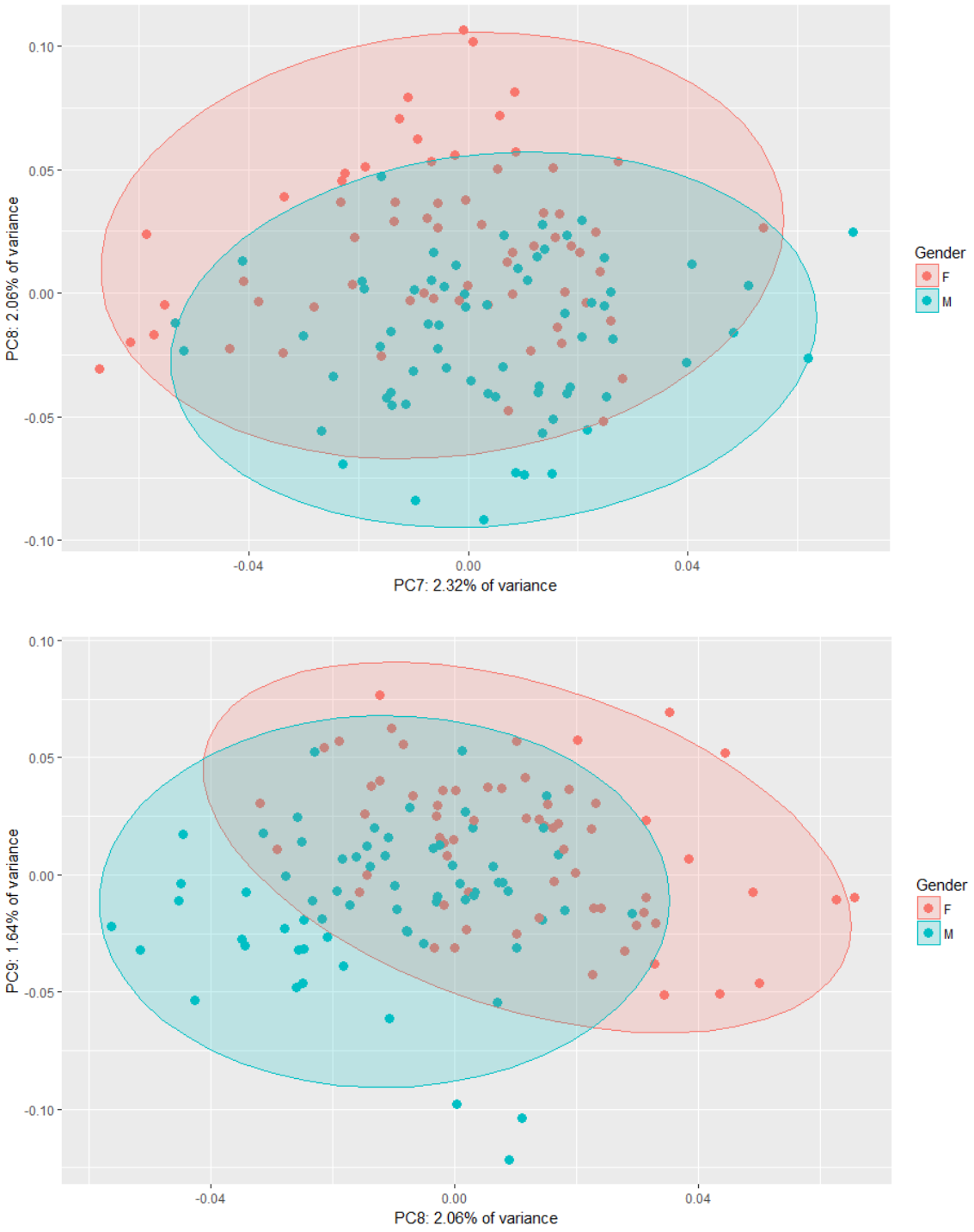


Figure 14: Principal component analysis of the prepared synovial fibroblast microarray dataset. Separation of the data on the basis of donor sex can be found in the principal components 7, 8, and 9 with approximately 6% of the variance in the dataset accounted for by these components.

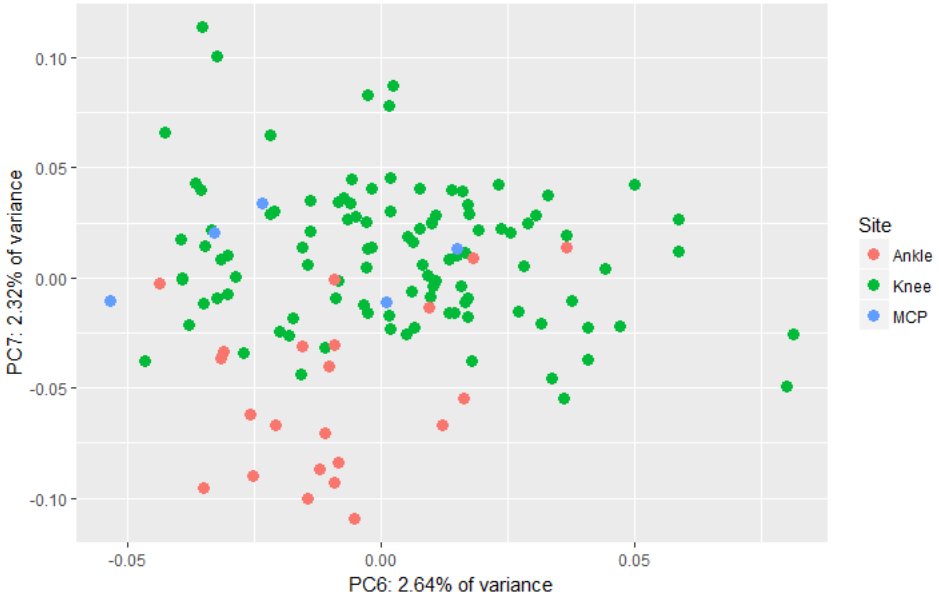


Figure 15: Principal component analysis of the prepared synovial fibroblast microarray dataset. Separation of the samples on the basis of site/joint from which the fibroblasts were isolated could be seen on the seventh principal component which accounted for 2.32% of the variance in the data.

Using this approach of analysing variance within each diagnostic outcome group across treatments did not provide any clearer clustering in the TNF treated samples however the normal samples still formed a subtly separated group of samples (Figure 16). Investigating the high serum unstimulated samples the normal, established RA, and joint replacement samples could each be seen forming subtly separated clusters in the second and third, fifth and sixth, sixth, and seventh principal components respectively mirroring the behaviour seen when running PCA on all of the samples (Figure 17). In the low serum samples separation between the disease groups could be seen in principal components 3 and 4 with the Jrep, Early and Res, and Norm and EstRA broadly forming separate clusters (Figure 18).

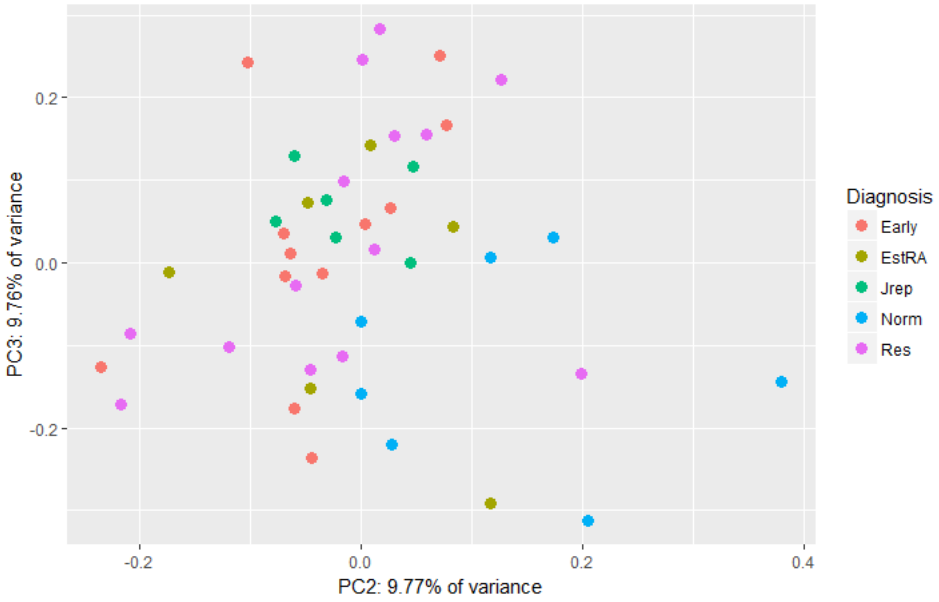


Figure 16: Principal component analysis of the TNF treated, high serum treated, synovial fibroblast microarray dataset. Although not a clear separation a tendency of the normal samples to cluster separately from the other samples could be seen in principal component 2.

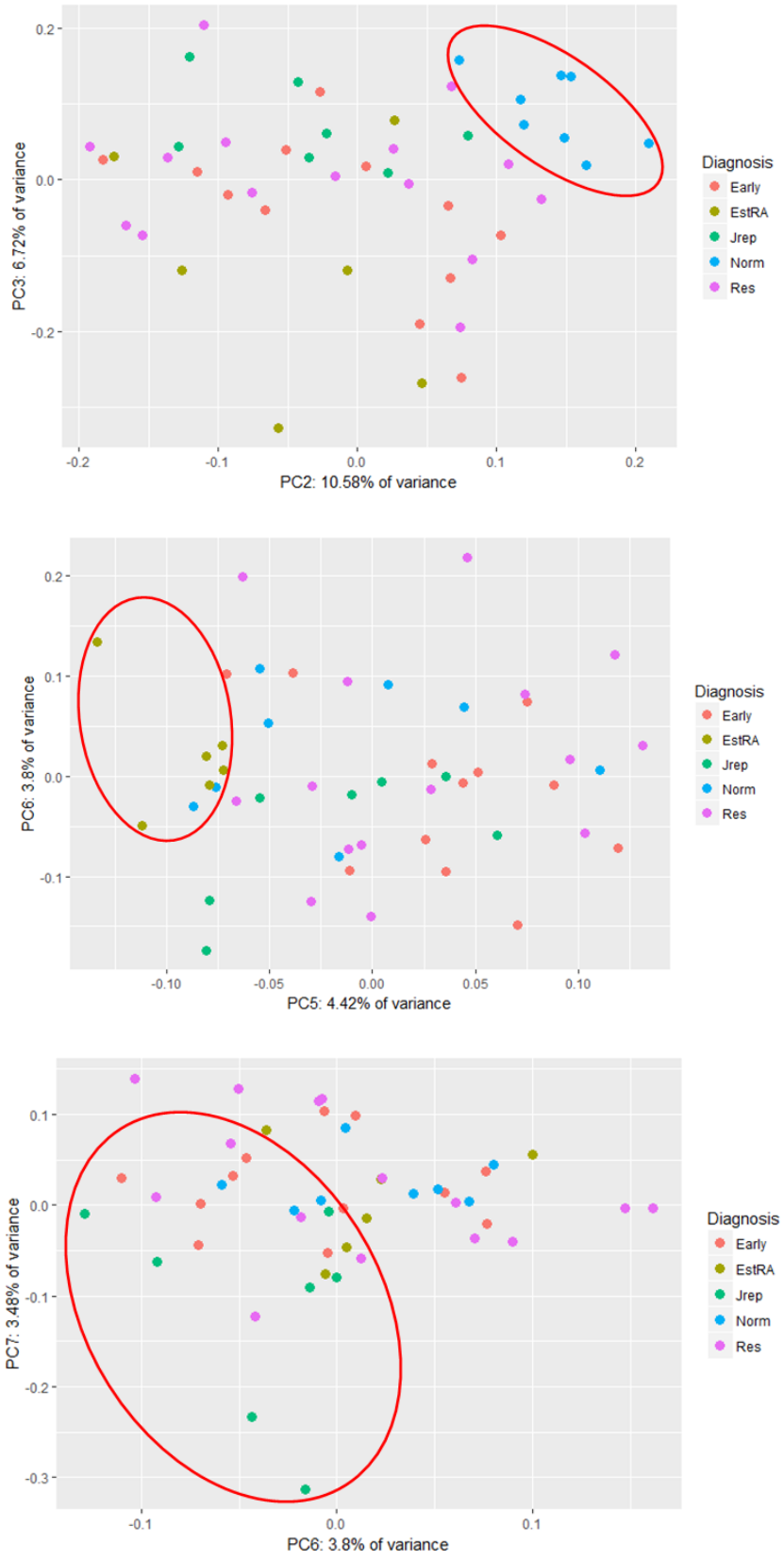


Figure 17: Principal component analysis of the high serum treated, unstimulated synovial fibroblast microarray dataset. The normal, established RA, and joint replacement samples could each be seen forming slightly separate clusters in the second and third, fifth and sixth, sixth, and seventh principal components respectively.

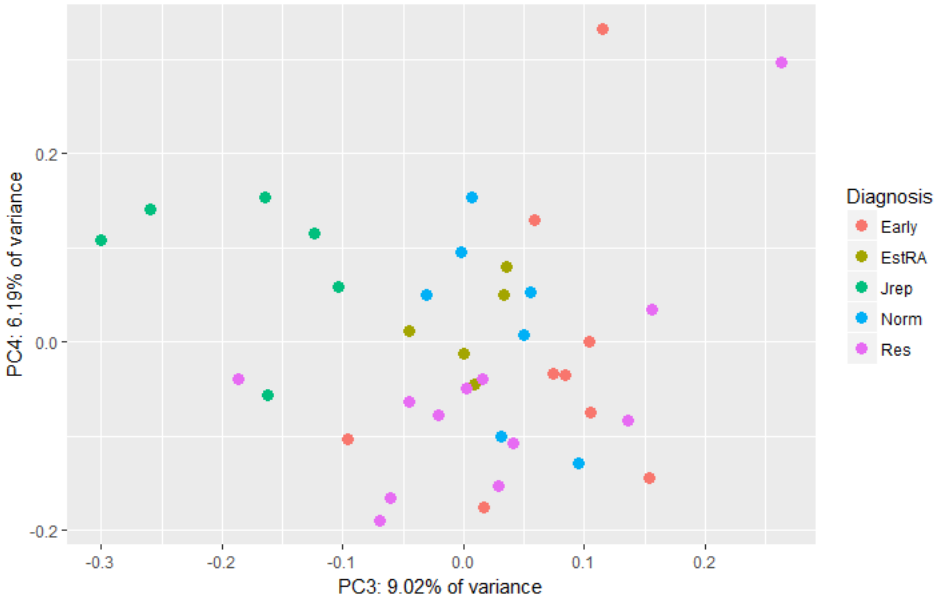


Figure 18: Principal component analysis of the low serum treated, unstimulated synovial fibroblast microarray dataset. Jrep samples, Early and Res samples, and EstRA and Norm samples broadly form separate clusters.

2.3.2 Differences between disease outcomes in each treatment

Due to the presence of unwanted (at least for the hypothesis being addressed) sex and site dependent variation within the dataset, a linear model was fitted to each probe using sex and site as the independent variables. The residuals were then taken from this model and a second model was fitted to the residuals using diagnosis and treatment as the independent variables. This strategy reduces the unwanted variation that may obfuscate the results of interest and was implemented after discussions and in collaboration with Kamil Slowikowski (Soumya Raychaudhuri lab, Brigham and Women's Hospital, Boston). The effect of this approach was apparent when looking at the principal components in which sex and site differences were visible in the original dataset. For example Figure 19 and Figure 20 show the clustering of TNF treated samples according to site and sex respectively both before and after the correction.

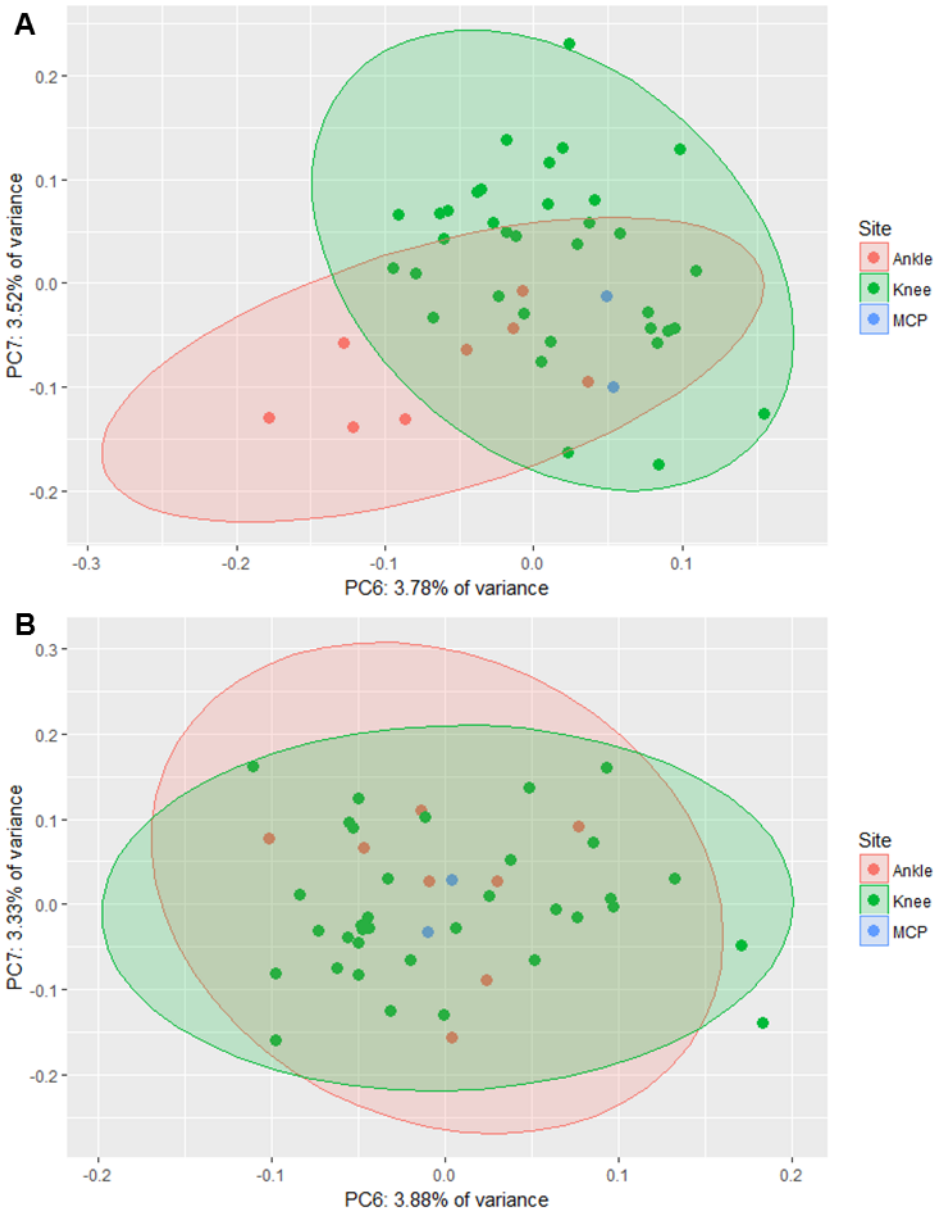


Figure 19: Principal component analysis of the TNF treated prepared synovial fibroblast microarray data in A, and of the dataset after fitting a model using sex and site as the independent variables and taking the residuals of the model for each probe in B. The effect of this approach was apparent as the separate clusters of knee and ankle samples present in principal components 6 and 7 in A were no longer visible in the 'corrected' data in B.

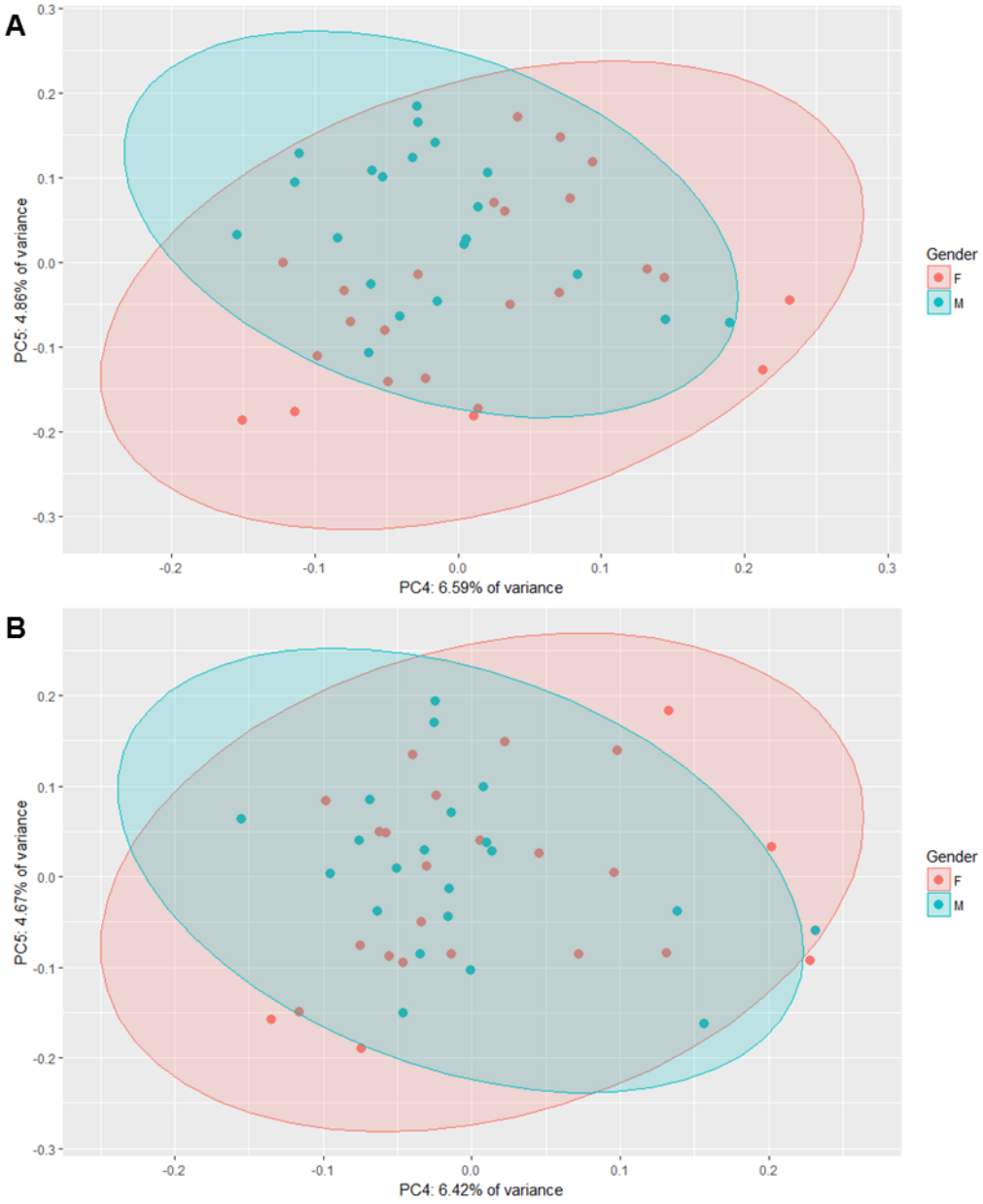


Figure 20: Principal component analysis of the TNF treated prepared synovial fibroblast microarray data in A, and of the dataset after fitting a model using sex and site as the independent variables and taking the residuals of the model for each probe in B. The effect of this approach was apparent as the separate clusters of female and male samples present in principal components 4 and 5 in A were no longer visible in the 'corrected' data in B.

Returning to the biological question of interest, how the fibroblasts from different disease stages differ in the different treatment groups, analyses of differentially expressed genes were made to address this question however a large number of differentially expressed genes were detected in many of the comparisons.

Whilst this indicated differences exist between the fibroblasts from different disease stages, the number of differentially expressed (DE) genes made interpretation complex. To expedite the analysis in each treatment group the DE genes from all comparisons were manually curated to select genes that may be of particular interest to RA and the immune system or genes that appear to be consistently expressed in one disease group compared to the others and only these selected genes were further investigated. Full tables of the selected differentially expressed genes are available in appendices 8.6, 8.7, and 8.8.

2.3.2.1 TNFα stimulated

Complete linkage hierarchical clustering of the TNF stimulated samples using Pearson's correlation distance calculated on the top 400 most variable genes between the samples did not provide clear separation of the samples depending on disease state (Figure 21). A cluster of resolving arthritis and very early RA samples could be found on the left-hand side of the dendrogram however the remaining samples appeared to be mixed throughout the plot.

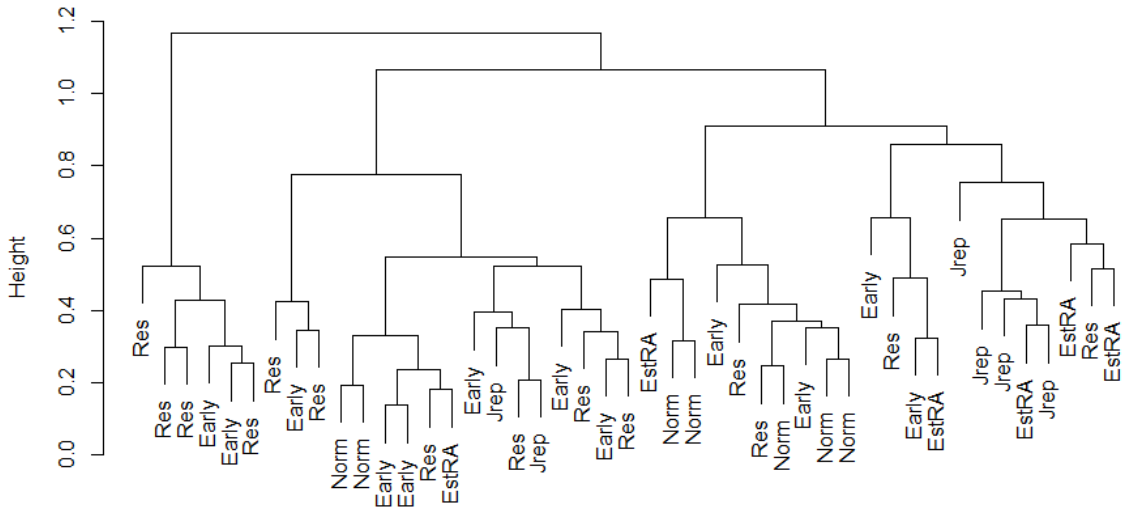


Figure 21: Dendrogram showing complete linkage hierarchical clustering of the TNF stimulated samples using Pearson's correlation distance calculated on the top 400 most variable genes between the samples after correction for sex and site influence. Clustering by diagnosis was unclear however possible clustering of resolving arthritis and very early RA samples was visible on the left-hand side of the plot.

Analysis of differentially expressed genes between the diagnosis groups in the TNF stimulated samples found a considerable number of DE genes in all comparisons with the exception of between very early RA and resolving arthritis samples in which no DE genes were detected (Table 6). The lack of DE genes between the Res and Early samples may reflect a lack of differences between these two disease outcomes in the fibroblasts during very early disease, a fact which is reflected by the clustering seen in PCA plots in Figure 13 and Figure 16.

	Jrep	EstRA	Early	Res
Norm	1013	1755	1951	1168
Res	58	992	0	
Early	262	2200		
EstRA	656			

Table 6: Number of differentially expressed genes between each disease group in the TNF stimulated samples

The lists of DE genes were manually curated as aforementioned and the resulting gene lists were investigated to find genes in each disease group that were consistently up- or down-regulated across the comparisons with other disease outcome groups and may represent an identifying phenotype. In the TNF stimulated Jrep samples this identified two upregulated and two downregulated genes that were expressed at higher or lower levels than all or most of the other disease groups. The upregulated genes were *HAS3*, encoding hyaluronan synthase 3, and *IL-6*, which was expressed at higher levels in the Jrep than EstRA, Early, and Norm samples. The genes downregulated in the Jrep samples were *FOS* and *FOSB*, downregulated when compared to all but Norm samples, which encode c-Fos and FosB respectively.

Using the same approach to investigate genes in the EstRA samples found five upregulated genes of interest and three downregulated that may represent a phenotype specific to the disease group. Of the upregulated genes *XIST* was detected as differentially expressed between EstRA and all other groups except the Jrep samples indicating that the approach taken to reduce the influence of sex and site on the dataset did not completely remove the variation contributed by these aspects (Figure 22). Investigating the number of male and female samples in each group clearly showed an imbalance particularly in the EstRA and Jrep samples (Table 7).

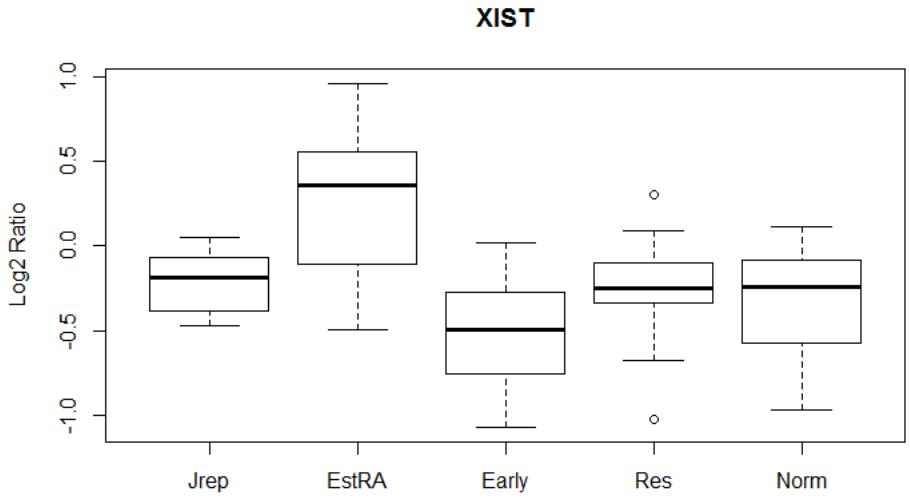


Figure 22: Expression of *XIST* in the TNF stimulated microarray samples after fitting a linear model to reduce the variation contributed by sex and site to the dataset. It is clear XIST expression remained higher in the EstRA samples than in the other groups.

	Male	Female
Jrep	2	4
EstRA	2	4
Early	5	7
Res	5	9
Norm	3	4

Table 7: Number of male and female samples in each disease group of the TNF stimulated microarray samples

The comparison of Norm samples with the other disease groups found several up- and down-regulated genes consistent across the comparisons. Both *CCL8* and *CCL20* were expressed at higher levels in Norm samples compared to all others, indicating differential regulation of leukocyte migration mediated by normal fibroblasts and fibroblasts from an inflammatory state, an unsurprising finding. This is further confirmed by the higher expression of *CXCL10* in Norm samples than in all other states except EstRA. The gene encoding cathepsin K, *CTSK*, was expressed at higher levels in the Norm samples than Res, EstRA, and Jrep samples and normal synovial fibroblasts also expressed higher levels of the gene encoding CD44, a receptor for hyaluronan that also may be involved

in process of activating and secreting TGF-β (Acharya et al., 2008; Aruffo et al., 1990). The expression of a reported fibroblast subset marker, *CD34*, was higher in Norm samples than Jrep or EstRA and the expression of another fibroblast subset marker, *PDPN*, was higher in Norm samples than all other groups except Early indicating that either the fibroblasts in each disease state had a homogenous population with different levels of subset markers between disease stages or that each disease group consisted of different proportions of fibroblast subsets (Mizoguchi et al., 2017). Of the genes downregulated in the Norm samples compared to other disease groups *HIST1H4C* and *HIST1H4L* were consistently lower expressed in the Norm samples.

Investigating the expression of known stromal markers showed differences in expression between disease stages such as the already mentioned higher *CD34* expression in the normal samples and also increased *PDPN* (GP38) expression in Norm and EstRA samples (albeit not detected as DE in EstRA samples) (Figure 23) (Choi et al., 2017; Mizoguchi et al., 2017). Expression of *THY1* (CD90) showed no clear modulation whereas *FAP* expression appeared highest in the Early samples.

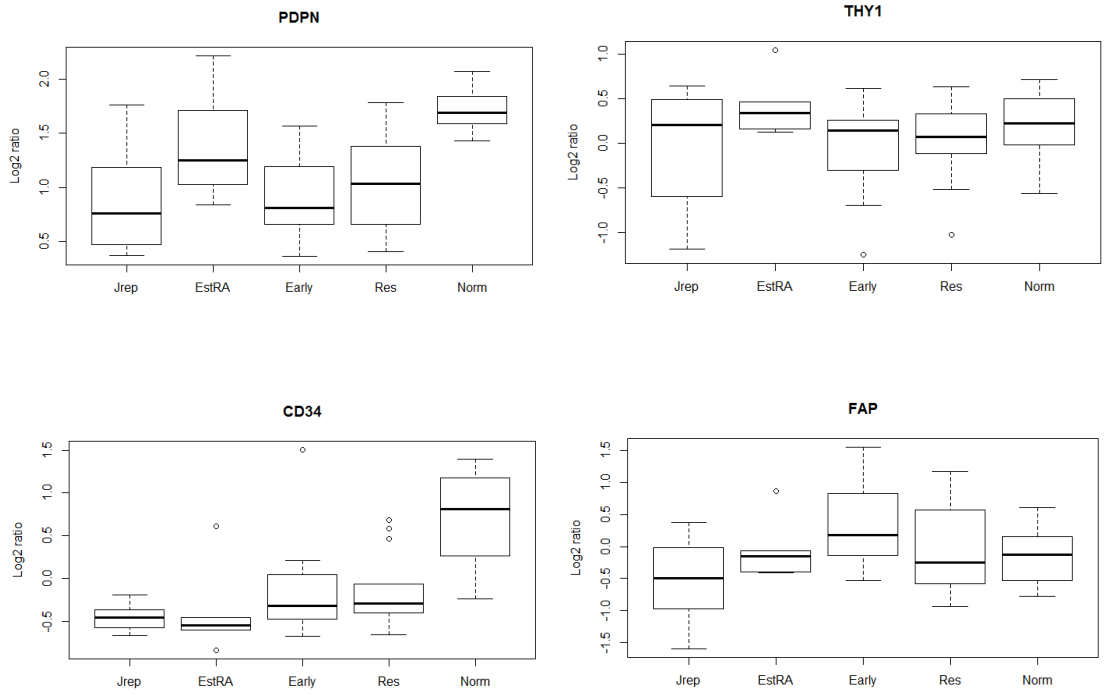


Figure 23: Expression of genes encoding markers of stromal populations in the TNF stimulated microarray data.

2.3.2.2 Unstimulated high serum.

As with the TNF stimulated samples complete linkage hierarchical clustering of the unstimulated high serum samples using Pearson's correlation distance calculated on the top 400 most variable genes between the samples did not find clear clusters according to diagnosis (Figure 24). However six of the eight Norm samples could be found in the cluster on the right hand side of the plot indicating some degree of clustering was occurring. The remaining samples were spread throughout the plot with no clear clustering apparent.

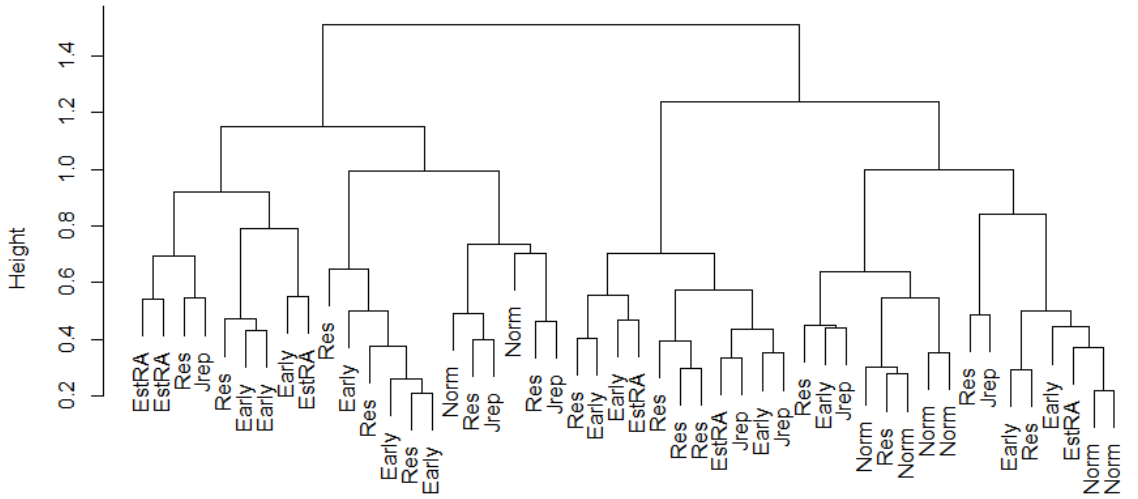


Figure 24: Complete linkage hierarchical clustering of the unstimulated high serum samples using Pearson's correlation distance calculated on the top 400 most variable genes between the samples after correction for sex and site influence. Clustering by diagnosis was unclear however a concentration of Norm samples could be seen in the right hand clusters.

When comparing different disease groups large numbers of DE genes were detected for all comparisons except Early versus Res, in which no DE genes were detected again, and Jrep versus Res, in which 46 genes were detected as DE (Table 8). These findings were similar to those in the TNF stimulated samples indicating that with or without the stimulus the disease groups maintained a degree of similarity.

	Jrep	EstRA	Early	Res
Norm	1063	3093	3466	2730
Res	46	2035	0	
Early	299	2348		
EstRA	1110			

Table 8: Number of differentially expressed genes between each disease group in the unstimulated high serum samples.

To investigate the large number of DE genes the same approach was taken to manually select relevant genes and then to investigate these genes for a conserved pattern of expression across comparisons that may aid in assigning a

phenotype or function to samples in each disease grouping. Jrep samples had significantly higher expression of *MOK*, encoding a protein kinase, than any other disease groups. *WNT16* was higher in Jrep samples compared to Res and Norm samples but not EstRA or Early samples. The increased expression of *HAS3* seen in the TNF stimulated samples was preserved in the unstimulated high serum samples with expression being higher in Jrep compared to Early, Res, or Norm. Considering genes downregulated in the Jrep samples, *FOS* and *FOSB* were both expressed at lower levels in the Jrep than the EstRA, Early and Norm samples or the EstRA, and Early samples respectively continuing the similarities with the TNF stimulated samples but with slight differences in which disease groups have lower expression. Jrep samples also had lower expression of the gene encoding hyaluronidase-1, *HYAL1*, than EstRA or Norm samples.

Investigating the Norm samples again highlighted similarities between this subset in the unstimulated high serum samples and TNF stimulated samples. The subset marker *CD34* was expressed at higher levels in the Norm samples compared to all other disease groups as was *CTSK*. *CTSH* which encodes cathepsin H was also higher in Norm compared to all other groups. Expression of the monocyte/macrophage chemoattractant *CXCL14* was also higher in Norm samples as was the expression of *MMP2* and *CCR3* (Augsten et al., 2009). *VEGFB* was expressed at higher levels in the Norm samples than all other groups perhaps indicating this gene is downregulated in response to inflammation. *FIGF*, also known as *VEGFD* was also higher expressed in the Norm samples which may indicate differences in vascularisation between normal and inflamed samples. Finally, Norm samples expressed higher levels of the IL1 receptor

encoding gene *IL1R1*, of *TLR5* which has been implicated in joint damage, and *CSF1* encoding the macrophage differentiation factor M-CSF (Fleetwood et al., 2007; Kassem et al., 2015).

Considering the genes downregulated in the Norm samples several keratin genes, *KRT14*, *KRT16*, and *KRT19* were expressed at lower levels in Norm compared to the other disease groupings along with *HAS2* and *SERPINE1* again hinting at differences in matrix constituents in between inflamed and non-inflamed tissues. *CDH13*, encoding cadherin-13, has been shown to inhibit proliferation and to be downregulated in tumours and is downregulated in the Norm samples (Sato et al., 1998; Takeuchi et al., 2002; Takeuchi et al., 2000). Expression of serine/threonine kinase 24 and integrin alpha-3 and *TGFB1* were all downregulated in the Norm samples also. Finally, as in the TNF stimulated samples several histone genes were downregulated in the Norm samples compared to the EstRA, Early, and Res samples. *HIST1H3H*, *HIST1H3G*, *HIST2H3D*, *HIST2H2BF*, and *HIST3H4C* all showed this downregulated pattern of expression.

Investigating the expression of fibroblast subset markers as with the TNF stimulated samples found similar patterns (Figure 25). The higher expression of *PDPN* (GP38) in the Norm samples seen in the TNF stimulated dataset was reduced in the unstimulated high serum samples however the higher level of *CD34* expression was maintained as was the increased expression of *FAP* in the Early samples.

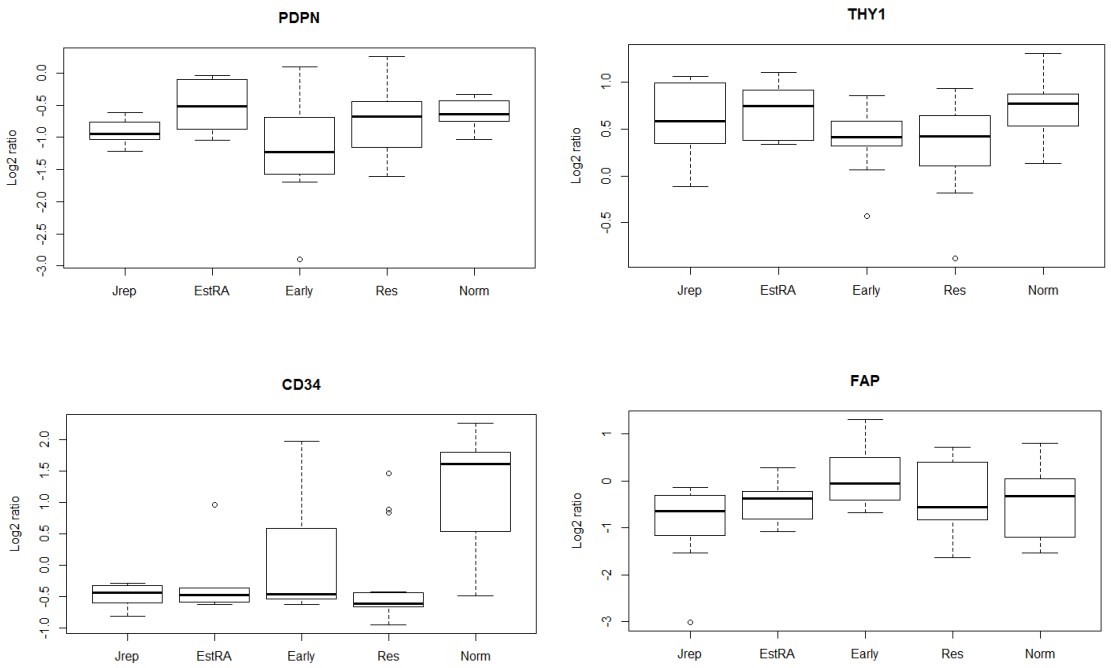


Figure 25: Expression of genes encoding markers of stromal populations in the unstimulated high serum microarray data.

2.3.2.3 Unstimulated low serum.

As with the previous samples hierarchical clustering did not find any clear clustering of clinical outcome groups in the unstimulated low serum dataset (Figure 26). EstRA samples in particular were spread throughout the dendrogram whereas some degree of clustering for Norm and Res samples could be seen in the middle region of the plot and some Jrep samples had clustered on the left-hand side of the plot.

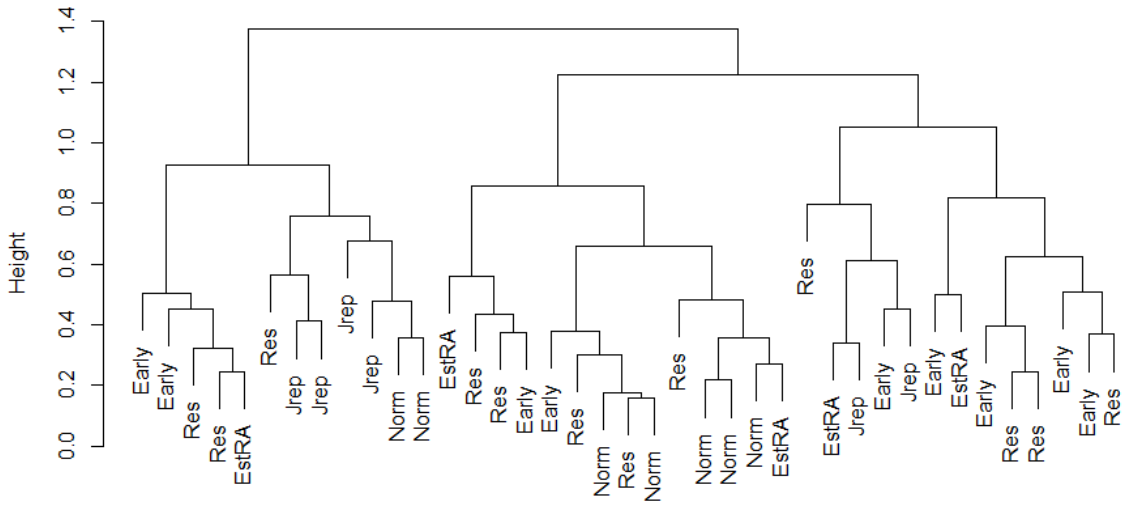


Figure 26: Complete linkage hierarchical clustering of the unstimulated low serum samples using Pearson's correlation distance calculated on the top 400 most variable genes between the samples after correction for sex and site influence. Clustering by diagnosis was unclear, some clustering of Norm and Res samples could be found towards the centre of the dendrogram with some Jrep clustering to the left hand-side but no clearly separated clusters.

The number of DE genes found in the comparisons of low serum samples was smaller than those found when investigating the TNF stimulated or unstimulated high serum samples (Table 9). In these samples comparisons with the EstRA group consistently found the lowest numbers of DE genes with only comparisons with EstRA versus Norm or Jrep v EstRA finding any DE genes (43 and two respectively). The comparison of Early and Res in the low serum dataset found only 1 DE gene following the pattern of no clear differences between these groups as in the previous datasets.

	Jrep	EstRA	Early	Res
Norm	868	2	369	30
Res	579	0	1	
Early	1134	0		
EstRA	43			

Table 9: Number of differentially expressed genes between each disease group in the unstimulated low serum samples.

Investigating the gene signature of Jrep samples found 14 genes that showed consistent regulation across the comparison of the different disease groups. *TAGLN*, encoding transgelin which is involved in cytoskeletal rearrangements, was higher in Jrep than in all other disease groups (Thompson et al., 2012). Another gene, *POGZ*, was upregulated in Jrep samples compared to the other disease groups and could be involved in the IL-1β signalling pathway (Ainscough et al., 2015). The final gene that was higher in Jrep samples than all others was *ADH4*, encoding alcohol dehydrogenase 4 pi subunit. There were several genes that were higher in Jrep samples than Early, Res, or Norm samples such as *ALDH3A1*; encoding aldehyde dehydrogenase 3 family member A1, *IL16*, encoding pro-interleukin-16, *AHRR*, encoding the aryl hydrocarbon receptor repressor, and *ZFAT*. No genes were found to be consistently downregulated in Jrep compared to all other samples. However *NFKBIZ*, *COL4A3BP*, *MAPK6*, *IRAK1*, *BCL2*, *DUSP14*, *CTSA* were all expressed at lower levels in the Jrep samples compared to the Early, Res and Norm samples. Expression of the subset marker *FAP* was lower in the Jrep samples when compared to Early samples. The genes involved in matrix regulation *MMP11* and *COL4A1* were both lower in Jrep samples than in Early or Res samples

Moving on to the Norm samples, *FIGF* was expressed at higher levels than in all other disease subsets and the related gene *VEGFB* was higher than in Jrep or Early samples. The gene *LGI1* was higher in the Norm samples than in EstRA, Early, or Res samples. As previously mentioned *CD34* was higher in this disease group than in the Early samples. With regards to genes downregulated in the Norm subset two patterns of expression were found, genes lower in Norm

samples compared to Jrep, Early, and Res samples, and genes lower in Norm samples compared to Jrep and Early samples. *TAGLN* and *CDH11* followed the first pattern of expression whereas *SERPINE2*, *ITGA10*, *SERPINI1*, and *ADAMTS4* showed the second pattern.

The expression of subset markers in the low serum samples highlighted the key differences found in the previous TNF stimulated and unstimulated high serum samples (Figure 27). Namely *CD34* expression was highest in Norm samples and *FAP* was increased in the Early samples.

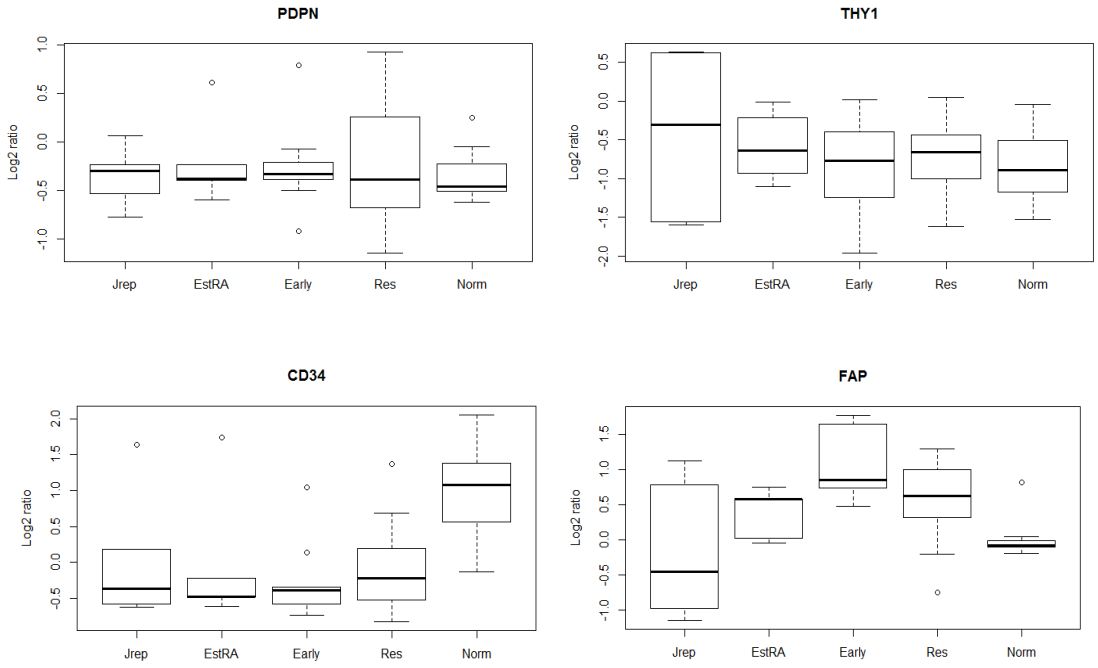


Figure 27: Expression of genes encoding markers of stromal populations in the unstimulated low serum microarray data.

2.4 Discussion

The first steps of analysis in big data approaches are normalisation/correction of the data followed by quality control and subsequent exploratory data analysis. Background correction and normalisation of the data acts to reduce unwanted variation, a very important step given the numerous sources of noise, and to prepare the data from each array to allow for comparisons between arrays (Coller, 2009). Many background correction and normalisation techniques are available for use on these kind of data each with advantages and disadvantages depending on the scenario. The normexp plus offset background correction was used here along with LOESS normalisation within arrays and G-quantile normalisation between arrays. The normexp plus offset has previously been shown to give a good consistent level of results with various datasets and a low false discovery rate (Ritchie et al., 2007). LOESS is a commonly used method to normalise microarray data removing technical variation (Cleveland, 1979; Smyth and Speed, 2003; Yang et al., 2002). The final normalisation step, G-quantile, is designed specifically for arrays in which a common reference was used on every array, which was the case here. The normalisation adjusts the distribution of the reference channels to match across all arrays and then adjusts the sample channel relative to the reference channel adjustment (see the documentation for the function normalizeBetweenArrays in the R Biocondutor package limma (Ritchie et al., 2015)). The normalisation successfully adjusted the distribution of the red and green fluorescence intensities and removed technical variation in the data as shown by a straightening in the curvature seen in some MA-plots. One additional benefit of the G-quantile normalisation method is that it could have helped to reduce batch effects introduced during the running of the arrays

although due to the lack of replicate samples in this dataset it was not possible to test how well this approach has worked.

Comparisons of differentially expressed genes between disease stages in all treatment groups detected a large number of genes leading to a targeted approach to the analysis in order to make the data more manageable. This approach had obvious caveats such as overlooking possibly important genes, overlooking new DE genes which may be of interest but have not previously been studied, and the reliance upon the knowledge of the operator to select genes of interest. However the advantage, particularly in time constrained situations, was that the differentially expressed gene lists were reduced to a more manageable and easily interpretable size and the genes being explored were easier to explore in the context of RA pathogenesis.

The comparison of most interest due to the possible benefits to early RA patients was between the very early RA and resolving arthritis group. As discussed in section 1.2 early treatment of RA is beneficial to patients and decreases joint damage over a substantial time period. However it is difficult to identify which patients are suffering from very early RA or from resolving arthritis and as such identification and treatment of RA patients during this early window is complex and improved biomarkers able to differentiate between resolving arthritis and persistent RA are still required. Unfortunately in the comparisons of these two disease stages in both the TNFα stimulated and unstimulated high serum samples no differentially expressed genes were detected and in the low serum samples only 1 differentially expressed gene, *RIIAD1*, was detected the

significance of which is unclear in the context of RA. The lack of differential gene expression between fibroblasts in the first three months of clinically evident arthritis is surprising given that functional differences have been found between fibroblasts from resolving arthritis and very early RA (Filer et al., 2017). In this publication fibroblasts from very early RA were found to recruit lymphocytes to stimulated endothelial cells whereas fibroblasts from resolving RA inhibited recruitment. This indicates functional assays may be required to understand the differences between cells from these outcome groups. Transcriptional analysis of freshly isolated fibroblasts may find more differences between these groups as long term culture is known to result in significant changes to gene expression (Zimmermann et al., 2000).

The Jrep sample signature defined as genes consistently found to be higher in most if not all of the comparisons with other samples included *HAS3*, encoding hyaluronan synthase 3, and *IL-6* which were upregulated and *FOS* and *FOSB* which were downregulated. Hyaluronan synthase 3 is responsible for the production of the synovial fluid constituent hyaluronic acid which is produced at higher levels in RA patients however doesn't lead to a difference in synovial fluid concentrations as the volume of synovial fluid also increases in RA patients (Ragan and Meyer, 1949). Additionally, the expression of hyaluronan synthase in endothelial cells and skin fibroblasts is known to be induced by proinflammatory mediators such as IL-1β and hyaluronan is known to interact with the immune system and also correlated with RA disease activity in early RA (Campo et al., 2006; De La Motte et al., 1999; Majeed et al., 2004; Mohamadzadeh et al., 1998). Increased expression of *IL-6* in the Jrep samples compared to the others is

unsurprising given the well-known involvement of IL-6 in the pathogenesis of RA (Srirangan and Choy, 2010). *FOS* and *FOSB* are members of the FOS gene family which are capable of dimerizing with members of the Jun family of proteins to form the transcription factor AP-1 (Chiu et al., 1988). c-FOS has been found to be expressed in RA synovial fibroblasts producing collagenases and is constitutively expressed in RA fibroblasts even after *in vitro* culture (Dooley et al., 1996; Trabandt et al., 1992).

In the normal samples *CTSK* and *CTSH* were at higher levels than in all other disease stages. The activity of cathepsin H has previously been shown to be unchanged in leukocytes from RA patients and the serum levels of cathepsin H are not correlated with disease activity or radiographic progression (Jorgensen et al., 2011; Sohar, Hammer and Sohar, 2002). However these findings do not exclude the possibility that levels and activity of cathepsin H in fibroblasts are affected in RA. *MMP2* and *CCR3* were both found at higher levels in the normal samples than all others, an interesting finding given that *MMP2* has been linked to erosive arthritis and *CCR3* is thought to be expressed at higher levels in RA synovial fibroblasts (Goldbach-Mansky et al., 2000; Liu et al., 2017; Xue et al., 2014b; Zhou et al., 2014). Interestingly, normal samples expressed higher levels of M-CSF which could be of importance given the role of this factor in differentiating anti-inflammatory macrophages (Fleetwood et al., 2007; Jaguin et al., 2013). Of the downregulated genes in the normal samples several can be linked to maintenance of matrix and synovial fluid. Three keratin genes, *KRT14*, *KRT16*, and *KRT19* were all lower as was *HAS2*. The gene *SERPINE1*, which encodes PAI-1, was also reduced and has been associated with both pro-fibrotic

responses and poor prognosis in cancers (Durand et al., 2004; Ghosh and Vaughan, 2012; Marudamuthu et al., 2015). *TGFB1* was lower in the normal samples which is in line with published data finding higher levels of TGF-β signalling in the synovium in RA (Gonzalo-Gil et al., 2013). Additionally the expression of the genes *HIST1H3H*, *HIST1H3G*, *HIST2H3D*, *HIST2H2BF*, and *HIST3H4C* which encode histone proteins was lower in the normal samples.

In the low serum samples an immediate difference was apparent simply by examining the number of differentially expressed genes in each comparison. In the other treatment groups the established RA samples demonstrated a decent number of differentially expressed genes in line with the magnitude of differences seen in comparisons of other disease groups. In the low serum samples either no or few differentially expressed genes were found in all comparisons with this disease stage. As all of the established RA samples in all treatment groups were generated in a separate experimental batch and also run in a separate batch of arrays with the low serum samples rather than the high serum samples an immediate concern is that the previously discussed signatures for this disease group in the stimulated and unstimulated high serum samples were actually an artefact of batch effects and not actually a true biological finding. Unfortunately no replicate samples were run between groups and so it was not possibly to easily test for this kind of noise in these datasets. The signatures present in all of the low serum samples were considerably different than those shown the stimulated and unstimulated high serum samples. The differences were unlikely to be due to different experimental batches as no comparisons between batches were made in the low serum samples. The differences in signature may be a biological

finding as the two treatment groups displaying similar signatures were grown in high serum with or without stimulation.

Of particular interest is that several subset markers were detected as differentially expressed between the disease stages. Across all treatment groups a pattern was seen in that *CD34* expression was highest in the normal samples and *FAP* highest in the very early RA samples. The cultured synovial fibroblasts used for the microarrays had not been examined for the expression of these markers so it is unclear if the protein expression in culture follows the pattern seen here. Several possibilities exist that may explain the differences in subset marker in the cultures. As the fibroblasts were isolated using the outgrowth method no selection of particular subsets was implemented meaning the fibroblasts isolated could be a heterogeneous population consisting of different proportions of fibroblasts from different subsets depending on the disease stage from which they were isolated which are maintained through culture and result in detection of different levels of expression of the genes encoding the markers between disease stages. Alternatively once isolated fibroblasts may become a homogenous population all expressing similar levels of subset marker genes that vary between disease stages.

The attempt to correct for variation contributed by sex using the linear modelling approach did not appear to completely eliminate confounding, as *XIST* was detected as differentially expressed in comparisons with the EstRA samples due to an imbalance in male and female samples. If the analysis was repeated techniques to correct for unwanted variation such as comBat or surrogate

variable analysis should be used and the results compared with the analysis in this thesis (Johnson, Li and Rabinovic, 2007; Leek and Storey, 2007). The risk of using these techniques is that some variation introduced by the biological differences of interest could also be removed. However, given the large number of differentially expressed genes detected this may not be a large issue with this dataset. As several samples were possibly mislabelled leading to the mismatch between sex and *XIST* expression it could also be possible that other samples may have been mislabelled but swapped with samples from the same sex donor meaning it is not possible to identify these samples.

The analysis of this dataset improves on the previous analysis due to rigorous quality control steps and exploration of confounding sources of variation (Juárez Pérez, 2014). The major sources of unwanted variation were sex and site, unsurprising given the known transcriptional differences between fibroblasts isolated from different sites (Chang et al., 2002; Frank-Bertoncelj et al., 2017; Rinn et al., 2006). The data appeared robust with the differentially expressed genes making intuitive sense given the origin of the fibroblasts and in the context of published work. If the work was to be completed again there is a convincing argument to use only samples from one sex and site thus reducing variation from unwanted sources without the need for computational correction. In addition, analysis of freshly sorted and treated fibroblasts would avoid the influence of prolonged *in vitro* culture upon gene expression and possibly uncover findings masked by this effect.

In summary the results of this chapter find that gene expression varied

significantly between different stages of RA but that resolving and very early RA are not significantly different. Gene signatures could be found in most disease stages in each treatment group that identified the stage of disease from which the fibroblasts were isolated. The expression of genes encoding previously identified subset markers varied between disease stages hinting that *in vitro* cultured fibroblasts may not be a homogeneous population. The expression of *FAP* could be a possible biomarker of very early RA allowing stratification of patients with early synovitis.

The work in this chapter supported the first alternative hypothesis stated in this body of work; fibroblasts isolated from different stages of RA progression maintain differential gene expression during *in vitro* co-culture and respond to stimulation in different ways. The finding of differential gene expression between late stage RA and early stages of RA progression changes the viewpoint with which one should approach the literature. Most studies in the field use synovial fibroblasts isolated from patients undergoing joint replacements and therefore use a highly selected for group of patients with destructive RA that is likely to have encountered many treatment regimes. Differential gene expression between fibroblasts from this late stage RA and early stages indicates that the 'imprinted aggressor' phenotype commonly referred to in the literature may be an artefact specific to destructive end stage RA (Bottini and Firestein, 2013). Additionally the lack of differential gene expression between fibroblasts from patients with resolving arthritis and patients with very early RA indicates that the imprinted changes that characterise RA fibroblasts do not occur during the first three months of apparent symptoms. These data could signpost that changes in

synovial fibroblasts are not the first features to occur in RA and instead occur subsequent to inappropriate activation and targeting of the immune system. This could explain why early therapeutic treatment has been shown to be alter overall disease progression and prevent joint damage more effectively than later treatment (Kyburz et al., 2011; Moura et al., 2015; Nell et al., 2004; Van Aken et al., 2004).

The differences in gene expression between fibroblasts from different stages of RA adds to the knowledge in this field. As differences were detected it could be that these differences extend to fibroblast-monocyte crosstalk resulting in more pro-inflammatory monocyte phenotypes in late stage RA than early and as such we used *in vitro* assays to test this hypothesis.

3 INVESTIGATING FIBROBLAST/MACROPHAGE INTERACTIONS: *IN VITRO* CO-CULTURES

3.1 Introduction

Given the sparsity of knowledge regarding the interaction between synovial fibroblasts and monocytes or macrophages, work was undertaken to address this using several different methods. Firstly, the technical aspects of co-culturing two cell types together including seeding density, removal and deconvolution, and optimal methods of monocyte isolation from peripheral blood were investigated. The effects of monocyte:fibroblast co-culture on cytokine release was interrogated to test the hypothesis that fibroblasts isolated from different BEACON outcome groups can differentially modulate monocyte phenotypes. Monocyte to macrophage differentiation was elicited with either M-CSF or GM-CSF during co-culture with BEACON fibroblast lines to assess if fibroblasts are capable of differentially modulating monocyte to macrophage differentiation and drive a more pro- or anti-inflammatory phenotype. Finally, M-CSF differentiated macrophages were co-cultured with fibroblasts in the presence of TNFα to test the ability of BEACON fibroblasts to differentially modulate a set of interferon stimulated genes previously shown to be regulated by co-culture with synovial fibroblasts (Donlin et al., 2014).

This chapter aimed to address the gap in knowledge regarding the influence of fibroblasts upon monocyte phenotype during RA progression using an *in vitro* system. As the previous chapter demonstrated fibroblasts from different stages of RA maintain differential gene expression it was hypothesised that this would translate into differences in protein expression that results in late stage RA fibroblasts driving a more pro-inflammatory phenotype in monocytes and

macrophages than fibroblasts from earlier stages. To test this hypothesis three questions were asked:

1. Do fibroblasts isolated from different stages of RA progression differentially modulate cytokine/chemokine release from peripheral blood monocytes?
2. Do fibroblasts isolated from different stages of RA progression differentially modulate monocyte-to-macrophage differentiation?
3. Do fibroblasts isolated from different stages of RA progression differentially modulate macrophage response to stimulation?

3.2 Materials and methods

3.2.1 Fibroblast and RAW264.7 co-cultures

Co-cultures of human synovial fibroblast lines and the murine monocyte/macrophage-like cell line RAW264.7 were investigated as an easy method for identifying the source of cytokines released during the co-culture by using species specific assays with the caveat that the species difference may also prevent or modify the interactions between the cell types (Raschke et al., 1978). 45,000 fibroblasts were seeded into wells of a 24 well plate and rested for 2 days before removal of the media and seeding either 50,000 or 600,000 RAW264.7 cells on top of the fibroblasts and culturing for various timepoints. Co-cultures were completed in standard complete fibroblast medium as specified in 2.2.1.

3.2.2 Monocyte isolation

For all experiments except the stimulated co-cultures in section 3.3.6 peripheral blood was obtained from apheresis cones provided as a commercial product by

the NHSBT. This facilitated the easy isolation of large numbers of monocytes for each experiment.

3.2.2.1 Method 1: Density gradient separation followed by subsequent negative selection using magnetic microbeads

Monocytes were isolated from whole blood via density gradient separation using Ficoll-Paque PLUS (GE Healthcare) followed by negative selection using a pan-human monocyte magnetic bead labelling kit (MiltenyiBiotec). The blood was emptied from a blood cone and any remaining blood flushed out with Hank's balanced Salt solution (HBSS, No Ca or Mg, Life Technologies). The total volume of the blood was made up to 40ml with HBSS and gently mixed before being layered onto Ficoll-Paque PLUS at a 1:1 ratio and centrifuged. The isolated peripheral blood mononuclear cells (PBMCs) were then labelled with the pan-human monocyte magnetic bead labelling kit and monocytes isolated by sequentially passing the labelled cell suspension through two MACs LS columns (MiltenyiBiotec). The purity and subset bias of isolated monocytes was determined by flow cytometric analysis of CD14(-PE) and CD16(-FITC) (BD Biosciences, clones MφP9 and 3G8) cell-surface expression using a Dako Cyan-ADP.

3.2.2.2 Method 2: Double density gradient separation followed by adherence to tissue culture plastic

The second method of monocyte isolation consisted of density gradient separation using Ficoll-Paque PLUS with a subsequent density gradient separation using Percoll PLUS (GE Healthcare) followed by seeding of the obtained cell suspension onto tissue culture plastic and washing off the non-adherent cells with PBS. The blood was collected from the blood cone and the

PBMCs isolated using Ficoll-Paque PLUS as in method 1. The washed PBMCs were diluted using RPMI 1640 containing 10% FBS to a concentration of $20x10^6$ cells/ml. A Percoll solution was made using 23ml Percoll PLUS, 4.75ml 10x Dulbecco's PBS (Gibco), and 26.25ml of distilled H_2O, and the solution was dispensed into 15ml Falcon tubes (5ml/tube). The cell suspension was then gently layered onto the Percoll solution (5ml/tube) and tubes were centrifuged at 400g for 30 minutes with minimum break and acceleration. The interphase was removed from the tubes and washed in RPMI 1640 containing 10% FBS and centrifuged at 650g for 5 minutes. The cells were resuspended in RPMI 1640 containing 10% FBS and seeded into 10cm diameter tissue-culture treated petri dishes at a density of $20x10^6$ cells/dish. The cells were allowed to adhere for 1h at 37^{o}C after which contaminating non-adherent cells were removed by washing twice with PBS with the remaining population being monocytes. The purity and subset bias of isolated monocytes was determined by flow cytometric analysis as in method 1.

3.2.3 Unstimulated monocyte:fibroblast co-cultures

Fibroblasts and heterologous peripheral blood monocytes were co-cultured together for 5 days, 24 hours, or less as follows. $0.2x10^5$ fibroblasts/well were seeded into a 24 well plate (Corning) and cultured for 2 days at 37^{o}C 5% CO_2 in complete fibroblast medium. Monocytes were isolated as in 3.2.2.1 and seeded onto the fibroblasts, after the fibroblasts had been washed once in PBS, at a density of $3x10^5$ monocytes/well in co-culture medium which is constructed as complete fibroblast medium with the adjustment of the foetal bovine serum concentration to 5% v/v. Fibroblast only and monocyte only control cultures were

subjected to the same protocol but omitting the relevant cell type at its addition and adding medium only. Supernatants were harvested at the end of co-culture and immediately stored at -80ºC. Clinical details of the fibroblast lines used are presented in Table 14.

3.2.4 Unstimulated macrophage:fibroblast co-cultures

Unstimulated macrophage:fibroblast co-cultures were completed in the presence of GM-CSF and M-CSF (PeproTech) with no additional stimuli. For these co-cultures monocytes were isolated as in 3.2.2.1 and were seeded at a density of $3x10^5$ monocytes/filter onto inverted Transwell filters for 24 well plates with 0.4µm pores in either co-culture medium containing 50ng/ml GM-CSF, 100ng/ml M-CSF, or neither and incubated at 37ºC 5% CO_2 for 1 hour. After 1 hour the filters were righted and placed into a 24 well plate and 200µl or 500µl of co-culture medium containing GM-CSF, M-CSF, or neither factor at the aforementioned concentrations was added above or below the filter respectively. The monocytes were left to differentiate for 2 days before gently washing once with PBS and adding $0.2x10^5$ fibroblasts into the upper chamber of the filters in 200µl of co-culture medium containing GM-CSF, M-CSF, or neither at the aforementioned concentrations and adding 600µl of the co-culture medium with the growth factors below the filter. After two days of co-culture all cells were washed once in PBS and removed by incubating with [2x] trypsin for 5 minutes at 37ºC 5% CO_2 before inhibiting the trypsin with co-culture medium with no added GM-CSF or M-CSF and centrifuging the cells at 300g for 6 minutes. Clinical details of the fibroblast lines used are presented in Table 15.

3.2.5 Removal and deconvolution of co-culture populations

In order to isolate RNA or assess cells by flow cytometry after direct co-culture methods to extract pure populations from the co-culture were compared. After five days in direct co-culture cells were removed from tissue culture plastic using three different treatments. The co-cultures were gently washed once with PBS and 500µl of either trypsin (used at 2x concentration), Accutase (Sigma), or cell-dissociation buffer (CDB) (Gibco) was added to the wells and plates were incubated for 15 minutes at 37ºC. 500µl of co-culture medium was then added to all wells. The co-cultures treated with CDB were scraped and all treatments were further disrupted by pipetting. The ability of the removal methods to isolate two clean populations of fibroblasts and monocytes with minimal disruption of surface markers was assessed via flow cytometry panels as in Table 10 and Table 11.

Target	Conjugate	Clone	Isotype	Supplier
CD16	FITC	3G8	Mouse IgG1, κ	BD Biosciences
CD14	PE	MφP9	Mouse IgG2b, κ	BD Biosciences
CD163	PE-CF694	GHI/61	Mouse IgG1, κ	BD Biosciences
CD68	Alexor Fluor 647	Y1/82A	Mouse IgG2b, κ	BD Biosciences
CD86	BV510	FUN-1	Mouse IgG1, κ	BD Biosciences
CD45	V450	2D1	Mouse IgG1, κ	BD Biosciences
CD90	PerCP-Cy5.5	eBio5E10	Mouse IgG1, κ	eBioscience
FAP	Biotin		Sheep IgG	R & D Systems
Biotin	Streptavidin PE-Cy7	N/A	N/A	Invitrogen

Table 10: Antibody panel used to assess the efficacy of co-culture separation and the effect of removal techniques on cell-surface markers

Isotype	Clone	Conjugate	Supplier
Mouse IgG1, κ	MOPC-21	FITC	BD Biosciences
Mouse IgG2b, κ	27-35	PE	BD Biosciences
Mouse IgG1, κ	X40	PE-CF594	BD Biosciences
Mouse IgG2b, κ	27-35	Alexa Fluor 647	BD Biosciences
Mouse IgG1, κ	X40	BV510	BD Biosciences
Mouse IgG1, κ	MOPC-21	V450	BD Biosciences
Mouse IgG1, κ	P3.6.2.8.1	PerCP-Cy5.5	eBioscience
Sheep IgG		Biotin	Southern Biotech
N/A	N/A	Streptavidin PE-Cy7	Invitrogen

Table 11: Isotype controls used in conjunction with the antibody panel in Table 10

3.2.6 Intracellular cytokine staining

Direct co-cultures were established in which the fibroblasts were labelled with 1μM of DDAO-SE (Life Technologies) prior to seeding. Cells were harvested from these 5 day co-cultures using Accutase (Sigma) then stained for CD45. Some co-culture wells were treated with 2μg/ml Brefeldin A (Sigma) for 12 hours before staining for intracellular TNF-α and IL-10 using a fixation and permeabilization kit (BD Biosciences). All samples were analysed on a MoFlowAstrios EQ (antibody panel used in Table 12).

Target	Conjugate	Clone	Isotype	Supplier
CD45	V450	2D1	Mouse IgG1, κ	BD Biosciences
TNF-α	Alexa Fluor® 700	MAb11	Mouse IgG1, κ	BD Biosciences
IL-10	DyLight 350	JES3-9D7	Rat IgG1	Novus Biologicals

Table 12: Antibodies used for analysis of intracellular cytokines

Isotype	Clone	Conjugate	Supplier
Mouse IgG1, κ	MOPC-21	V450	BD Biosciences
Mouse IgG1, κ	MOPC-21	Alexa Fluor® 700	BD Biosciences
Rat IgG1, kappa L Chain	KLH/G1-2-2	DyLight 350	Novus Biologicals

Table 13: Isotype control antibodies used in conjunction with those in Table 12.

3.2.7 Blocking CD18 interactions

An anti-CD18 blocking antibody (Biolegend clone TS1/18) was added to direct 5 day co-cultures to investigate the role of common membrane-bound protein interactions between fibroblasts and monocytes. The direct co-cultures were completed as in 3.2.3 with the addition of a single treatment with 5µg/ml at the beginning of the co-culture when the monocytes are combined with the fibroblasts. An isotype control (Biolegend clone MOPC-21) was used as appropriate to control for non-specific effects of the anti-CD18 antibody.

3.2.8 Stimulated macrophage:fibroblast co-cultures

TNFα stimulated macrophage:fibroblast co-cultures were completed following a similar protocol as used by Donlin et al. (2014). $0.2x10^5$ fibroblasts were seeded into the upper chamber of a Transwell filter with 0.4µm pores in a 24 well plate with 200µl and 600µl of complete fibroblast medium above and below the filter respectively and were incubated at 37ºC 5% CO_2 for two days. Monocytes were isolated from peripheral blood as in 3.2.2.1 and were seeded at a density of $0.5x10^5$ monocytes/well in a 24 well plate in complete fibroblast medium supplemented with 10ng/ml M-CSF and were cultured at 37ºC 5% CO_2 for 2 days to induce monocyte to macrophage differentiation. After the 2 days of culture both cell types were washed twice in PBS and the Transwell filters containing the fibroblast cultures were placed into the 24 well plate above the macrophage cultures before adding 200µl or 600µl of complete fibroblast medium supplemented with 20ng/ml TNFα above and below the filter respectively. The cells were then co-cultured at 37ºC 5% CO_2 for 16 hours. Both cell types were then separated and washed once with PBS, the fibroblasts were removed by

treatment with 600µl 2x Trypsin for 15 minutes before inhibiting the enzyme with an equal volume of complete fibroblast medium. The macrophages were removed by adding fresh PBS to the wells, putting the entire plate on ice and scraping the cells with cell scrapers. Clinical details of the fibroblast lines used to generate samples in preparation for RNA-sequencing are presented in Table 16.

ID	Group	Disease classification	Site	CCP +/-	RF +/-	Disease duration (Weeks)	Age	Sex	DMARD	Prednisolone	DAS28 ESR baseline	ESR (mm/ hr)	CRP (mg/l)
BX013	VeRA	RA	Knee	-	-	10	45	F	None	None	3.8	24	12
BX033	Res	Pseudogout	Ankle	-	-	7	81	F	None	None	6.7	60	52
BX049	VeRA	RA	Ankle	+	+	3	48	F	None	None	3.9	10	0
BX071	Res	Parvovirus	Ankle	-	-	4	41	F	None	None	1.7	5	9
BX076	Res	Reactive	Knee	-	-	6	28	M	None	None	4.5	18	8
BX084	VeRA	RA	MCP	+	+	6	49	M	None	None	6.8	25	18
BX087	Res	Reactive	Ankle	-	-	4	27	M	None	None	3.8	37	28
BX092	VeRA	RA	MCP	+	+	4	48	M	None	None	6.0	63	38
BX119	VeRA	RA	MCP	-	-	6	61	M	None	None	5.3	20	25
RA11SY	Jrep	RA	Knee	NA	+	1040	62	F	None	None	6.1	63	34
RA18SY	Jrep	RA	Knee	NA	+	1196	47	F	Etanercept	None	3.8	57	NA
RA20SY	Jrep	RA	Knee	NA	+	1040	72	M	Leflunomide	None	4.8	18	3
RA28SY	Jrep	RA	Knee	NA	-	1040	41	F	Gold	10mg	3.3	20	26
RA29SY	Jrep	RA	Knee	+	+	364	67	F	Etanercept, MTX	7.5mg	6.6	57	66

Table 14: Clinical characteristics of the patients from which fibroblast lines were isolated and used for the unstimulated monocyte:fibroblast co-cultures. NA = Not available, MTX = Methotrexate, DAS28 = Disease activity 28 score, ESR = Erythrocyte sedimentation rate, CRP = C-reactive protein, CCP = Anti-cyclic citrullinated peptide, RF = Rheumatoid factor, MCP = Metacarpophalangeal, UA = Undifferentiated arthritis.

ID	Group	Disease classification	Site	CCP +/-	RF +/-	Disease duration (Weeks)	Age	Sex	DMARD	Prednisolone	DAS28 ESR baseline	ESR (mm/ hr)	CRP (mg/l)
BX005	VeRA	RA	Knee	-	-	5	70	F	None	None	6.0	68	26
BX013	VeRA	RA	Knee	-	-	10	45	F	None	None	3.8	24	12
BX033	Res	Pseudogout	Ankle	-	-	7	81	F	None	None	6.7	60	52
BX063	VeRA	RA	Knee	+	-	9	74	F	None	None	4.4	20	32
BX071	Res	Parvovirus	Ankle	-	-	4	41	F	None	None	1.7	5	9
BX072	Res	UA	Knee	-	-	10	32	M	None	None	3.6	15	0
BX076	Res	Reactive	Knee	-	-	6	28	M	None	None	4.5	18	8
BX084	VeRA	RA	MCP	+	+	6	49	M	None	None	6.8	25	18
BX092	VeRA	RA	MCP	+	+	4	48	M	None	None	6.0	63	38
RA11SY	Jrep	RA	Knee	NA	+	1040	62	F	None	None	6.1	63	34
RA25SY	Jrep	RA	Knee	NA	+	1560	53	f	HCQ	10mg	3.6	na	na
RA28SY	Jrep	RA	Knee	NA	-	1040	41	F	Gold	10mg	3.3	20	26
RA29SY	Jrep	RA	Knee	+	+	364	67	F	Etanercept, MTX	7.5mg	6.6	57	66

Table 15: Clinical characteristics of the patients from which fibroblast lines were isolated and used for the unstimulated macrophage:fibroblast co-cultures. NA = Not available, MTX = Methotrexate, DAS28 = Disease activity 28 score, ESR = Erythrocyte sedimentation rate, CRP = C-reactive protein, CCP = Anti-cyclic citrullinated peptide, RF = Rheumatoid factor, MCP = Metacarpophalangeal, UA = Undifferentiated arthritis, HCQ = Hydroxychloroquine.

ID	Group	Disease classification	Site	CCP +/-	RF +/-	Disease duration (Weeks)	Age	Sex	DMARD	Prednisolone	DAS28 ESR baseline	ESR (mm/ hr)	CRP (mg/l)
BX005	VeRA	RA	Knee	-	-	5	70	F	None	None	6.0	68	26
BX010	Res	Parvovirus	Knee	-	-	4	40	F	None	None	3.9	5	0
BX011	VeRA	RA	Knee	-	-	2	49	F	None	None	4.7	12	8
BX017	RA	RA	Knee	-	-	30	61	F	None	None	4.9	8	9
BX018	RA	RA	Knee	-	-	52	69	F	None	None	4.6	11	0
BX033	Res	Pseudogout	Ankle	-	-	7	81	F	None	None	6.7	60	52
BX038	Res	Parvovirus	Knee	-	-	1	45	F	None	None	4.0	4	0
BX049	VeRA	RA	Ankle	+	+	3	48	F	None	None	3.9	10	0
BX063	VeRA	RA	Knee	+	-	9	74	F	None	None	4.4	20	32
BX071	Res	Parvovirus	Ankle	-	-	4	41	F	None	None	1.7	5	9
BX072	Res	UA	Knee	-	-	10	32	M	None	None	3.6	15	0
BX075	RA	RA	Knee	-	-	52	22	F	None	None	6.4	81	79
BX081	Normal	Normal	Knee	-	-	NA	59	F	None	None	NA	NA	NA
BX089	Normal	Normal	Knee	-	-	NA	38	F	None	None	NA	NA	NA
BX095	Normal	Normal	Knee	-	-	NA	47	F	None	None	NA	NA	NA
BX097	Normal	Normal	Knee	-	-	NA	41	F	None	None	NA	NA	NA
BX098	Normal	Normal	Knee	-	-	NA	41	F	None	None	NA	NA	NA
BX121	VeRA	RA	Ankle	+	+	11	60	F	None	None	4.3	32	45
BX125	RA	RA	Knee	+	+	16	56	F	None	None	5.5	48	47
BX130	VeRA	RA	Ankle	-	+	12	61	F	None	None	5.4	43	18
BX143	Res	UA	Knee	-	-	16	45	F	None	None	3.0	8	4
BX157	Res	UA	Knee	-	-	2	45	F	None	None	2.6	8	17
RA05SY	Jrep	RA	Knee	NA	+	1040	62	F	None	None	6.8	63	62
RA07SY	Jrep	RA	Knee	NA	+	52	71	F	MTX	10mg	7.2	31	54
RA16SY	Jrep	RA	Knee	NA	+	780	30	F	Azathioprine	3mg	6.5	37	33
RA18SY	Jrep	RA	Knee	NA	+	1196	47	F	Etanercept	None	3.8	57	NA

RA19SY	Jrep	RA	Knee	NA	+	1040	62	F	Etanercept	None	4.1	8	8
RA22SY	Jrep	RA	Knee	NA	+	1560	70	F	Adalimumab	None	4.419537	20	NA

Table 16: Clinical characteristics of the patients from which fibroblast lines were isolated and used for the stimulated macrophage:fibroblast co-cultures. NA = Not available, MTX = Methotrexate, DAS28 = Disease activity 28 score, ESR = Erythrocyte sedimentation rate, CRP = C-reactive protein, CCP = Anti-cyclic citrullinated peptide, RF = Rheumatoid factor, MCP = Metacarpophalangeal, UA = Undifferentiated arthritis.

3.2.9 Enzyme-linked immunosorbent assay (ELISA) of co-culture supernatants

IL-6 levels present in the supernatants harvested from all co-cultures were measured by ELISA (Thermo Fisher Scientific). After thawing supernatants for use in an ELISA the remaining supernatant was aliquoted into working volumes before storing at -20^{o}C to prevent any further freeze/thaw cycles.

3.2.10 Luminex of co-culture supernatants

Supernatants from co-cultures were used in a Bio-Plex Pro™ Human Cytokine 27-plex Assay (BIORAD) as per the manufacturer's instructions. The full panel of factors assessed can be found in appendix 8.2.

3.2.11 Quantitative real-time polymerase chain reaction (qRT-PCR)

RNA was isolated from the cells using the PicoPure® RNA Isolation Kit (Thermo Fisher Scientific). The PicoPure RNA Isolation Kit was used according to the manufacturer's protocol with the modifications specified in the protocol appendix VI for isolation from cell pellets and DNase treatment and with the following additional modification; after step h the columns were left open to air dry for 5 minutes before elution of the isolated RNA in 11µl of elution buffer. Eluted RNA was stored at -80^{o}C. RNA was converted to complementary DNA (cDNA) using the iScript™ cDNA Synthesis Kit (Bio-Rad). cDNA was then used for various qRT-PCR reactions depending on the experiment. The primer sequences were designed using Primer-BLAST (NCBI) or obtained from John O'Neil (both synthesised by Eurofins) and are listed in Table 17. The reaction mix consisted of 2.4µl of sample, 0.3µl (0.05pmol/µl final concentration) of forward primer, 0.3µl (0.05pmol/µl final concentration) of reverse primer, and 3µl of SYBR® Premix Ex

Taq™ II (Tli RNase H Plus) (Takara) or 3µl of Absolute qPCR SYBR Green ROX Mix (Thermo Scientific). The thermocycler used was a LightCycler® 480 II (Roche) and the reaction parameters were 15 minutes at 95ºC before 40 cycles of 15 seconds at 95ºC followed by 30 seconds at 55ºC and finally 30 seconds at 72ºC.

Target	Forward primer 5`->3`	Reverse primer 5`->3`
GAPDH	GTCAGCCGCATCTTCTTTTGC	AATCCGTTGACTCCGACCTTCC
UBC	CGGGATTTGGGTCGCAGTTCTTG	CGATGGTGTCACTGGGCTCAAC
STAT1	GCGCGCAGAAAAGTTTCATTT	TGAGACATCCTGCCACCTTG
IL-6	GAGGCACTGGCAGAAAACAA	TGGCATTTGTGGTTGGGTCA
IL-8	CAGAGACAGCAGAGCACACA	GGCAAAACTGCACCTTCACA
TNF	CCCCAGGGACCTCTCTCTAAT	TCTCTCAGCTCCACGCCATT
IL-10	GCCTAACATGCTTCGAGATC	TGATGTCTGGGTCTTGGTTC

Table 17: Primers used for qPCR reactions

3.3 Results

3.3.1 Co-cultures of human synovial fibroblasts with the murine RAW264.7 cell line

Initial co-cultures were undertaken in order to establish a suitable protocol to allow interrogation of monocyte:fibroblast co-cultures. A starting co-culture duration of 5 days was used, the reasoning behind this decision being that many monocyte-macrophage differentiation protocols, including one commonly used in-house, culture monocytes for 5 days in the presence of M-CSF or GM-CSF to allow for full differentiation of the cells and in order to capture any potential differentiation effects that fibroblast:monocyte co-culture may have it was decided to co-culture for a similar amount of time.

The murine monocyte/macrophage cell line RAW264.7 was used during pilot qualitative experiments for two reasons. First, due to the high proliferative rate of

the cell line it provided a plentiful source of monocytic cells without the need for isolation from peripheral blood. Secondly, if the co-culture proved functional then the species difference may have allowed for easy identification of the source of soluble factors released during the co-culture. However, the obvious caveat to this experiment is that the species difference may result in spurious interactions and results that fail to represent the situation in human co-cultures.

Seeding fibroblasts at an initial density of 45,000 cells/well and resting the cells for 2 days prior to co-culture resulted in an over-confluent monolayer of cells before the co-culture had begun (Figure 28 **A** and **C**). Regardless of if 600,000 or 50,000 RAW cells were seeded into the co-culture by day five the cells had aggressively proliferated and obscured the underlying fibroblasts making any observations of these cells difficult. The qualitative analysis of co-culture images informed the choice of seeding density for fibroblasts in future co-cultures with 20,000 cells/well giving a reasonable level of confluence within the wells. The RAW264.7 co-cultures were discontinued in favour of human monocyte:fibroblast co-cultures due to (a) the difficulties of using a murine cell line and (b) the difficulties in managing proliferation of the RAW264.7 cell line.

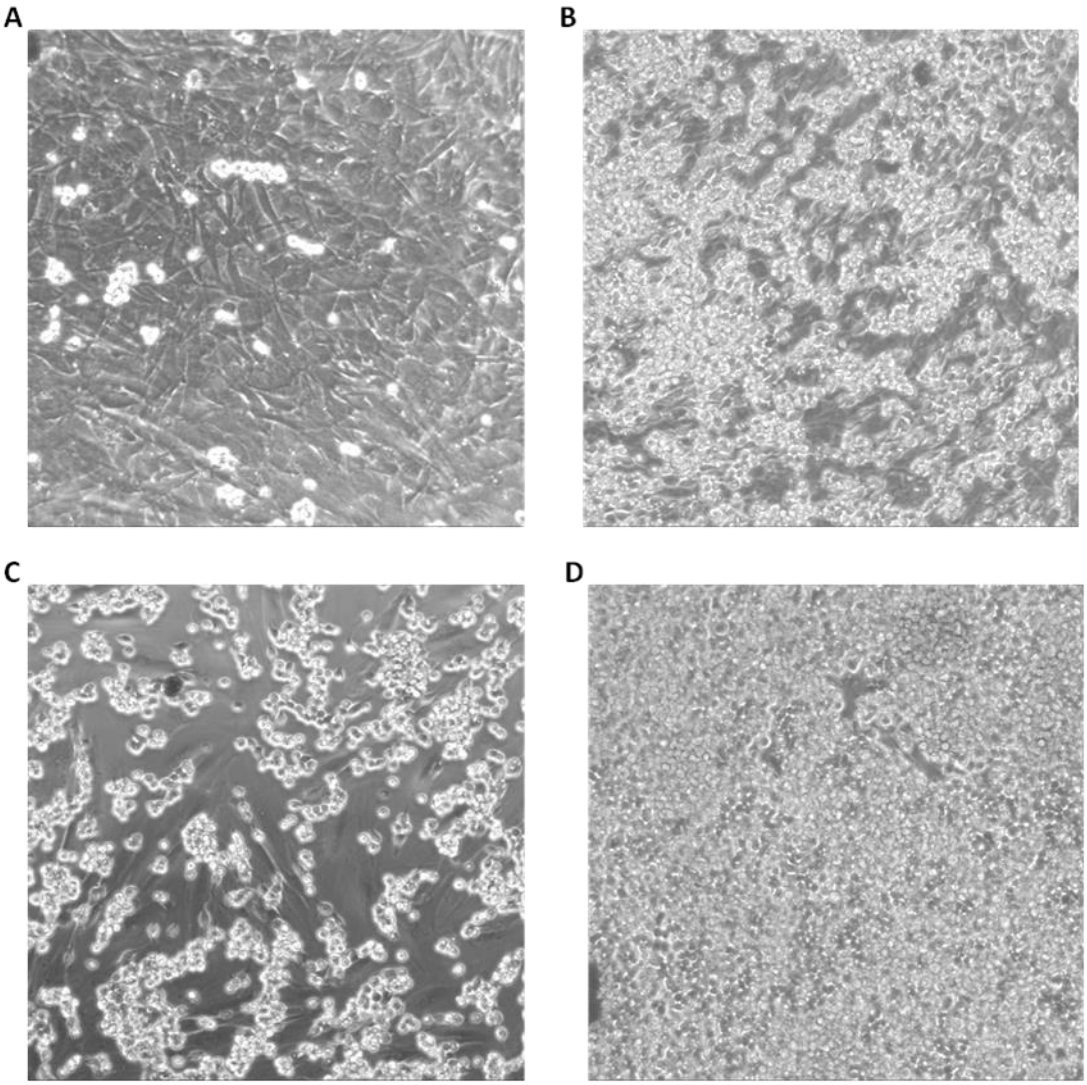


Figure 28: Microscopy analysis of 5 day co-cultures of RAW264.7 cells with synovial fibroblasts Rapid proliferation of the monocyte/macrophage line and a seeding density of 45,000 fibroblasts (A and C) resulted in an over-confluent monolayer of cells at the start of co-culture. All images were taken using differential interference contrast at 200x magnification. A 50,000 RAW264.7 cells on 45,000 fibroblasts cells at the start of the co-culture and B after 5 days of co-culture, and C 600,000 RAW264.7 cells on 45,000 fibroblasts at the start of the co-culture and D after 5 days of co-culture

3.3.2 Negative selection from blood cones using magnetic microbeads isolates unbiased monocyte subsets at a high purity

Two methods to isolate human monocytes from peripheral blood commonly used in the literature were assessed in three ways, the yield produced from total PBMCs, the purity of the isolated cells as assessed by flow cytometric analysis of CD14 and CD16 surface expression, and any bias for isolating one or two

monocyte subsets over others. Isolating monocytes using density-gradient separation followed by subsequent negative selection using MiltenyiBiotec MACs magnetic bead technology consistently resulted in a much higher yield and purity of monocytes compared to double density-gradient separation with subsequent adherence to tissue culture plastic for one hour with the average purity of the negative selection approach being 82.71% ± 4.96 (Mean ± 95% confidence interval (CI), n=6) and for the adherence approach being 30.51% ± 6.99 (Mean ± 95% CI, n=3) (Figure 29, representative figure). Both approaches isolated monocyte subsets in similar proportions in-keeping with those published in the literature; using the negative selection approach as a representative the proportion of classical monocytes was 87.39% ± 3.35 (Mean ± 95% CI, n=6), the proportion of intermediate monocytes was 3.79% ± 2.20 (Mean ± 95% CI, n=6), and the proportion of non-classical monocytes was 8.83% ± 1.91 (Mean ± 95% CI, n=6) (Ziegler-Heitbrock et al., 2010; Ziegler-Heitbrock and Hofer, 2013).

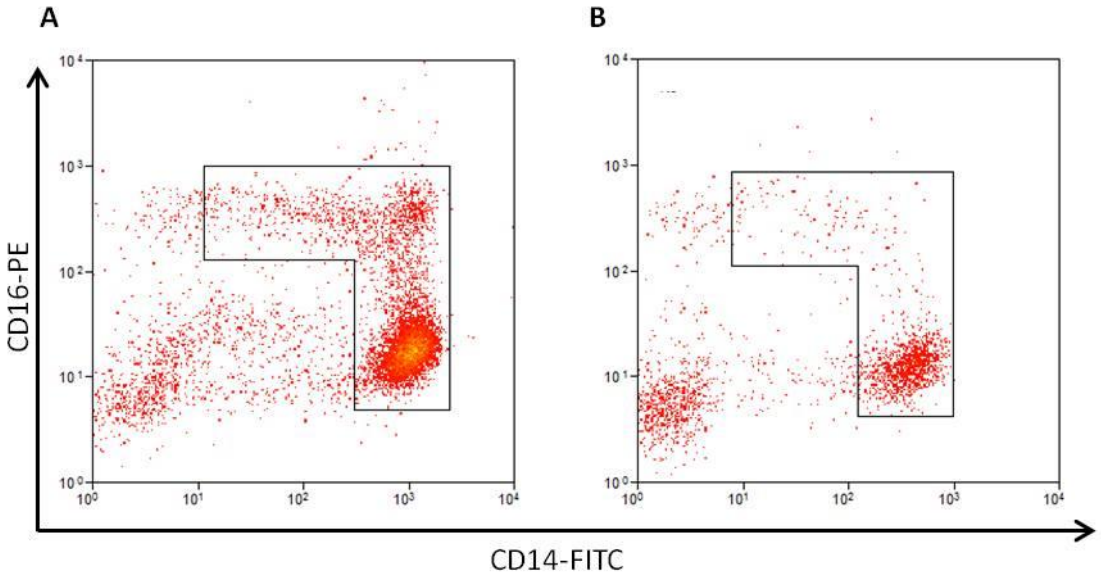


Figure 29: Representative dot plots of A monocytes isolated using negative selection, and B monocytes isolated using the adherence method. The gating demonstrates the method used to assess the purity of the monocyte populations

As density-gradient separation of peripheral blood using Ficoll-Paque followed by negative isolation of monocytes from PBMCs using magnetic bead technology provided a higher yield and purity of monocytes compared to the adherence method this technique was used for all future monocyte isolations both for direct co-culture and macrophage differentiation.

3.3.3 Accutase and trypsin remove co-cultures from culture vessels with similar efficacy whereas cell-dissociation buffer performs worse

As it was planned to isolate separate populations post co-culture for downstream analysis, the efficacy of different methods of cell removal was tested. Three methods of removal; trypsinisation, treatment with accutase, and treatment with CDB, were used to remove fibroblasts and monocytes from 24 well plates after 5 days of co-culture.

Treatment with trypsin or Accutase removed similar populations of fibroblasts from monoculture with the presence of two broad forward-scatter (FSC) low or FSC high populations visible by flow cytometry (Figure 30). However, the FSC high population was absent when the cells were removed from culture using CDB indicating this method does not remove all fibroblasts from tissue culture plasticware with equal effectiveness. With regards to the isolation of monocytes from monoculture all three methods remove similar populations from culture with the population isolated using trypsin displaying a reduction in variation in both the FSC and the side-scatter (SSC) channel. When assessing the number of doublets present in the samples isolated from monocyte-only cultures, trypsin produced an almost pure population of single cells whereas accutase and CDB both produced samples with a population of doublets with CDB producing the

largest doublet population. In the co-culture samples the FSC SSC profile resembled a combination of the fibroblast and monocyte monoculture profiles with the co-culture profile of the cell-dissociation buffer sample still showing a reduction in the number of FSC high cells in-keeping with that found in the fibroblast-only culture. It was apparent that the FSC SSC profiles of monocytes and fibroblasts overlap negating the possibility of using FSC SSC profiles to cell sort pure populations of each cell type post co-culture without the use of further surface marker staining, intracellular labelling of cells, or physically separated co-culture.

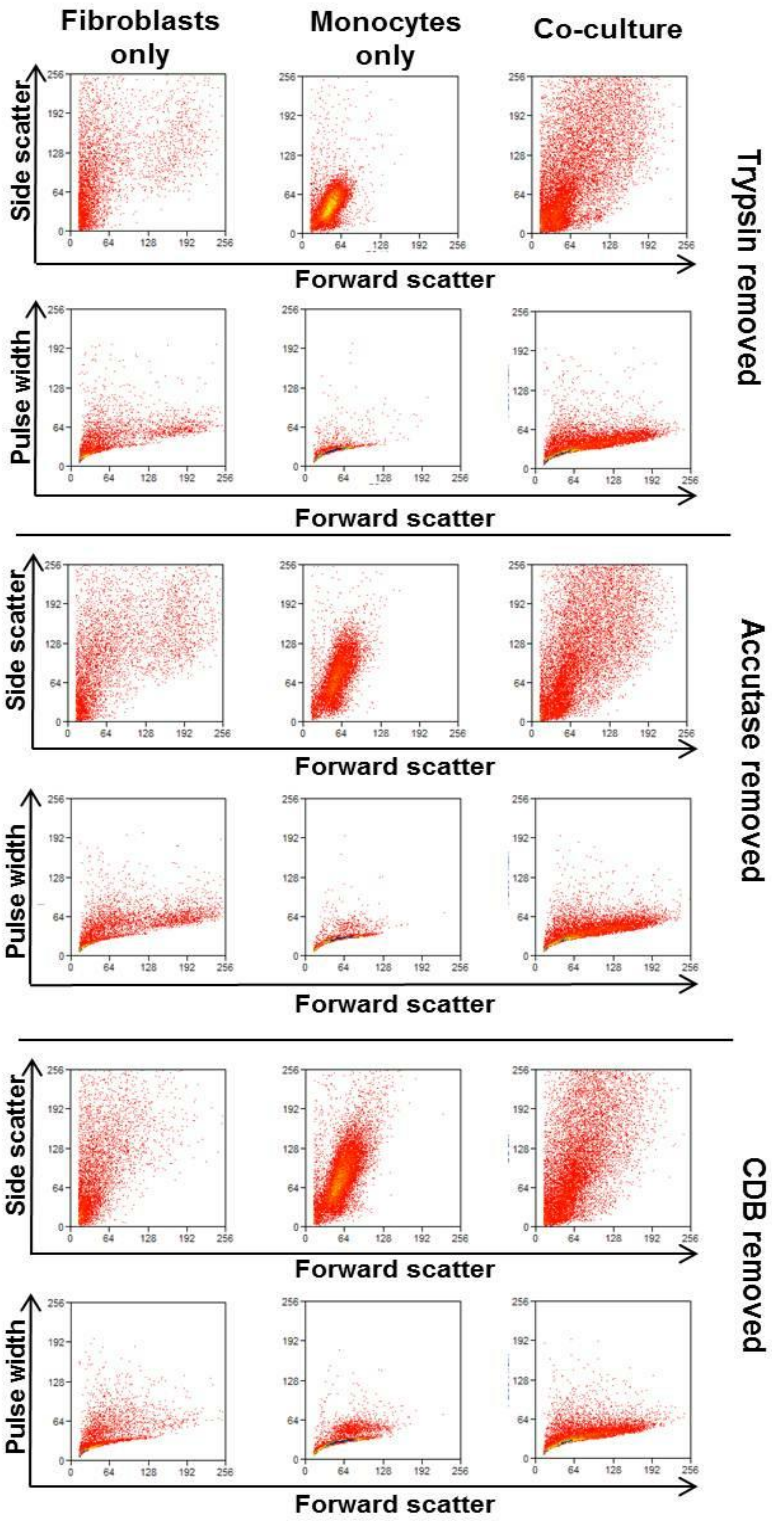


Figure 30: Representative dot plots of flow cytometric analysis of co-cultured cells removed from co-culture using Trypsin, Accutase, or Cell-dissociation buffer (CDB). It was not possible to identify clear populations of fibroblasts or monocytes from co-cultures due to the overlapping nature of the forward and side scatter profiles of the two cell types.

3.3.4 Direct and separated monocyte:fibroblast co-cultures stimulate a synergistic increase in IL-6 release

Co-cultures of freshly isolated human monocytes with fibroblasts from the BEACON cohort, specifically: (1) fibroblasts isolated from tissue removed during arthroplasty of RA patients (Joint replacement fibroblasts, Jrep), (2) fibroblasts isolated from treatment naïve patients who either presented with or would subsequently develop persistent RA but who had had symptoms for less than 3 months (Very early RA fibroblasts, VeRA), and (3) fibroblasts isolated from treatment naïve patients who were suffering from arthritis symptoms for less than 3 months which spontaneously resolved (Resolving fibroblasts, Res), were completed. Three techniques were used to co-culture the cells in order to interrogate the role of cell contact in the interactions between fibroblasts and monocytes (Figure 31). In direct co-culture fibroblasts and monocytes were seeded into the same well enabling full contact between cell types. In the Transwell co-cultures fibroblasts were seeded and allowed to adhere onto the underside of a filter with 0.4µm pores. The filter was then placed into a well with monocytes seeded in the upper chamber allowing the fibroblasts suspended below and the monocytes above a limited amount of contact. In separate co-cultures the fibroblasts were seeded into the upper chamber of a Transwell filter whereas the monocytes were seeded into the plate below thus preventing cell to cell direct contact.

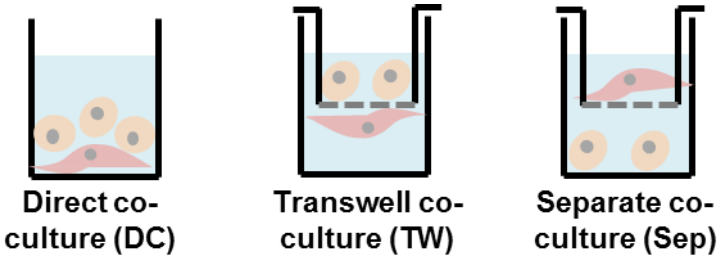


Figure 31: Schematic of the three different assay setups used to investigate the contribution of contact-dependent and contact-independent interactions to the release of soluble factors. In direct co-culture fibroblasts and monocytes were seeded into the same well enabling full contact between cell types, Transwell co-cultures allowed a reduced level of contact whereas separate co-cultures prevented all direct cell contact.

Monocytes cultured alone for 5 days with no stimuli typically release basal levels of IL-6 with median levels released in these controls being 29.26pg/ml ± 33.45 (Median ± 95% CI) (Figure 32). Unstimulated fibroblast monocultures also results in low levels of IL-6 release, albeit higher than monocyte monoculture, with levels in direct monoculture being 120.28pg/ml ± 69.07 (Median ± 95% CI). Direct and separate co-culture of synovial fibroblasts with monocytes elicited a synergistic and statistically significant increase in the concentration of IL-6 released into the culture media regardless of the disease outcome group from which the fibroblast line was isolated (direct co-culture IL-6 levels 3116.578pg/ml ± 2824.02 (Median ± 95% CI), separate co-culture IL-6 levels 931pg/ml ± 892.83 (Median ± 95% CI)). The increase in transwell co-culture IL-6 levels did not reach statistical significance (461.22pg/ml ± 153.03 (Median ± 95% CI)).

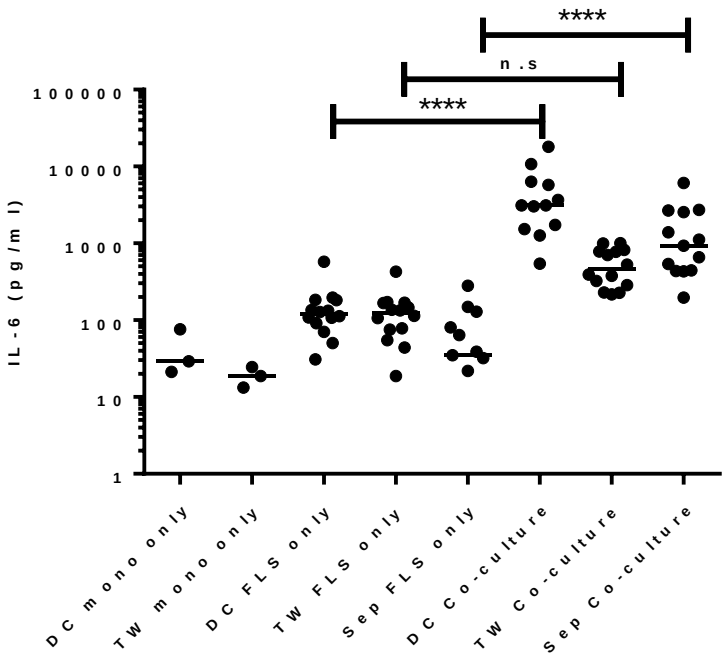


Figure 32: IL-6 release from monocyte and fibroblast cultures after 5 days as assessed by ELISA. Statistical analysis was performed using a Kruskal-Wallis test with Dunn's multiple comparisons test. **** = p<0.0001. Bars show the median. DC = Direct co-culture, TW = Transwell co-culture, Sep = Separate co-culture, Mono = monocyte, FLS = Fibroblast-like synoviocyte.

When the results of the co-cultures are stratified according to the disease group of the fibroblast line it becomes apparent that regardless of the type of co-culture IL-6 levels are similar (Figure 33). This finding favours the null hypothesis that fibroblasts isolated from different clinical outcome groups all modulate monocyte phenotype similarly. However, as it was possible that factors other than IL-6 produced by monocytes and/or fibroblasts could be differentially modulated but not detected by this approach, a 27-plex luminex assay was performed on supernatants from the direct co-cultures.

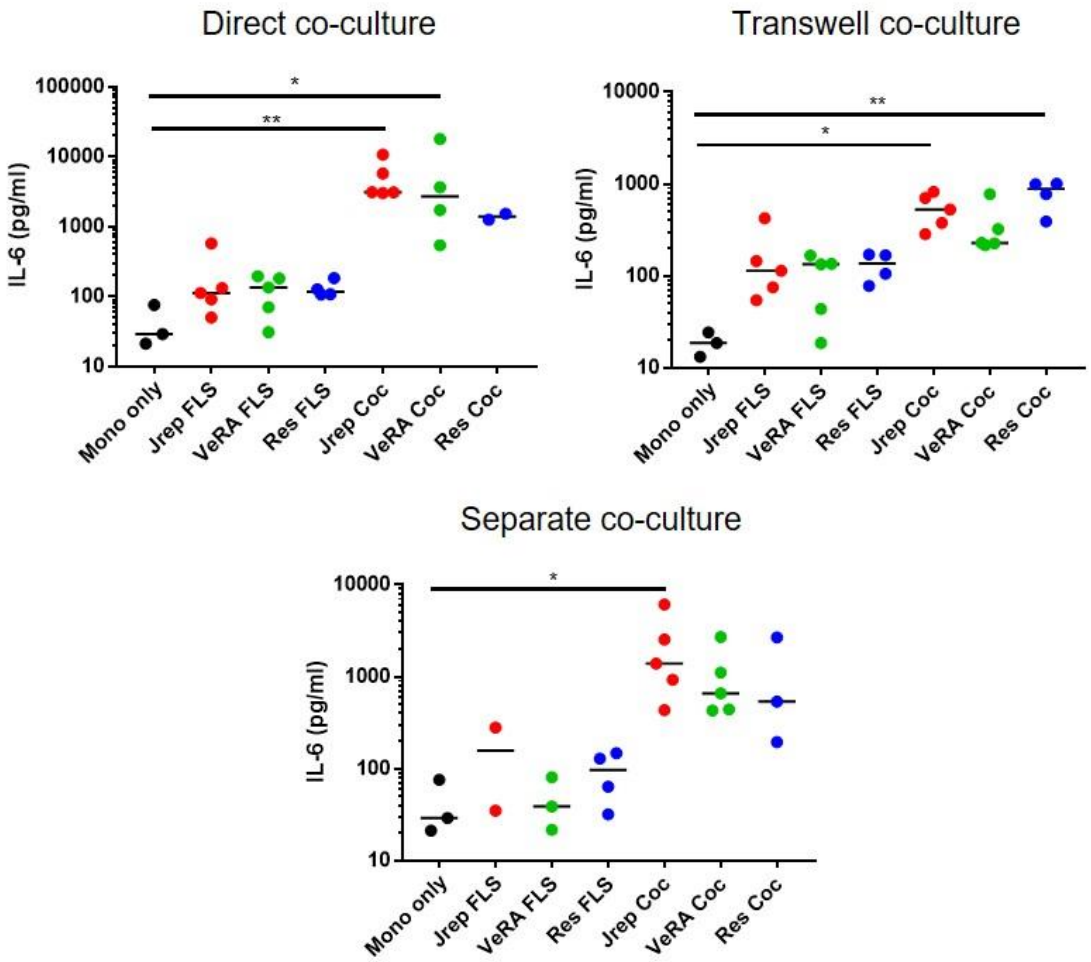


Figure 33: Co-culture of monocytes with fibroblasts from different disease outcome groups did not induce differential IL-6 release regardless of if the cells were in contact or not. Tested using Kruskal-Wallis test with Dunn's multiple comparisons test. * = p<0.05, ** = p<0.01. Bars show the median. Mono = monocyte, FLS = Fibroblast-like synoviocyte, Coc = Co-culture.

Results from the luminex assay found several factors that were either up- or down-modulated by direct co-culture. However, the concentration of many of the cytokines, chemokines, or growth factors often remained low making the interpretation of the biological significance of the findings questionable (Figure 34, Table 18, full results in 8.2). Also, as CCL2 and CCL4 are both monocyte chemoattractants it could be argued that the decrease in the extracellular concentration of these chemokines is due to uptake of these factors by monocytes. Regardless of the biological significance of the results it appeared

that no differential modulation of soluble factors by disease group of the fibroblast line in co-culture could be seen. The factors with the highest levels of production were IL-6 (1583.54 ± 1343.67 (Median ± 95% CI)) and CXCL10 (1478.98 ± 917.96 (Median ± 95% CI)) with VEGF being the next highest (437.51 ± 126.91 (Median ± 95% CI)) but considerably lower than both IL-6 and CXCL10.

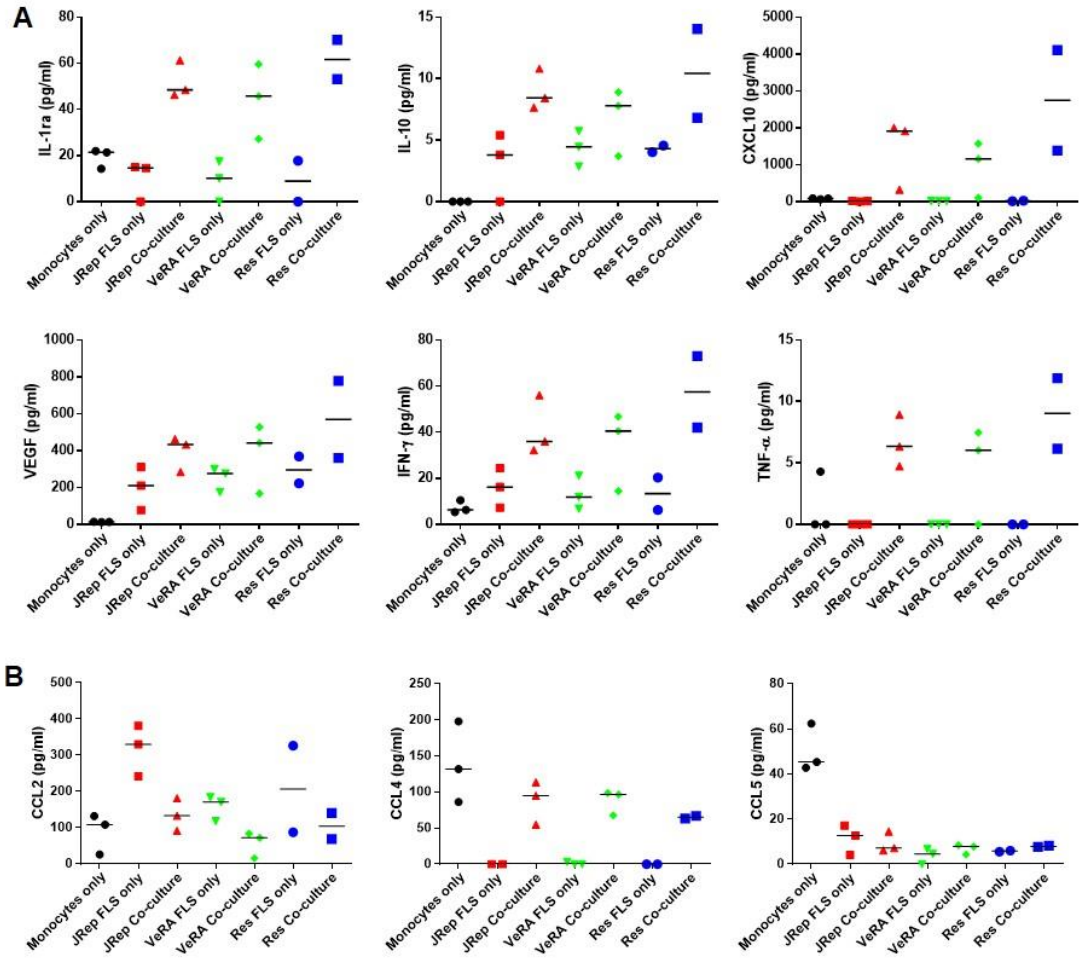


Figure 34: Co-culture of monocytes and fibroblasts resulted in the up- and down-regulation of the release of several soluble factors as assessed by a Luminex assay but not differential modulation dependent on the disease group from which the fibroblast line was isolated. (A) Cytokines and chemokines upregulated by co-culture, (B) cytokines and chemokines decreased or unregulated by co-culture. Although the concentration of some factors were increased or decreased by co-culture the levels often remained low making the biological significance of these findings questionable. No significant differences between any of the co-cultures as tested by a Kruskal-Wallis test with Dunn's multiple comparison test. Bars show the median.

Mediator	Direction of regulation	Median concentration in co-culture (pg/ml)	Statistically significant?
IL-1Ra	Up	50.835	FLS v Coc
IL-6	Up	1583.54	Mo v Coc
IL-10	Up	8.11	Mo v Coc
TNFα	Up	6.255	FLS v Coc
IFNγ	Up	41.26	Mo or FLS v Coc
VEGF	Up	437.51	Mo v Coc
CXCL10	Up	1478.98	FLS v Coc
CCL5 (RANTES)	Down	7.755	Mo v FLS
CCL2 (MCP1)	Down	87.585	FLS v Coc
CCL4 (MIP-1β)	Unregulated	81.425	Mo v FLS, FLS v Coc

Table 18: Summary of the luminex results plotted in Figure 34 with IL-6 luminex results added and with disease stages pooled. FLS = fibroblast like synoviocyte only, Coc = co-culture. Tested using Kruskal-Wallis with Dunn's multiple comparisons test.

Variability between cone samples remained a problem. The blood apheresis cones provided by the NHSBT are not provided with information on the time of blood collection, the time the apheresis was completed, nor the number of apheresis cycles used to produce the cone. Endeavours were made to obtain cones at similar times and days of the week to try to reduce variability however it was decided to investigate whether the increase in IL-6 elicited by co-culture was an artefact present due to the source of monocytes or a finding consistent regardless of the source. To this end three to four fibroblast lines were co-cultured in parallel with monocytes obtained from either freshly drawn peripheral blood from a healthy donor or with monocytes obtained from an apheresis cone, with monocytes from both sources isolated in parallel. No significant difference was found between IL-6 levels from co-cultures (Figure 35).

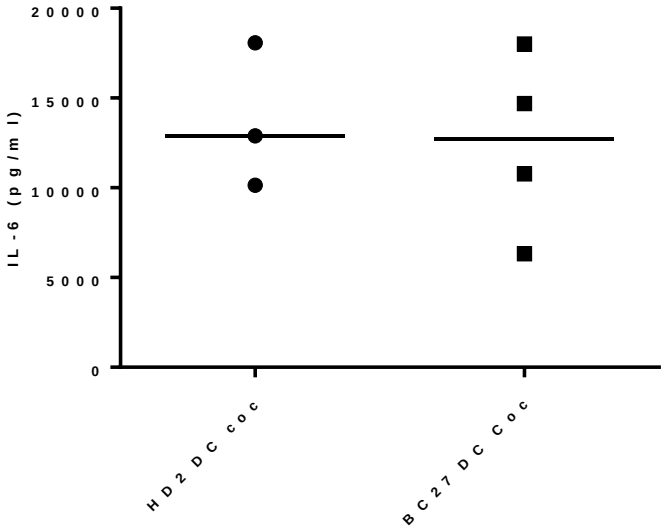


Figure 35: IL-6 levels released during 5 day co-culture of the same 3-4 fibroblast lines with monocytes obtained from either freshly drawn peripheral blood from a healthy donor (HD2) or from a peripheral blood apheresis cone (BC27). No significant difference existed between co-cultures with monocytes from the two sources. Tested using a Mann-Whitney test. Bars show median.

All co-cultures were run for 5 days in a similar manner to monocyte to macrophage differentiation protocols to allow for any differences that may depend on differentiation to manifest. As such the temporal dynamics of IL-6 release during co-culture were unknown. To investigate the kinetics of IL-6 release co-cultures were run for durations of 30 minutes, 1 hour, 2 hours, 4 hours, 18 hours, 1 day, 2 days, 3 days, and 5 days and supernatants were harvested before measuring the concentration of IL-6 by ELISA. The results show that the IL-6 concentration in fibroblast or monocyte monoculture increased over a 24 hour period with maximal levels in the monocyte monocultures approached within 24 hours of culture and with a continuous increase in levels in the fibroblast monocultures over 5 days (Figure 36). IL-6 levels during co-culture increased in a similar manner, but with a greater magnitude, to the monocyte monocultures. This increased concentration of IL-6 was then maintained for a duration of at least 5 days.

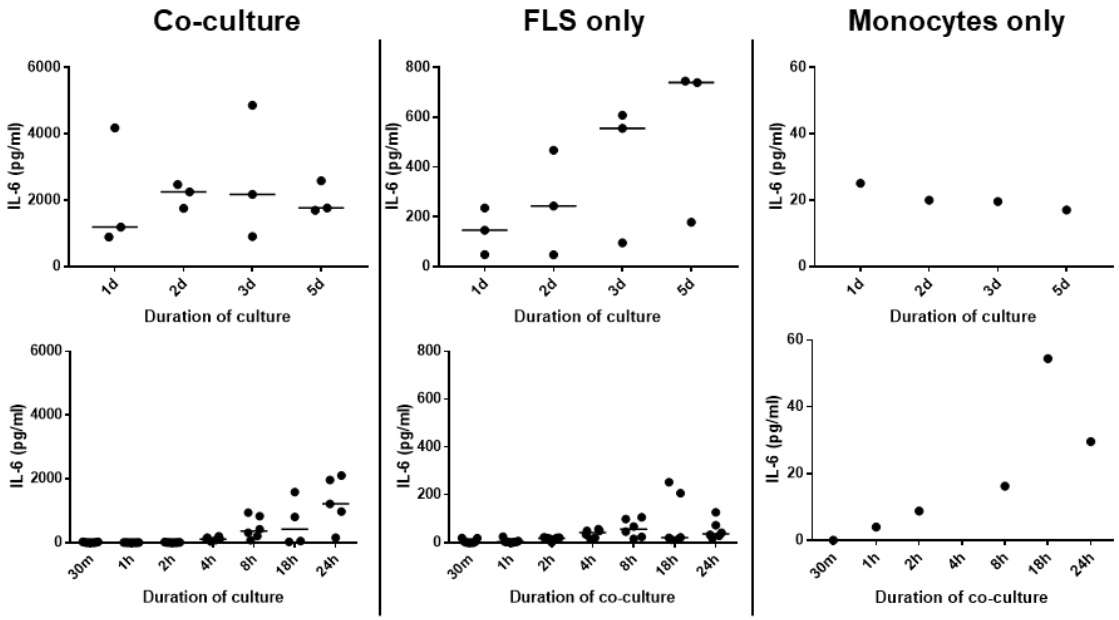


Figure 36: Concentration of IL-6 released into the medium during fibroblast and monocyte direct co-culture for varying timescales. The majority of IL-6 release occured within 24 hours of the initiation of co-culture and was maintained for 5 days. No significant increase was detected between any of the daily timepoints in co-culture compared to monocultures presumably due to low numbers affecting the power of the test. Although not shown on the plot significant differences did exist between the timepoints earlier than 1 day of co-culture with a difference between 1h and 4h being the earliest detected. Significant increases in the basal IL-6 production by FLS were also found with a difference between 30m and 4h being the earliest detected. All tested using a Kruskal-Wallis test with Dunn's multiple comparison test. Bars show the median.

In order to assess the contribution of contact dependent or contact independent interactions to the increase in IL-6 release a blocking antibody against the integrin CD18, which is expressed by monocytes, was used (Campbell et al., 1998). ICAM-1 is known to be expressed on fibroblasts and ICAM-1 and CD18 are known to bind and interact (Liu and Piela-Smith, 2000; Yang et al., 2010; Zhou et al., 1994). It was attempted to block CD18 ICAM-1 interactions in 5 day co-cultures with a single addition of 5µg/ml anti-CD18 blocking antibody at the initiation of co-culture and assessing the impact of treatment by measuring IL-6 concentration in the co-culture media at the end of the co-culture. The effect of the blocking antibody was unclear with a seemingly inhibitory effect on IL-6

release in the direct co-cultures, hints of a stimulatory effect in the Transwell co-cultures, and no effect in the separate co-cultures (Figure 37). However, the isotype control antibody also reduced IL-6 release in the direct co-culture, suggesting that the effect seen was a result of non-specific antibody binding in the co-culture.

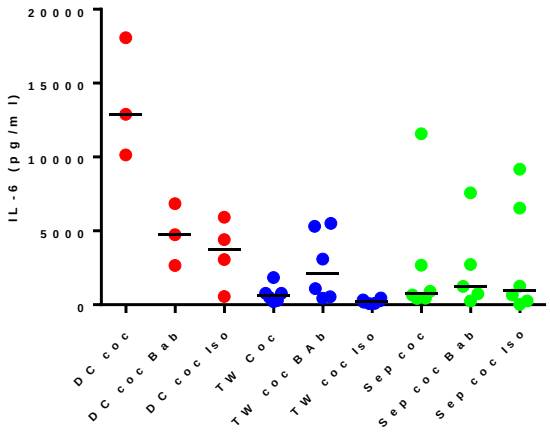


Figure 37: Blocking CD18-ICAM-1 interactions using 5µg/ml of an anti-CD18 blocking antibody at the start of 5 day co-cultures had no clear effect on the final concentration of IL-6 released into the co-culture medium as assessed by ELISA. No significant differences were detected using a Kruskal-Wallis test with Dunn's multiple comparisons test. Bars show the median. DC = direct co-culture, TW = transwell co-culture, Sep = separated co-culture, Bab = blocking antibody, Iso = isotype control.

As examination of extracellular cytokine levels had not found significant differences in the regulation of monocyte phenotype by fibroblasts from different disease outcome groups it was decided to investigate the intracellular levels of cytokines. The advantage of such an approach in the unstimulated co-cultures is that although the profile of cytokines being synthesised may be altered by co-cultures with different fibroblast lines, the release of cytokines may not be significantly altered without additional stimuli to elicit the effect. Therefore it was decided to stain for two signature intracellular macrophage cytokines, TNFα and IL-10 respectively, to assess if any differential modulation/differentiation was

driven by different fibroblast lines. Fibroblasts were labelled with an intracellular dye prior to co-culture to allow for easy identification of fibroblasts and monocytes after 5 days of co-culture.

Figure 38 shows the results of the intracellular cytokine staining of co-cultured monocytes and fibroblasts. In both cell types the median fluorescent intensities (MFIs) for both cytokines were either similar to or only slightly higher than those of the relevant isotype controls indicating that the intracellular levels of these cytokines are low. The mean MFI of TNFα in co-cultured monocytes was not significantly higher than that of the isotype control with variation in MFIs between monocytes co-cultured with different fibroblasts lines being low, indicating that co-cultured monocytes do not express intracellular TNFα. No significant difference was found between either fibroblasts or monocytes in co- or mono-culture compared to relevant isotype controls. With regards to intracellular IL-10 levels in co-cultured monocytes the average MFI is slightly higher than that of the relevant isotype control but not significantly so with a similar low level of variation between samples indicating a lack of differential regulation by different fibroblast lines. Interestingly fibroblasts seem to have higher levels of IL-10 than monocytes with both mono- and co-cultured fibroblasts having significantly higher MFIs than the relevant isotype controls but not than each other when tested using a Kruskal-Wallis test with Dunn's multiple comparisons test. Whether IL-10 is produced by the fibroblasts or taken up during co-culture is unknown.

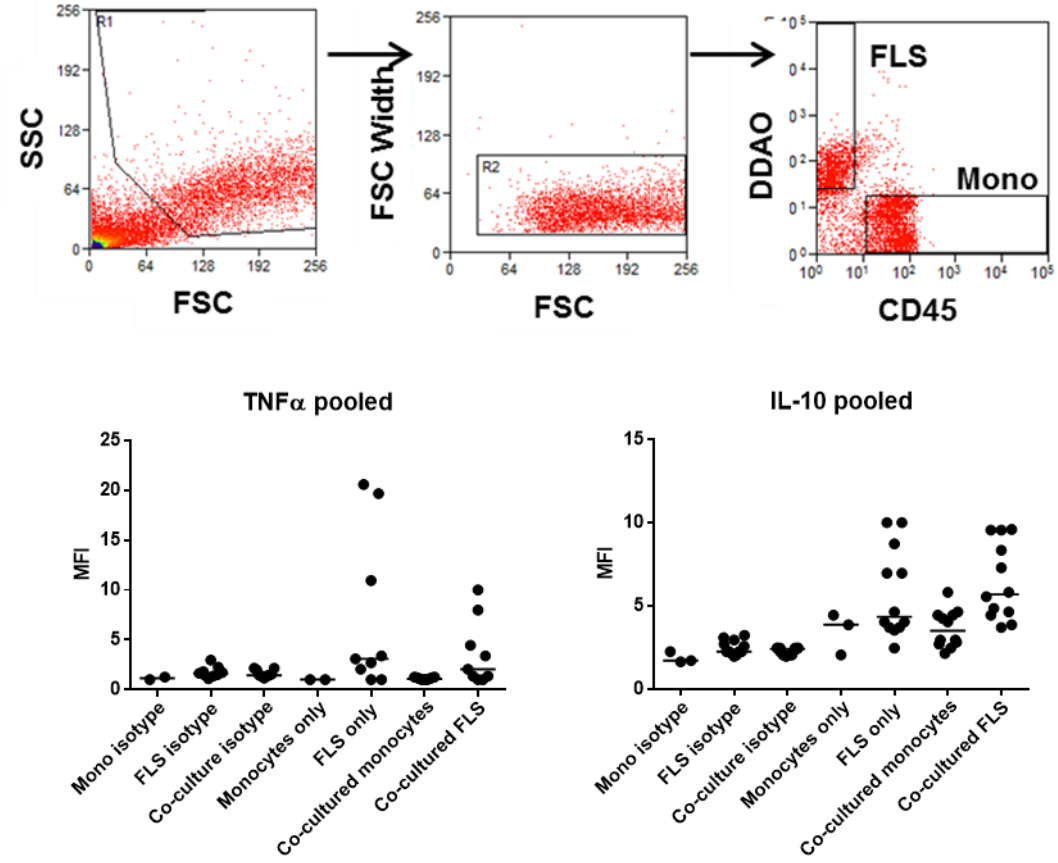


Figure 38: Flow cytometric identification of DDAO labelled fibroblast and CD45$^+$ monocyte populations after 5 days of co-culture and measurement of intracellular TNFα and IL-10 levels. Levels detected remained low with median fluorescence intensities (MFIs) in the stained samples being in-line with or only slightly above the MFIs from isotype control stained samples. No differential expression of TNFα or IL-10 was apparent between fibroblasts from different disease outcome groups or from monocytes co-cultured with these fibroblasts. Tested using Kruskal-Wallis test with Dunn's multiple comparisons test. Bars show the median.

Considering the evidence so far it was apparent that fibroblasts isolated from different disease outcome groups did not appear to differentially regulate monocyte cytokine or chemokine expression during unstimulated 5 day co-cultures as assessed by investigation of the cytokine profile. There were several caveats to the investigations so far including the fact that all co-cultures were run for 5 day periods and no other timepoint was investigated and that all co-cultures were unstimulated. It was therefore decided to address the second point using both differentiating and stimulating factors.

3.3.5 Co-culture of macrophages with fibroblasts from different disease outcome groups does not differentially affect macrophage differentiation

To assess if fibroblasts isolated from different stages of RA could differentially affect the *in vitro* differentiation of monocytes to macrophages freshly isolated monocytes from apheresis cones were cultured with either 50ng/ml GM-CSF or 100ng/ml M-CSF, to induce a pro- or anti-inflammatory phenotype respectively, for two days before co-culturing with fibroblasts from different outcome groups for a further two days still in the presence of either GM-CSF or M-CSF. At the end of this timepoint RNA was isolated from the macrophages and used in qRT-PCR experiments to analyse the expression of genes differentially expressed in pro- and anti-inflammatory macrophage subsets. *TNF* is an obvious pro-inflammatory gene and *STAT1* has been found to drive the expression of the pro-inflammatory inducible-nitric oxide synthase whereas *IL-10* encodes a well-known anti-inflammatory cytokine (Ganster et al., 2001). No significant differences were found comparing the fold change in gene expression induced by co-culture with fibroblasts from different clinical outcome groups for any of the three genes tested: *TNF*, *IL-10*, or *STAT1* (Figure 39). There was a suggestion that macrophages differentiated in GM-CSF co-cultured with fibroblasts from resolving arthritis may downregulate *STAT1*; increasing the number of biological replicates could have clarified this effect.

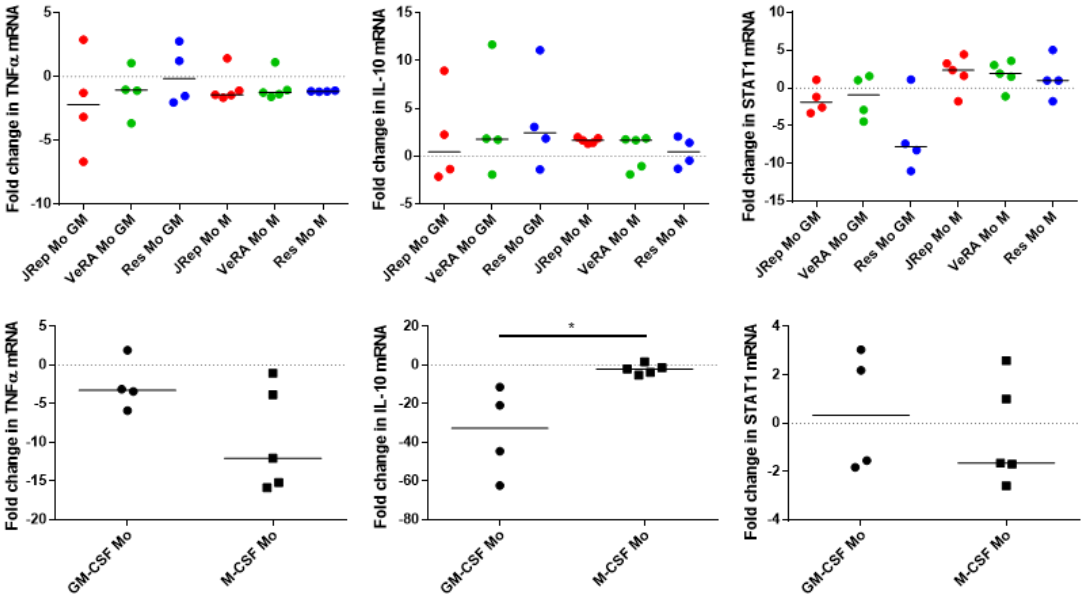


Figure 39: Fold-change in expression of the pro-inflammatory genes *TNF* and *STAT1* and anti-inflammatory gene *IL-10* expression in macrophages co-cultured with fibroblasts from different disease stages in the presence of GM-CSF or M-CSF relative to expression in macrophage mono-cultures in the presence of these growth factors. Gene expression normalised to *UBC* expression. No significant difference was detected in the modulation of gene expression by fibroblasts from different disease stages when analysed using a Kruskal-Wallis test with Dunn's multiple comparison test. However when pooling data from different disease stages a significant difference existed *in IL-10* mRNA expression between treatments as assessed using a Mann-Whitney test. * = p<0.05. Bars show the median. Jrep = Joint replacement, VeRA = very early RA, Res = Resolving arthritis, Mo = Monocyte, GM = GM-CSF, M = M-CSF.

As the co-culture of differentiating monocytes with fibroblasts from different disease stages did not appear to result in differential modulation of signature pro- or anti-inflammatory genes it was hypothesised that differences elicited by the fibroblasts may not be apparent in the absence of pro-inflammatory stimuli. Although no differential modulation of phenotype by fibroblasts from different clinical outcome groups had yet been demonstrated work on the BEACON cohort has demonstrated functional differences both in fibroblast-endothelial crosstalk and fibroblast motility (Filer, 2017; Juárez Pérez, 2014). Recent work indicated that fibroblasts are able to modulate the expression of genes in macrophages normally up- or down-regulated by stimulation (Donlin et al., 2014). It was decided

to adapt the approach used in this work to begin to assess if fibroblasts from different disease stages have different impacts upon stimulated gene expression in macrophages.

3.3.6 Stimulated macrophage:fibroblast co-cultures

Peripheral blood monocytes were differentiated for two days in 10ng/ml M-CSF before co-culturing with fibroblasts in the presence of 20ng/ml TNFα for 16 hours. After 16 hours of co-culture and stimulation RNA was extracted from the macrophages and the expression of a set of interferon stimulated genes and non-interferon stimulated genes was assessed using qRT-PCR. Only 16 hours of co-culture and stimulation was allowed in order to capture genes expressed in the early/intermediate gene response cassette as in Donlin et al. (2014). Across 4 healthy donors results were mixed compared with the expected findings; downregulation was sometimes seen in the genes *CXCL9*, *CXCL10*, *IFIT1*, and *MX1* and *NKG7* and *IL1B* remained unchanged, as expected from the published results (Figure 40). However modulation was often seen in the two 'control' genes, *NKG7 and IL1B*, and no modulation seen in the interferon stimulated genes. Additionally there was heterogeneity in response to fibroblast co-culture between monocyte donors in *CXCL10* expression.

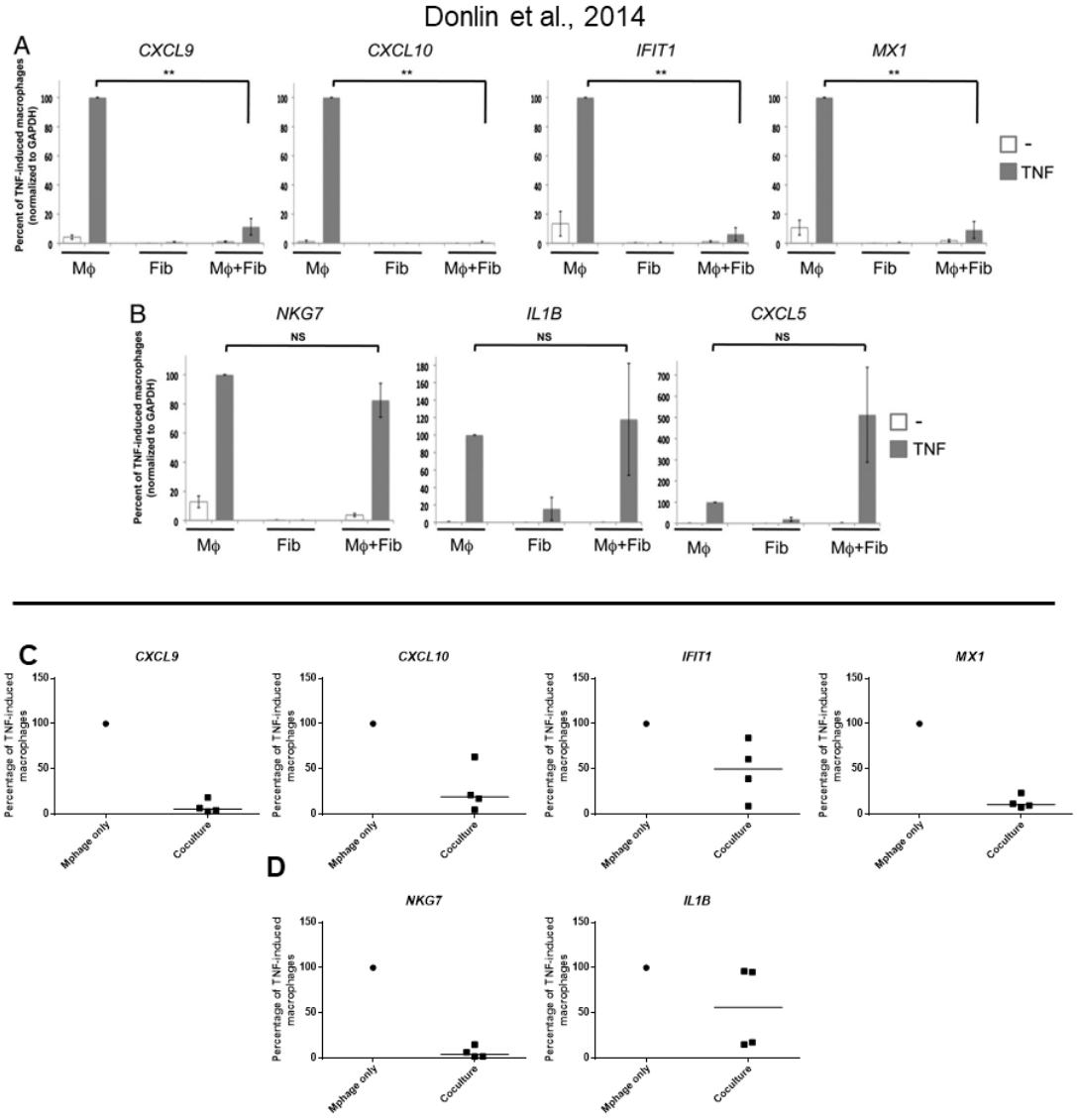


Figure 40: Representative results of gene expression modulation in macrophages co-cultured with fibroblasts for 16 hours in the presence of 20ng/ml TNFα. Expression was normalised to *GAPDH* and calculated relative to expression in TNFα stimulated macrophages in monoculture. In the publication by Donlin et al. (2014) (top plots adapted from this publication, **$p < 0.001$) *CXCL9*, *CXCL10*, *IFIT1*, and *MX1* (A) were all downregulated by co-cultured compared to macrophage mono-culture whereas *NKG7* and *IL1B* (B) were both unchanged by co-culture however the same pattern of modulation was not always seen in the work in these experiments (C and D) as exemplified by *NKG7* and *IL1B* expression. Bars in the bottom plots show the median.

Despite the more variable findings of this work it was decided to prepare co-culture samples for RNA-sequencing. This prepatory work was completed, but sequencing work was not completed as part of this thesis, other work having taken priority (see chapter 4).

3.4 Discussion

The work in this chapter successfully established a protocol for the co-culture of heterologous peripheral blood derived monocytes/macrophages and synovial fibroblasts. Co-cultures of fibroblasts and monocytes modulated the production of IL-6, CXCL10, and CCL4 amongst other factors independent of the clinical outcome group from which the fibroblast lines were isolated. Macrophage:fibroblast co-cultures modulated macrophage gene expression both in the presence or absence of TNFα independent of fibroblast clinical outcome group.

Regarding the choice of monocyte source peripheral blood monocytes isolated from commercially available apheresis cones provided by the NHSBT service were used for all co-cultures. There is logic to using autologous monocytes isolated from the peripheral blood of the same patients from which the fibroblast lines they were to be co-cultured with however the logistical challenges of this approach, given the 2-3 month duration required for the isolated fibroblasts to reach a useable passage, necessitated the use of other sources of monocytes. Additionally the use of a different monocyte donor for every fibroblast line may introduce a large amount of variation into the data. Furthermore it is possible that monocyte phenotype will change as disease progresses from early to later stages further confounding experimental results. For these reasons heterologous monocytes were used with several fibroblasts from different clinical outcome groups being co-cultured with each batch of monocytes. A final point of concern with the use of apheresis cones is that no information is provided regarding the length of time since blood was drawn and the cones were generated which may

lead to increased variation in the data or even spurious results. To address this question four fibroblast lines were co-cultured in direct contact with monocytes either from freshly drawn peripheral blood or apheresis cones, with no significantly different IL-6 levels found, suggesting that the true difference between these allogeneic approaches may not be great.

During the transwell co-cultures, where cell-cell contact is reduced but not prevented, the increase in IL-6 in co-cultures does not reach statistical significance. The reasoning behind this is unclear however an obvious difference between the transwell co-cultures and the other co-cultures is that fibroblasts are suspended from the underneath of the filters whereas in the direct and separate co-cultures the fibroblasts are resting on either a membrane or the cell culture plastic. This may induce cytoskeletal changes leading to changes in the regulation of gene expression similar to the changes seen by Van Den Berg et al. (2006) in which the disruption of elements of the cytoskeleton lead to a decrease in translational efficacy of IL-6 mRNA in response to TNFα stimulation in epithelial-like cells. Additionally, the transwell membrane may modulate the amount of cell-cell contact in such a way as to limit the IL-6 release. Importantly the increase in IL-6 release in the direct and separate co-cultures did not significantly differ between co-cultures with fibroblasts isolated from different BEACON outcome groups meaning this data fails to support the alternative hypothesis of differential regulation.

The synergistic increase in IL-6 levels elicited by monocyte:fibroblast co-culture is not unique to this study and has previously been demonstrated in other work.

Chomarat et al co-cultured synovial fibroblasts and peripheral blood monocytes together for 48 hours and also noted a synergistic increase in IL-6 production and found that this increase was comparable between co-cultures with RA synovial fibroblasts and normal synovial fibroblasts and was dependent on the cell numbers of both cell types in co-culture (Chomarat et al., 1995). This study also characterised a rapid increase in IL-6 levels during the first 48 hours of co-culture and upon extending the co-culture found that maximal IL-6 levels in the supernatant occurred around day 7 reaching approximately 10ng/ml. Furthermore, the authors began to tease apart the mechanisms behind the IL-6 increase finding that fixing monocytes prevents the IL-6 increase and using IL-1R agonist, anti-CD14 or anti-CD13 blocking antibodies, or treatment with IL-4, IL-10 or IL-13 inhibited the IL-6 release. It would be of interest to attempt similar experiments on future co-cultures with the BEACON fibroblast lines to assess if different mechanisms are involved in interactions with fibroblasts from different clinical outcome groups. The pre-monocytic cell line U937 co-cultured with RA synovial fibroblasts results in an increase in *in vitro* degradation of cartilage that is dependent on fibroblast contact with cartilage and dependent on TNF-α, IL-1β, IL-6, and CD44 but not ICAM-1 in the co-culture (Scott et al., 1997). Co-culture of RA synovial fibroblasts with the U937 monocytic cell line has also been shown to increase CD14 and CD11b expression on the U937 cells indicating some degree of differentiation may be driven by the fibroblasts (Chen et al., 1998).

The increase in IL-6 elicited by co-culture of monocytes and fibroblasts is important in the context of RA given the multiple findings implicating this cytokine in the pathogenesis of the disease. The semiquantitive score of the expression

of sublining CD68 and IL-6 in RA synovium as measured by immunohistochemistry correlates with knee pain in a multivariate analysis (Tak et al., 1997). Additionally, IL-6 protein is constitutively expressed by RA synovial fibroblasts and is found at higher levels in the serum and synovial fluid of RA patients than in OA patients and expression in RA synovial fibroblasts can be increased with IL-17 stimulation, an important point given the increased IL-17 levels present in the RA synovium (Chabaud et al., 1999; Hwang et al., 2004; Okamoto et al., 1997; Scaife et al., 2004). Importantly even though genes such as *TNF*, *CXCL13*, and basic fibroblast growth factor (*FGF2*) are differentially expressed in the published synovial phenotypes in RA *IL-6* is consistently expressed across all phenotypes indicating the cytokine plays an essential role in all forms of RA (Dennis et al., 2014).

Although only a few factors appearing to increase in co-culture were detected at biologically significant levels those detected have previously been implicated in RA. CXCL10 has been found to stimulate increased invasiveness of RA synovial fibroblasts and arthritic rat synovial fibroblasts via CXCR3 signalling (Laragione et al., 2011). Additionally, plasma CXCL10 levels in untreated early RA are higher than in healthy controls and correlate with disease activity and a polymorphism in the CXCL10 gene is associated with RA and extra-articular manifestations of the disease (Kotrych et al., 2015; Pandya et al., 2017). Genetic knockout of the chemokine in the murine collagen antibody induced arthritis reduced macrophage and $CD4^+$ T cell infiltration into the synovium (Lee et al., 2017). Regarding the role of IFNγ in RA the cytokine has been found at higher levels in the RA synovium than in osteoarthritis and increases the motility and migration of

synovial fibroblasts in *in vitro* assays (Dolhain et al., 1996; Karonitsch et al., 2010). In contrast to the findings regarding CXCL10, IFNγ has been found to reduce the increase in MMP1 and MMP3 production in RA synovial fibroblasts elicited by IL-1β stimulation *in vitro* and reduce fibroblast mediated cartilage damage in an *in vitro* assay (Page et al., 2010). The final of the three main upregulated cytokines, VEGF, is found at higher levels in the synovial fluid of RA patients than of patients with other arthritides (Fava et al., 1994; Koch et al., 1994; Lee et al., 2001). Additionally serum VEGF levels are elevated in RA compared to osteoarthritis patients, correlate with CRP levels, decrease with successful treatment, and are correlated with radiological progression (Ballara et al., 2001; Harada et al., 1998; Lee et al., 2001). Of the downregulated chemokines discussed CCL2 is known to be produced by synovial fibroblasts and is higher in both the serum and synovium of RA patients compared to healthy controls (Hayashida et al., 2001; Zhang et al., 2015). Also CCL5 is higher in the synovial fluid and serum of RA than OA patients and serum levels correlated with both CRP and ESR measurements (Stanczyk et al., 2005; Yang et al., 2009).

Published studies have also investigated a range of other factors produced during fibroblast:monocyte/macrophage co-culture. Chomarat et al found significant synergistic increases in GM-CSF, LIF, and IL-8 during co-culture of human peripheral blood monocytes and RA synovial fibroblasts but also noted that increases in IL-10, both IL-1α and IL-1β, and TNFα were only additive indicating no modulation of these cytokines (Chomarat et al., 1995). In co-cultures of murine 3T3 fibroblasts and RAW264.7 or murine bone-marrow derived macrophages similar results to the work in this thesis were seen (Holt, Chamberlain and

Grainger, 2010). The cytokines IL-2, IL-3, IL-4, IL-5, IL-9, IL-10, IL12p70, IL-13, IFNγ, and GM-CSF were all expressed at levels lower than 40pg/ml in co-cultures and so were not investigated by Holt, Chamberlain and Grainger. Co-cultures of the murine fibroblast cell line CL.7 with RAW264.7 cells results in an upregulation of CCL3 which is dependent on ICAM-1 interactions indicating that for some factors cell-cell contact may be required (Steinhauser et al., 1998). Another study successfully highlighted the importance of different aspects of the co-culture for the production of different chemokines as in human lung fibroblast co-cultures with peripheral blood monocytes both CCL3 and CCL2 were synergistically increased by co-culture however the CCL3 production can be inhibited by fixing the monocytes prior to co-culture or with antibodies targeting β3-integrins whereas the CCL2 increase can be inhibited by fixing the fibroblasts or using an anti-TNFα antibody (Zickus et al., 1998). Monocytes are not the only cell types to elicit responses during co-cultures with fibroblasts as fibroblast co-cultures with natural killer cells have been shown to increase IL-6, IL-8, GM-CSF, CCL2, TNFα, IL-5, IL-15, granulocyte-colony stimulating factor (G-CSF), CXCL10, MMP9, MMP13, and MMP3 (Chan et al., 2008).

As no changes in extracellular cytokine levels between co-cultures with different fibroblast lines were discovered a different approach was used to assess the intracellular levels of cytokines based on the reasoning that cells in unstimulated co-cultures may not release the full spectrum of cytokines available to them and that fibroblasts from different disease lines may ‘prime’ monocytes by changing levels of cytokines maintained in intracellular stores. However, neither TNFα nor IL-10 were detected at levels significantly different from the isotype control in

monocyte samples and IL-10 was detected at similar levels in both FLS mono- and co-culture indicating TNFα was not being produced by either cell type and IL-10 may have been produced by FLS but levels were not modulated by co-culture.

The data so far indicated that co-culture of fibroblasts from different disease outcome groups with monocytes did not result in differential regulation of cytokine and chemokine release but all elicit a similar response. Regardless it is still of interest to investigate the mechanisms which facilitate this response. ICAM-1 and CD18 are known to bind each other and are expressed on fibroblasts and monocytes respectively (Campbell et al., 1998; Liu and Piela-Smith, 2000; Yang et al., 2010; Zhou et al., 1994). To test if interactions between ICAM-1 and CD18 are involved in the interactions during co-culture a CD18 blocking antibody known to prevent interactions was added at the start of 5 day co-cultures and IL-6 levels in the co-culture medium were measured at the end of co-culture. The results of this work are inconclusive as in direct co-cultures the blocking antibody appears to reduce the IL-6 release but this reduction is also seen when the isotype control antibody is used which hints that the effect of the blocking antibody is non-specific. In the separate co-cultures, which behaved similarly to the direct co-cultures with regards to IL-6 release, neither the blocking antibody nor the isotype control had any impact upon IL-6 levels in the co-culture, further complicating the picture. Finally in the transwell co-culture blocking CD18 interactions led to an increase in IL-6 release and the isotype control antibody reduced levels. Statistical analysis was unable to confirm or refute the findings due to low sample numbers and inadequate power. These data are difficult to interpret and draw any

conclusions from however several improvements could aid in generating clearer data. As most of the IL-6 release occurs within 24 hours of co-culture initiation it would be pertinent to analyse IL-6 levels after 24 hours of co-culture with the blocking antibody rather than the full five days. This would also reduce the chance that any effects may normalise and be missed due to clearance or degradation of the antibody as could happen in the 5 day co-cultures. If the 5 day co-cultures must be used in this assay then regular additions of the antibody to the co-culture during the 5 day period may be required.

Returning to the hypothesis that fibroblasts from different disease outcome groups differentially modulate monocyte/macrophage phenotype it was decided to test the capability of fibroblasts to skew macrophage differentiation in pro- or anti-inflammatory directions depending upon the disease group of the fibroblast line. Monocytes were cultured in the presence of either GM-CSF or M-CSF for 2 days before co-culturing the monocytes with fibroblasts in separated co-cultures for a further 2 days still in the presence of the growth factors. Changes in the expression of genes associated with pro-inflammatory (*TNF* and *STAT1*) or anti-inflammatory (*IL-10*) macrophage functions were investigated in the macrophages. The expression of *TNF* was reduced in all co-cultured monocytes with both GM-CSF and M-CSF relative to monocyte monocultures with the factors apart from in co-cultures of monocytes with resolving fibroblasts in the presence of GM-CSF in which expression was not altered from the monocultures. *IL-10* showed the opposite pattern of regulation with a slight increase in all co-cultures and *STAT1* demonstrated a decrease in all GM-CSF co-cultured monocytes, with those co-cultured with resolving fibroblasts showing a larger decrease, and an

increase in all M-CSF co-cultures. Although the data show that fibroblasts can influence gene expression during the differentiation process no clear differences between disease groupings can be seen with exception of co-culture of GM-CSF differentiating monocytes with resolving fibroblasts which appears to downregulate *STAT1* to a larger extent than other co-cultures but does not achieve statistical significance.

Published work has demonstrated that fibroblasts are capable of influencing the monocyte differentiation process, as may be evidenced herein, by the modulation of gene expression compared to the differentiating monocytes alone. Differentiation of human peripheral blood monocytes to dendritic cells using a combination of GM-CSF and IL-4 can be shifted to a monocyte to macrophage differentiation process by co-culture with dermal fibroblasts (Chomarat et al., 2000). This shift was assessed using CD14 expression as a monocyte/macrophage marker and CD1a expression as a DC marker and required cell-cell contact, fibroblast derived IL-6, and also M-CSF to facilitate the change. Also of note is that the modulation of STAT1 expression by fibroblasts may be relevant to RA pathogenesis as STAT1 levels are higher in monocytes found in the synovial fluid of RA patients than those found in the peripheral blood of the same patients (Isomäki et al., 2015).

In summary the work in this section established a co-culture protocol that was used to assess the role of contact dependent and independent mechanisms in fibroblast:monocyte co-cultures. Co-culture resulted in increased IL-6, CXCL10, VEGF, and IFNγ levels and decreased CCL2 and CCL5 levels in the co-culture

supernatant however fibroblasts isolated from different BEACON disease outcome groups did not differentially modulate these factors i.e. all fibroblast lines elicited a similar response. The release of IL-6 during co-culture was not contact dependent however inducing changes in the cytoskeleton in fibroblasts may change IL-6 release as evidenced by co-cultures in which fibroblasts were suspended from a membrane. Fibroblasts were able to modulate the expression of genes during monocyte to macrophage differentiation however they did not display differential modulation dependent on the clinical outcome group from which the fibroblasts were taken. Therefore there is a lack of evidence to support the alternative hypothesis and as such the null hypothesis, that fibroblasts isolated from different stages of RA progression do not differentially modulate monocyte or macrophage phenotype *in vitro*, was accepted.

Although the co-culture of fibroblasts and monocytes did not result in clear differential modulation dependent on the stage of RA from which the fibroblasts were isolated the variation that was observed is of interest in the context of RA. IL-6 is a pro-inflammatory cytokine known to be highly involved in RA demonstrated by the successful targeting of IL-6 signalling as a therapy for RA patients and also the importance of IL-6 in the murine collagen-induced arthritis model (Alonzi et al., 1998; Castell et al., 1988; Choy et al., 2002; Sasai et al., 1999). The interactions between monocytes and fibroblasts in the synovium of RA patients may contribute to the inflammatory burden through upregulation of IL-6 release along with increased release of the T cell chemoattractant CXCL10 increasing the involvement of the adaptive arm of the immune system in the disease (Dufour et al., 2002a). Additionally secretion of VEGF was upregulated

by co-culture which indicates that monocyte-fibroblast interactions could increase vascularisation of the synovium facilitating increased inflammatory infiltrate into afflicted joints. However, the modulation of secreted factors was not a blanket increase as the monocyte chemoattractants CCL2 and CCL5 decreased in concentration during co-culture which may demonstrate an feedback loop employed by monocytes to prevent an unrestrained influx of monocytes into inflammatory sites (Rollins, Walz and Baggiolini, 1991; Schall et al., 1990). In contrast to the pro-inflammatory effects of fibroblast-monocyte co-culture fibroblasts appear to dampen the differentiation of macrophages towards a pro-inflammatory phenotype. Regardless of the fibroblast disease stage or macrophage differentiation factor (GM-CSF or M-CSF) co-culture resulted in lower levels of *TNF* mRNA and higher levels or *IL10* mRNA in macrophages than in macrophages differentiated in monoculture. This highlights the complexity of fibroblast interactions with other cell types, although the co-cultures with monocytes seemingly resulted in a pro-inflammatory phenotype fibroblasts are capable of downregulating expression of the critical factor for RA, *TNF*, and increasing mRNA levels of the anti-inflammatory factor *IL10*.

The lack of any significant differences in monocyte or macrophage phenotype is particulary interesting given that large numbers of differentially expressed genes were detected between fibroblasts cultured from different stages of RA in chapter 2. It could be that the differences in gene expression are somewhat normalised through post-transcriptional methods leading to similar levels of protein expression which in turn has a similar impact on monocytes regardless of the fibroblast disease stage and it is known that the correlation between mRNA and

protein levels is complex (Edfors et al., 2016; Koussounadis et al., 2015; Liu, Beyer and Aebersold, 2016). Regardless, the lack of differential modulation can be interpreted as the interactions between fibroblasts and monocytes not being an initiating factor for RA. However, the work in this chapter finds limited evidence that fibroblasts from resolving arthritis may actually induce lower levels of IL-6 in co-cultures with monocytes and supress *STAT1* expression during GM-CSF driven monocyte-to-macrophage differentiation to a higher degree than fibroblasts from very early or late stage RA demonstrating that resolving fibroblasts may indeed regulate the level of pro-inflammatory factors and macrophages within the joint. As monocyte/macrophage populations derived from peripheral blood did not appear to be differentially modulated by fibroblasts from different stages of RA it was decided to turn the investigation towards synovial populations of macrophages to eliminate the confounding factor of using peripheral blood derived cells.

4 IDENTIFICATION AND ISOLATION OF SYNOVIAL FIBROBLAST AND MACROPHAGE POPULATIONS

4.1 Introduction

To address the third hypothesis, that different fibroblast and macrophage subsets exist within the synovium with different distinct functions, two approaches were used, immunofluorescent staining of tissue sections and enzymatic digestion of synovial biopsy material. As mentioned in section 1.5.5 different subsets of fibroblasts are thought to exist in the synovium, particularly between the lining and sublining layers. In addition macrophages are also thought to exist as different subsets in the synovium, possibly identified by either lining or sublining location or CD163 expression, as discussed in section 1.4.3.2, and may differentiate from infiltrating or tissue-resident populations (section 1.4.2.2). The expression of stromal cell markers in synovial tissue from the BEACON cohort were investigated along with possible macrophage markers. However, to investigate function and gene expression in synovial stromal and macrophage populations one must be able to isolate viable cells from the tissue and as such a protocol for the digestion of synovial tissue was developed and used to facilitate the flow cytometric analysis of synovial cell populations.

Protocols for the isolation of cell suspensions from the synovium exist for both human and mouse tissue. In mouse tissues the protocols have a degree of standardisation presumably due to the difficultly in accessing the small amount of synovial tissue. Protocols typically include microdissection stages where all tissue is removed from bones but leaving the joint capsule intact and then, upon exposing the synovium by cutting the joint capsule, digestion with collagenase IV and in some case deoxyribonuclease I (DNaseI) (Armaka et al., 2009; Hardy et

al., 2013; Zhao et al., 2016). In humans, the published protocols are more varied with tissue source, arthroplasty material or biopsy, and enzyme used for digestion changing between publications e.g. hyaluronidase and collagenase IV, 0.2% collagenase, or clostridiopeptidase A (Chang et al., 2013; Dayer et al., 1976; Gullo and De Bari, 2013; Van Landuyt et al., 2010). Many collagenase preparations are derived from bacteria and may not have been adequately purified to remove endotoxin contamination. Working in collaboration with other members of the National Institutes of Health (NIH) accelerating medicines partnership (AMP) (https://www.nih.gov/research-training/accelerating-medicines-partnership-amp) this work aimed to develop a unified digestion protocol using a widely available enzyme mix with low endotoxin and high collagenase activity.

The aim of the work in this chapter was to identify fibroblast and macrophage subpopulations using both tissue sections and dissagregated biopsy material. To address this aim three questions will be addressed:

1. Does the expression of markers identifying synovial fibroblasts vary depending on the stage of RA?
2. Can subpopulations of macrophages be identified in the synovium?
3. Can a protocol be developed to allow the isolation of pure fibroblasts and macrophages from synovial tissue?

4.2 Methods and materials

4.2.1 Arthroplasty, biopsy material and clinical outcome groups

Synovial tissue from patients from the BEACON clinical outcome groups was collected by ultrasound guided biopsy. Patients donating tissue samples collected by ultrasound guided biopsy in Birmingham gave written, informed consent under the following ethics: West Midlands Black Country REC Reference 07/H1203/57 “Ultrasound guided synovial biopsy in patients with arthritis”. Tissue from patients in the Jrep group or from osteoarthritis patients was collected during arthroplasty at the Royal Orthopaedic Hospital NHS Foundation Trust (Birmingham, UK) under the following ethics: West Midlands Black Country REC Reference 07/H1204/191 “Origins and roles of stromal cells in chronic inflammatory arthritis”. To increase the number of samples for stromal marker staining, available tissue from the Dutch Synoviomics cohort, collected by mini-arthroscopy, were included using the same criteria as for the BEACON samples (De Hair et al., 2011). Characterisitics of the patients from which tissues sections used for the quantification of stromal markers are presented in Table 19.

ID	Group	Disease classification	CCP +/-	RF +/-	Sex	Age	DAS28 ESR baseline	ESR (mm/hr)	CRP (mg/l)
58	Res	UA	NA	NA	M	57	NA	15	19
167	EstRA	RA	-	-	F	58	NA	44	10
389	EstRA	RA	-	-	F	18	3.8	102	5
390	Res	Gout	-	-	F	69	3.95	38	24
394	EstRA	RA	+	+	F	68	4.5	25	8
573	Res	Gout	-	-	M	44	3.72	101	88
587	EstRA	RA	-	-	M	52	6.05	20	3
676	VeRA	RA	+	+	M	59	2.42	8	5
760	EstRA	RA	-	-	M	74	4.07	8	12
790	EstRA	RA	+	+	M	55	6.4	18	16
1181	VeRA	RA	-	-	F	22	7.33	111	113.8
1182	VeRA	RA	-	-	F	43	5.7	35	4.6
1190	VeRA	RA	-	+	F	55	5.37	37	4.2
1192	VeRA	RA	-	-	F	54	6.7	76	42.5
1258	Res	Gout	-	-	M	61	3.07	8	3.6
1262	Res	Gout	-	-	M	66	4.58	36	20.1
1264	Res	UA	-	-	F	32	3.49	13	6.7
1282	VeRA	RA	+	+	F	54	5.2	32	16.5
1299	VeRA	RA	+	+	M	60	1.81	3	1.7
1403	EstRA	RA	-	-	F	72	4.48	5	3
1457	EstRA	RA	-	-	F	66	4.38	13	1.8
1515	Res	PsA	-	-	F	40	2.99	11	1
1543	VeRA	RA	+	+	M	33			
1554	VeRA	RA	+	+	M	38	3.62	19	6.4
1556	Res	PsA	-	-	M	65	4.24	15	1
1611	VeRA	RA	+	+	M	74	5.35	68	79.9
1717	VeRA	RA	-	-	F	82	5.1	8	6.3
1732	VeRA	RA	+	+	F	48	2.92	5	2.1
Bx003	VeRA	RA	+	+	M	50	5.7	31	26

Bx004	Res	UA	-	-	M	33	3.6	9	6
BX016	RA	RA	+	+	M	46	6.7	34	7
BX018	RA	RA	-	-	F	69	4.6	11	0
BX024	Res	Reactive	-	-	M	32	2.9	10	10
BX028	Res	RA	-	-	M	74	4.8	45	13
BX032	RA	RA	+	-	F	46	5.2	24	32
BX038	Res	Parovirus	-	-	F	45	4	4	0
BX042	VeRA	RA	+	-	M	55	3.5	58	45
BX048	Res	Reactive/post strep	-	-	M	35	4.1	51	7
BX054	Res	UA	-	-	M	55	3.5	2	6
BX063	VeRA	RA	+	-	F	74	4.4	20	32
BX064	Res	UA	-	-	M	35	1.6	2	9
BX065	Res	UA	-	-	F	37	3.3	7	0
BX070	Norm	Normal	-	-	M	44	NA	NA	NA
BX071	Res	Parovirus	-	-	F	41	1.7	5	9
BX075	RA	RA	-	-	F	22	6.4	81	79
BX076	Res	Reactive	-	-	M	28	4.5	18	8
BX077	RA	RA	-	-	M	72	7.4	53	43
BX082	Norm	Normal	-	-	F	49	NA	NA	NA
BX083	Norm	Normal	-	-	M	42	NA	NA	NA
BX084	VeRA	RA	+	+	M	49	6.8	25	18
BX089	Norm	Normal	-	-	F	38	NA	NA	NA
BX092	VeRA	RA	+	+	M	48	6	63	38
BX093	RA	RA	+	+	F	65	5.3	72	81
BX097	Norm	Normal	-	-	F	41	NA	NA	NA

Table 19: Clinical characteristics of the patients from which samples used for the quantification of stromal markers originated. NA = Not available, DAS28 = Disease activity 28 score, ESR = Erythrocyte sedimentation rate, CRP = C-reactive protein, CCP = Anti-cyclic citrullinated peptide, RF = Rheumatoid factor, UA = Undifferentiated arthritis.

4.2.2 Preparation of synovial tissue sections

Frozen tissue sections used for immunofluorescent staining of synovial cellular markers were cut from frozen tissue blocks by Holly Adams and Jennifer Marshall as follows. Ultrasound guided biopsy material was immediately mounted in Tissue-Tek® O.C.T™ compound (Sakura) and frozen in the gas-phase of liquid nitrogen. The resulting tissue block was cut into 6µm sections on a cryostat and the sections were subsequently acetone fixed and permeabilised at 4°C in histological grade acetone (J.T.Baker) for 20 minutes. Tissue sections were stored at -20°C until used in staining protocols.

4.2.3 Immunofluorescent staining of fibroblast markers in synovial tissue sections

Staining and imaging of tissue sections was undertaken as described in the supplementary methods of Misharin et al. (2014). Briefly tissue sections were blocked with PBS v/v 10% FBS for 10 minutes to reduce non-specific binding of antibodies. The primary antibodies listed in Table 20 were then used to stain the tissues overnight at 4°C before washing the slides in PBS and staining for 1 hour at room temperature using the secondary antibodies in Table 21. A nuclear counterstain was then added in the form of 4µg/ml Hoechst 33258, Pentahydrate (bis-Benzimide) for 20 minutes (Invitrogen) before washing and mounting slides in 2.4% w/v 1,4-diazabicyclo (2,2,2) octane (Sigma) in 90% v/v glycerol (Sigma, 49767) in PBS (pH, 8.6). The conjugates FITC, Cy3, and Cy5 were excited with a 488nm argon laser, a 543nm HeNe laser, and a 633nm helium laser respectively and images were acquired on a LSM 510 confocal microscope using LSM 510 version 3.2 SP2 software (Carl Zeiss). Negative controls were stained

with secondary antibodies only. Each section contained 6-8 biopsy pieces and staining for each of the markers was assessed by calculating the number of pixels which have intensity measurements above the background and dividing this value by the pixel area of the region being assessed (giving the units pixels/unit area). Three batches of staining were completed with samples restained in each batch to account for batch effects. Analysis was completed by myself, Debbie Hardie, Andrew Filer, and Jennifer Marshall.

Target	Clone	Isotype	Supplier
CD90	F15-42-1	Mouse IgG1	Chemicon
FAP	F11-24	Mouse IgG1	eBioscience
GP38	D2-40	Mouse IgG1	AbD Serotec
CD68	Y1/82A	Mouse IgG2b	BD Pharmingen

Table 20: Primary antibodies used for fibroblast/macrophage staining in synovial tissue sections

Target	Conjugate	Isotype	Supplier
Mouse IgG1	FITC	Goat IgG	SouthernBiotech
Mouse IgG2b	Cy5	Goat IgG	SouthernBiotech

Table 21: Secondary antibodies used for fibroblast/macrophage staining in synovial tissue sections

4.2.4 Immunofluorescent staining of macrophage markers in synovial tissue sections

Staining was completed as in section 4.2.3 but only on a small number of tissues using the primary antibodies listed in Table 22 and the secondary antibodies in Table 23.

Target	Clone	Isotype	Supplier
CD163	EDHu-1	Mouse IgG1	AbD Serotec
CD206	15-2	Mouse IgG1, κ	Biolegend
CD68	Y1/82A	Mouse IgG2b, κ	BD Biosciences
HLA-DR	L243	Mouse IgG2a, κ	BD Biosciences

Table 22: Primary antibodies used for macrophage staining in synovial tissue sections

Target	Conjugate	Isotype	Supplier
Mouse IgG1	FITC	Goat IgG	SouthernBiotech
Mouse IgG2a	TRITC	Goat IgG	SouthernBiotech
Mouse IgG2b	Cy5	Goat IgG	SouthernBiotech

Table 23: Secondary antibodies used for macrophage staining in synovial tissue sections

4.2.5 Digest optimisation

Several parameters were evaluated in order to establish an optimal digestion protocol and to assess the feasibility of cryopreservation of tissue to facilitate later digestion or shipment of synovial tissue samples between sites. An additional step of trituration in the digestion protocol was suggested by collaborators in the Brenner Lab (Brigham and Women's Hospital, Boston). The final established protocol was as follows. Synovial tissue, either obtained during arthroscopy and provided without any buffer or obtained by ultrasound guided biopsy procedures and provided in saline solution, was minced using scalpels until the average size of pieces was less than $1mm^3$. Tissue was then transferred into 5ml polypropylene tubes (Falcon) containing a magnetic stirring bar at a maximum quantity of 0.02g of tissue/tube. 1ml of digestion buffer consisting of serum-free RPMI-1640 (Sigma), 0.97mM sodium orthopyruvate (Sigma), 1.94mM glutamine (Sigma), 97U/ml penicillin (Sigma), 97µg/ml streptomycin (Sigma), [0.97x] MEM non-essential amino acids (Sigma), 50µg/ml Liberase TM (Roche), and 40µg/ml DNaseI (Sigma) was added to each 5ml tube followed by immediate incubation in a 37°C water bath with stirrer agitation set to 380rpm. After 15 minutes of incubation the digestion buffer and tissue was triturated using a 16G needle before continuing the incubation and agitation at 37°C for another 15 minutes. The digestion mix was then passed through a 100µm filter (Greiner Bio-One) to remove any remaining pieces of tissue and the filter was rinsed with an equivalent

volume of complete fibroblast medium containing serum (see section 2.2.1) to the volume of digestion mix in order to inhibit the enzymatic digestion. The suspension was then centrifuged at 300g for 6 minutes before discarding the supernatant and resuspending in PBS (Oxoid). The cell suspension was centrifuged again at 300g for 6 minutes before resuspending in PBS containing LIVE/DEAD® Fixable Near-IR Dead Cell Stain Kit (Thermo Fisher Scientific) at the manufacturers recommended concentration and incubating at 4^{o}C for 30 minutes. The sample was then washed twice by centrifuging at 300g for 6 minutes before adding 200µl of staining buffer (PBS 2% v/v heat-inactivated foetal bovine serum (Biosera) and 2mM Ethylenediaminetetraacetic acid (EDTA, Sigma)), centrifuging again at 300g for 6 minutes, and repeating the previous two steps. The cells were then resuspended in staining buffer containing an antibody mix consisting of several or all of the antibodies described in Table 24. Single colour controls of all antibodies in Table 24 were prepared at the same concentration as the full mix in staining buffer and added to VersaComp Antibody Capture Beads (Beckman Coulter) used as specified by the manufacturer's instructions. Both the cell suspension and single colour controls were incubated at 4^{o}C for 30 minutes before washing all stained cells and beads as before. Samples were resuspended in 200µl of staining buffer before passing the cell suspension through a 70µm filter (Sysmex) and rinsing the filter with 300µl of staining buffer. The cell suspension was then sorted into 6 populations using a MoFloAstrios EQ (Beckman Coulter).

Target	Conjugate	Clone	Isotype	Supplier
CD1c	FITC	L161	Mouse IgG1, κ	Biolegend
CD11c	BV650	3.9	Mouse IgG1, κ	Biolegend
CD14	BV785	M5E2	Mouse IgG2a	Biolegend
CD15	BV510	W6D3	Mouse IgG1, κ	Biolegend
CD31	FITC	TP1/15	Mouse IgG2a	Abcam
CD31	PerCP	9G11	Mouse IgG1	R&D Systems
CD34	FITC	581	Mouse IgG1, κ	Biolegend
CD45	V450	2D1	Mouse IgG1, κ	BD Biosciences
CD64	BV605	10.1	Mouse IgG1, κ	Biolegend
CD90	PerCP-Cy5.5	eBio5E10	Mouse IgG1, κ	eBioscience
CD117	FITC/PE-Cy7	104D2	Mouse IgG1, κ	Biolegend
CD163	PECF594	GHI/61	Mouse IgG1, κ	BD Biosciences
CD197	BV711	G043H7	Mouse IgG2a, κ	Biolegend
CD206	APC	15-2	Mouse IgG1, κ	Biolegend
CD235a	FITC/PE-Cy7	HI264	Mouse IgG2a, κ	Biolegend
CD326	FITC/PE-Cy7	9C4	Mouse IgG2b, κ	Biolegend
GP38	PE	NZ-1.3	Rat IgG2a	eBioscience
HLA-DR	BV510	L243	Mouse IgG2a, κ	Biolegend
Mer	Alexa Fluor 350	125518	Mouse	Novus Biologicals

Table 24: Antibodies used for fibroblast/macrophage staining in synovial tissue digestions

4.3 Results

4.3.1 Fibroblasts and macrophages can be identified by immunofluorescent staining of tissue and the expression of fibroblast markers varies with disease stage

The exact phenotype of fibroblast subsets in the synovium during RA is unclear with recent publications only beginning to address this question. However, it was of interest to investigate the expression of the known stromal markers, CD90, FAP, and GP38, in synovial tissue from different clinical outcome groups to assess if the markers are differentially expressed between disease stages, particulary given the differential expression of genes encoding these markers in cultured synovial fibroblasts from different disease stages (2.3.2). To this end synovial tissue sections obtained from patients at different stages of RA were stained for several prospective stromal subset markers as well as the macrophage marker CD68 and the levels of markers were quantified and compared between different disease stages the results of which formed part of a publication (Choi et al., 2017). Of the stromal markers assessed FAP showed a significant increase in expression in very early RA compared to normal, uninflamed controls (Figure 41). GP38 demonstrated a similar pattern to FAP expression but differences were not statistically significant. CD90 expression was significantly higher in the resolving arthritis samples compared to the normal samples and CD68 remained unchanged across all clinical outcome groups.

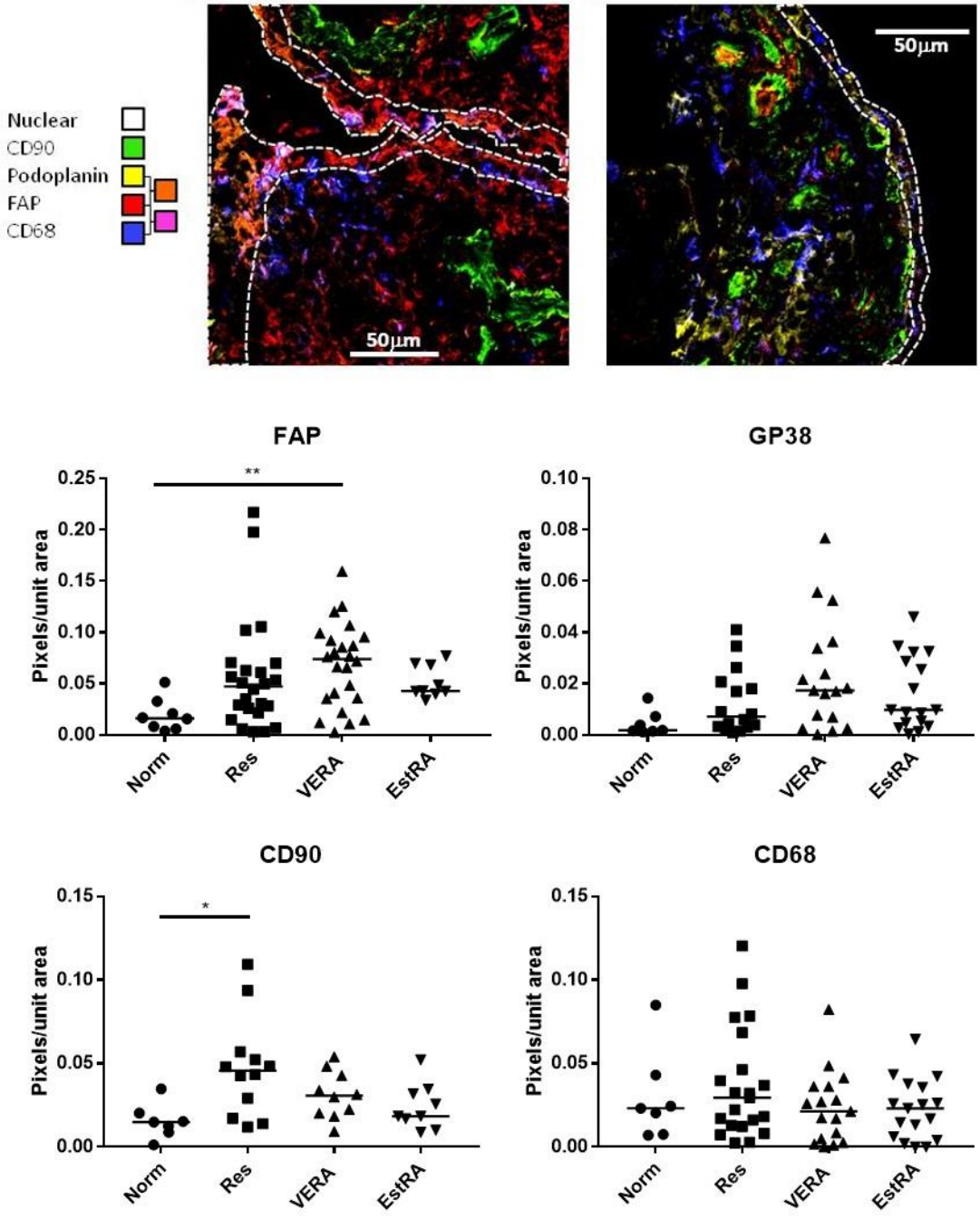


Figure 41: Immunofluorescent staining and quantification of fibroblast subsets and macrophages in the synovium of uninflamed control patients (Norm), patients with resolving arthritis (Res), patients with very early RA (VERA), and patients with established RA (RA). Representative images of staining of VERA tissue with regions of interest (ROIs) drawn around the lining layer present in the image (adapted from (Choi et al., 2017)). Quantification of staining in the tissue found significantly higher levels of FAP in VERA samples compared to the Norm samples and significantly higher CD90 expression in the Res samples than the Norm samples. Tested using Kruskal-Wallis tests and Dunn's multiple comparisons tests. *= p<0.05, ** = p<0.01. Representative immunofluorescence images are from a patient with RA (left) and a patient with resolving arthritis (Res).

In order to identify macrophages within the synovium, synovial tissue sections were stained for CD68, CD163, CD206 and HLA-DR, markers reported to differentiate between infiltrating and resident macrophages in the literature, and imaged used confocal microscopy, the results of which formed part of a publication (Misharin et al., 2014). Two populations of $CD68^{+}CD163^{+}$ macrophages could be discerned using this method, a $HLA\text{-}DR^{hi}CD163^{lo}$ population hypothesised to arise from infiltrating cells and a $HLA\text{-}DR^{lo}CD163^{hi}$ population hypothesised to arise from tissue resident cells. In addition a $CD68^{+}HLA\text{-}DR^{+}CD206^{+}$ population could also be identified in the tissue.

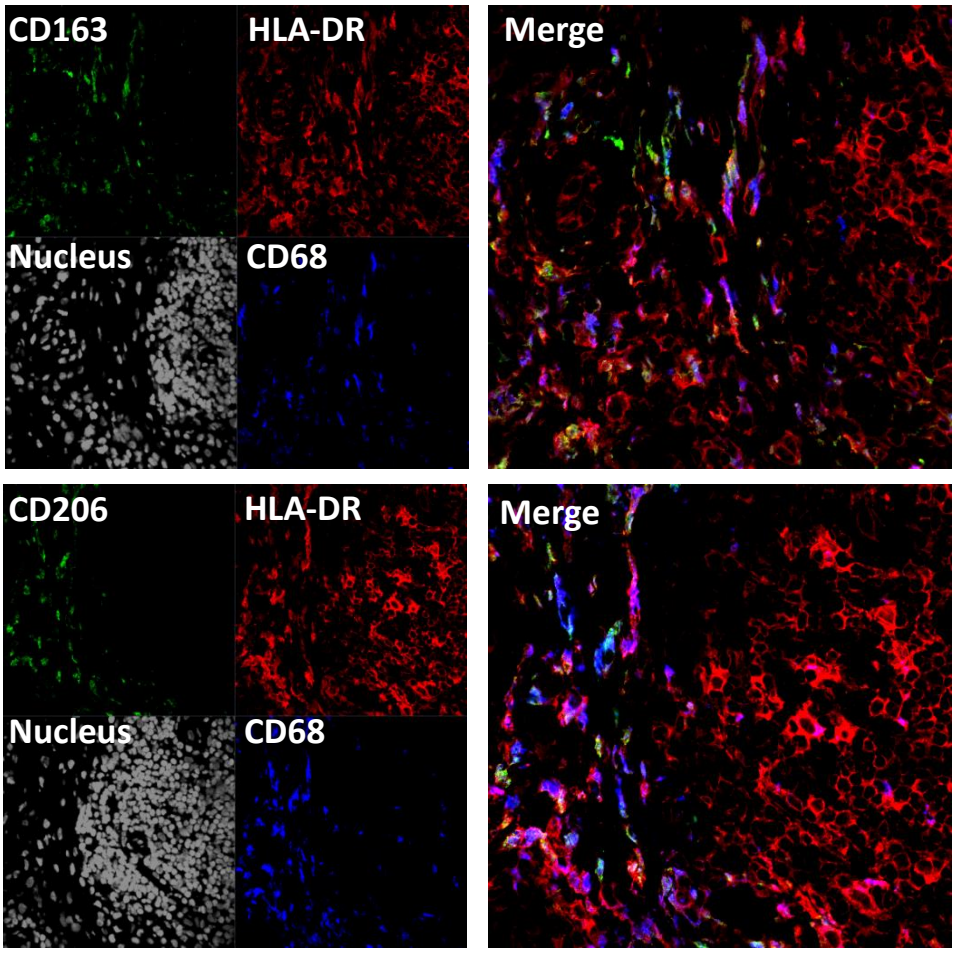


Figure 42: Representative images demonstrating the presence of $CD68^{+}HLA\text{-}DR^{+}CD163^{+}$ and $CD68^{+}HLA\text{-}DR^{+}CD206^{+}$ populations in joint replacement RA synovial tissue.

Due to the limited dynamic range of confocal microscopy it was technically challenging to image the synovial tissue populations using the same gain settings for the photomultiplier tube. This was due to the differences in expression of all the markers varying greatly between the populations with even CD68 expression being lower on the 'resident' population. Therefore quantification of imaging data for the macrophage populations was not attempted, however a protocol for the dissociation of synovial biopsy material and flow cytometric assessment of fibroblast and macrophage populations was developed.

4.3.2 Development of a synovial digestion protocol optimised for recovery of cellular subpopulations

In order to further investigate synovial fibroblast and macrophage populations a tissue digestion protocol was developed in collaboration with members of the NIH AMP to be used on arthroplasty and synovial biopsy tissue. Synovial tissue was minced using scalpels and two commercially available enzyme mixes, Liberase TM and Liberase TL which contain a medium or low concentration of thermolysin respectively, were used to digest the tissue for 30 minutes at 37^{o}C with agitation with a magnetic stirrer with the enzymes used at a concentration of 50µg/ml and with DNaseI present in both digestions at a concentration of 40µg/ml. The resulting stromal populations were investigated by flow cytometry by investigating CD90 and GP38 expression in the $CD31^{-}CD45^{-}$ population. Liberase TM appears more efficient at isolating the $CD90^{-}GP38^{+}$ and $CD90^{+}GP38^{-}$ populations than Liberase TL but not significantly (Figure 43). Neither digestion method appears to isolate $CD31^{-}CD45^{-}$ stromal or $CD45^{+}$ hematopoietic populations in different proportions relative to each other indicating a similar efficacy at isolating overall stromal and hematopoietic populations. Perhaps most important is that although

Liberase TM has a higher concentration of thermolysin it does not decrease the viability of the cells isolated any more than that obtained by the liberase TL preparation with a mean viable proportion of the entire cell yield and 95% confidence interval of 76.5 ± 30.8 and 74.4 ± 29.3 for the liberase TM or liberase TL respectively.

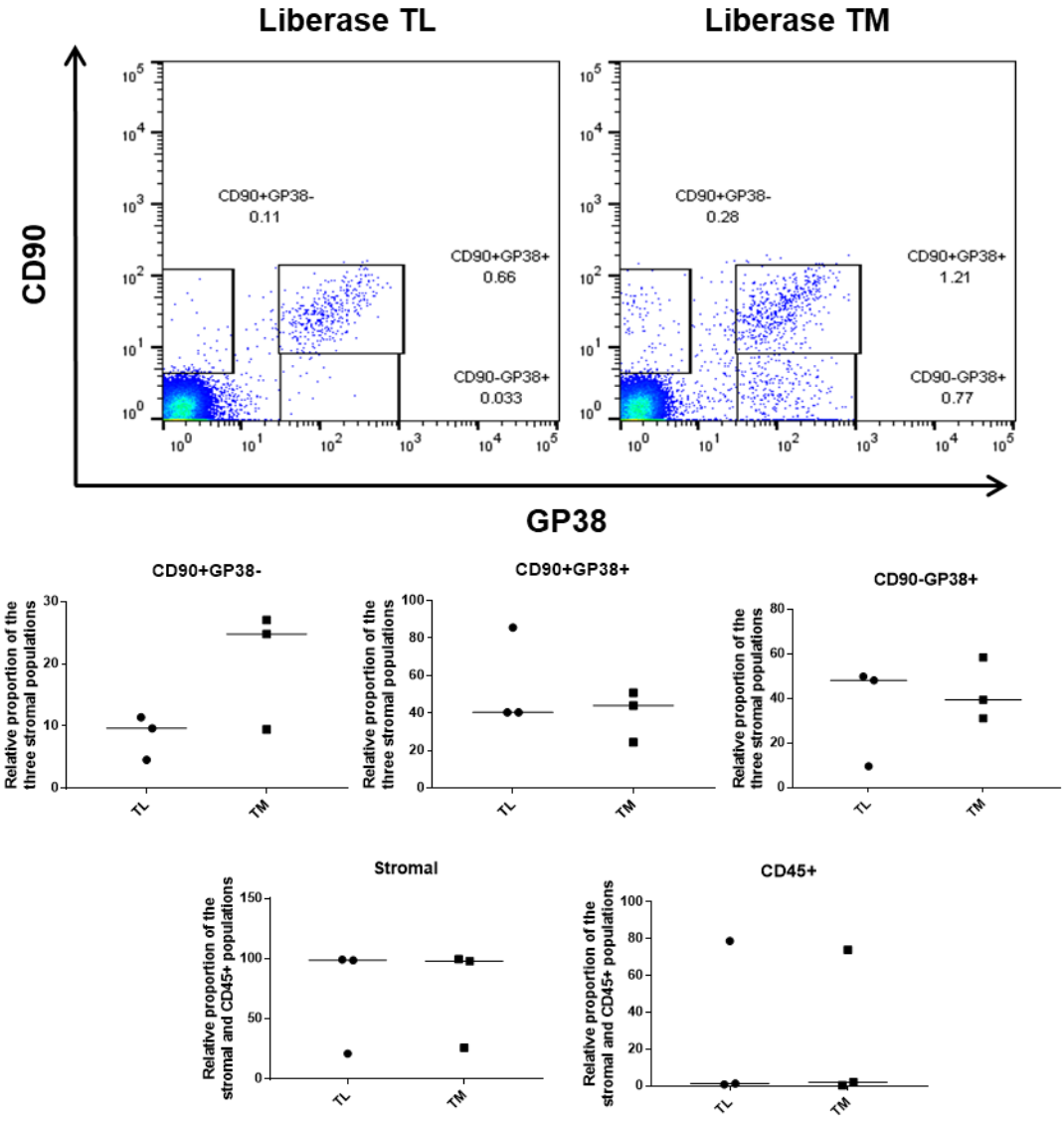


Figure 43: Comparison of digestion of synovial tissue using either Liberase TM or Liberase TL at 50µg/ml, with 40µg/ml DNaseI at 37ºC for 30 minutes with agitation by a mechanical stirrer. Liberase TM appeared to isolate higher proportions of the $CD31^{-}CD45^{-}CD90^{+}GP38^{-}$ and the $CD31^{-}CD45^{-}CD90^{-}GP38^{+}$ populations than liberase TL but not significantly. Both enzymes isolated similar proportions of stromal and hematopoietic populations, defined as $CD31^{-}CD45^{-}$ or $CD45^{+}$ respectively. n=3 representative flow dot plots shown. Tested using Mann-Whitney test. Bars show median.

In order to minimise the impact of enzymatic digestion upon surface marker expression, mechanical disruption of synovial tissue using a gentleMACS™ Dissociator (MiltenyiBiotec) with gentleMACS C Tubes (MiltenyiBiotec) using the pre-programmed protocol m_spleen0.40 or a Stomacher® 80 Biomaster (Seward) for 5 minutes with or without subsequent digestion using liberase TM were investigated using the stromal populations identified with CD90 and GP38 as the readout. Using only mechanical disruption resulted in a relatively poor isolation of stromal populations with the $CD90^{+}GP38^{+}$ population seeming to be completely absent (Figure 44). Using either mechanical method with a subsequent enzymatic digestion with magnetic stirrer agitation resulted in the isolation of similar stromal populations as were seen in digestions without mechanical disruption by stomaching or GentleMACs.

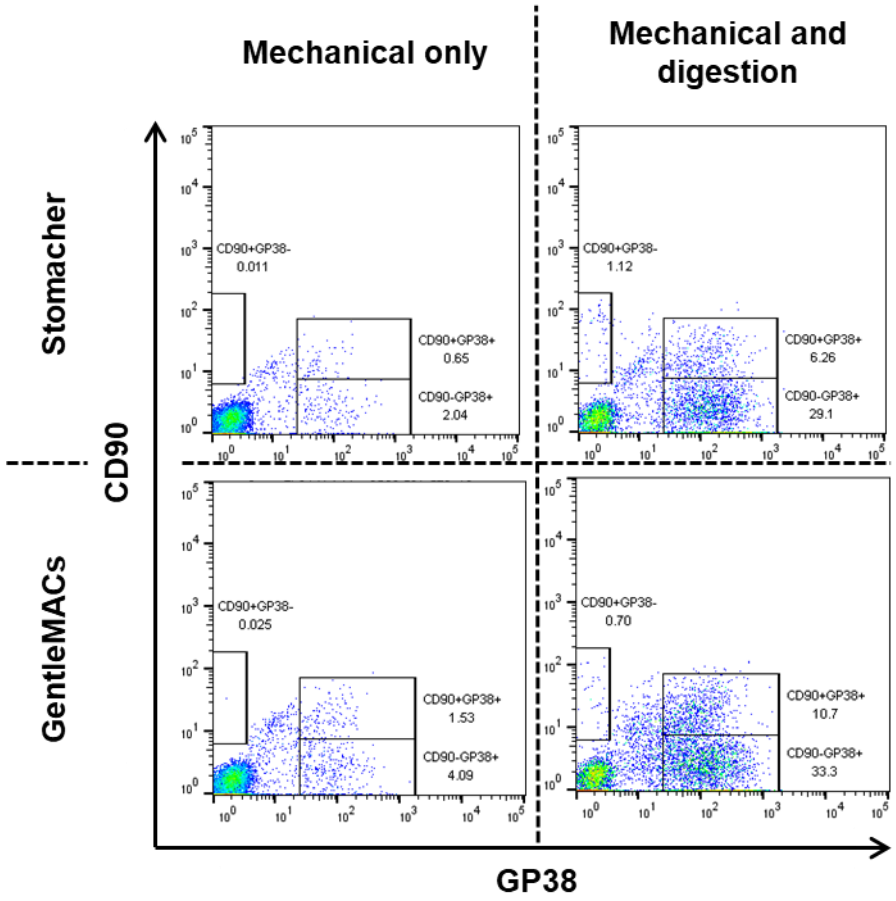


Figure 44: Mechanical digestion by either stomaching or GentleMACs C tube dissociation alone did not efficiently isolate stromal populations. When used as a prior step to enzymatic digestion using Liberase TM (50µg/ml and 40µg/ml DNase I) stomaching increased the mean proportion of $CD90^{+}GP38^{+}$ stromal cells isolated by 35.85% however GentleMACs increased the mean proportions of the $CD90^{+}GP38^{+}$ and the $CD90^{-}GP38^{+}$ stromal populations by 59.73% and 16.21% respectively (n=3, representative plots shown).

As digesting synovial biopsy material just after collection would result in many small experimental batches, possibly introducing unwanted variation into downstream analyses, the ability to freeze tissue and digest several samples in one batch would be advantageous. To this end a digestion protocol using Cryostor 10 was optimised. As the tissue would be frozen prior to digestion there was a risk that the tissue would not be fully perfused by protective freezing media causing the death of cells present in the centre of tissue pieces. To prevent this from occurring freezing protocols allowing the Cryostor 10 to perfuse into the

tissue at 4ºC for either 10 or 20 minutes before transfer to -80ºC were compared with freshly digested tissue with regards to the viability of cells isolated by digestion, cell yield, and proportions of cell populations. As seen in Figure 45 freezing the tissue with a 10 minute perfusion step did not appear to decrease the viability of cells isolated relative to the number of cells isolated from a digestion of the same tissue immediately after collection. The proportion of stromal subsets relative to the total of the CD90-GP38+ and CD90+GP38+ populations did not appear to be consistently altered by the freezing protocols; however when considering the proportions of the stromal and CD45+ populations relative to the total, it appeared that the freezing protocols increased the proportion of stromal cells isolated and hence decreased the proportion of $CD45^+$ cells isolated. It seems unlikely that this change was due to death of $CD45^+$ cells as the overall viability did not appear to be affected using the 10 minute protocol; the CD45+ and stromal proportions suggested that the effect could have been due to an increase in cell yield after freezing.

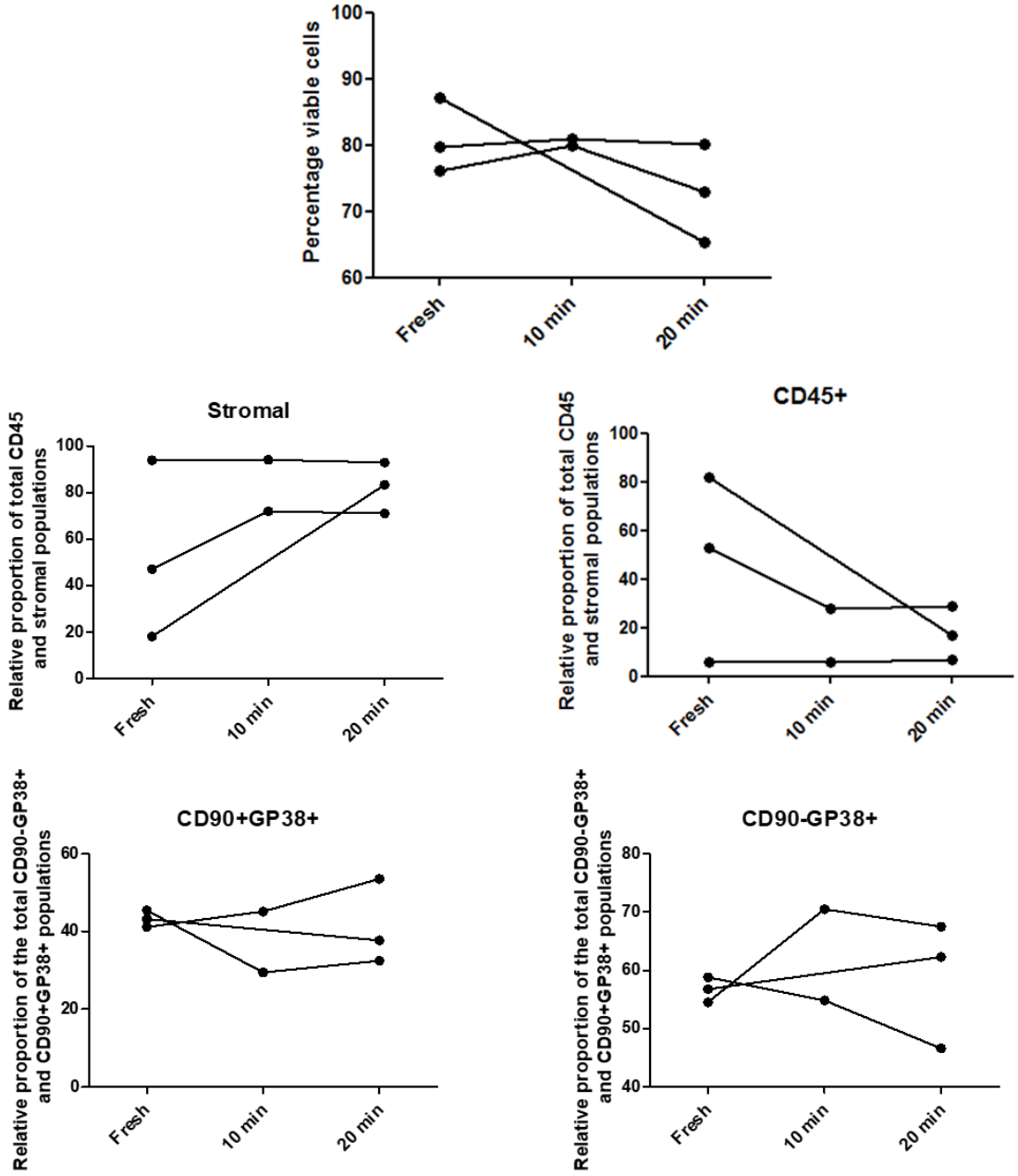


Figure 45: Comparison of the viability and proportion of cell populations isolated by digestion of either fresh synovial tissue or cryopreserved synovial tissue frozen in Cryostor 10 with perfusion at 4ºC for either 10 or 20 minutes prior to freezing at -80ºC. Freezing with 10 minutes of perfusion did not appear to impact cell viability whereas 20 minutes of perfusion did possibly reduce viability. Freezing did not appear to change the proportions of stromal subsets isolated relative to one another. The proportion of stromal cells isolated relative to the total stromal and CD45$^+$ populations could possibly increase after freezing for either 10 or 20 minutes. All digestions were completed using Liberase TL 100ug/ml and 100ug/ml DNaseI. n = 2 for 10 minutes, 3 for fresh and 20 minutes.

With a functional digestion protocol optimised (full protocol in appendix 8.3) the stromal and macrophage populations could now be investigated in greater detail.

4.3.3 Digested macrophage and fibroblast populations can be identified by flow cytometry

Using the protocol developed in section 4.2.5 flow cytometric and *ex vivo* analysis of synovial stromal and macrophage populations could be attempted. The cell surface markers for stromal populations had been previously identified and during this thesis, work on single cell sequencing of synovial cells isolated by digestion from arthroplasty samples was completed by Mizoguchi et al. (2017) which identified three major stromal populations in the $CD31^-$ compartment, $CD90^-GP38^+$ cells, $CD90^+GP38^+CD34^+$ cells, and $CD90^+GP38^+CD34^-$ cells. Using these markers the same populations could be identified in synovial biopsy material digested herein (Figure 51).

The identification of synovial macrophages was less clear, previous work having identified macrophage populations however one of the common markers used, CD68, could not be used in this work as CD68 is predominantly expressed intracellularly and the necessary protocols to stain for this marker would result in reduction of the quality of RNA (Micklem et al., 1989; Parwaresch et al., 1986; Pulford et al., 1989). Using CD163 and CD64 as putative macrophage markers $CD45^+CD64^+CD163^-$ and $CD45^+CD64^+CD163^+$ populations could be identified (Figure 46). One of the caveats of using CD64 as a macrophage marker is that it can also be expressed on activated mast cells, however by using the mast cell marker CD117 these cells could be excluded (Meininger et al., 1992; Woolhiser, Brockow and Metcalfe, 2004). In the digestions no $CD45^+CD64^+CD117^+$

population was found indicating mast cells were not present in the $CD45^+CD64^+$ cell fraction (Figure 46, B). Using the $CD45^+CD163^+$ approach to identify possible macrophages, an additional $CD45^+CD163^{lo}GP38^+$ population could be found.

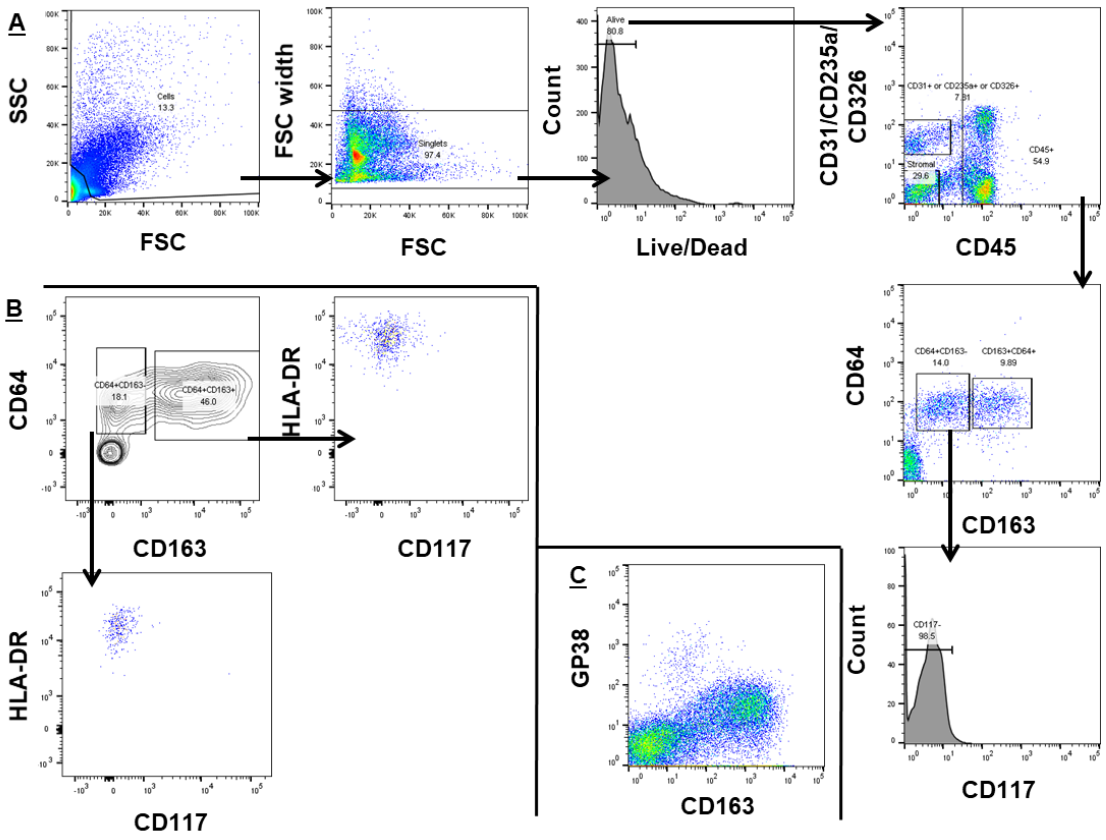


Figure 46: Representative plots of $CD45^+$ cells obtained from digestion of human synovial tissue. A, Demonstration of the gating strategy used to identify macrophage populations; after exclusion of debris and doublets live cells are identified through the use of an amine binding dye, staining for the cell surface markers CD45, and a combination of CD31, CD235a, and CD326 in the same channel allows identification of stromal populations (negative for these markers) and hematopoietic populations ($CD45^+$ and either positive or negative for the other markers). B, two macrophage populations identified by CD64 and CD163 expression showed a uniform level of HLA-DR expression and as such could not be further stratified by the use of this marker. C, A $GP38^{hi}CD163^{lo}$ population was present in some synovial samples but not in others, the nature of this intermittent population was not known.

Continuing the investigation of macrophage marker expression in the cell suspensions generated by digestion of synovial tissue, the expression of HLA-DR and CD206 on macrophages was assessed. HLA-DR is expressed by antigen presenting cells and would aid in the identification of macrophages and CD206

had been used as a marker of less inflammatory macrophages (Ambarus et al., 2012a; Verreck et al., 2004). In addition to these markers autofluorescence in the FITC channel was assessed as a potential marker, as autofluorescence had been shown to be high in macrophages and had previously been used as a parameter to identify macrophages (Aubin, 1979; Pankow et al., 1995; Umino et al., 1999). When investigating the expression of CD45 versus CD235a/CD326/CD117, two $CD45^+$ populations could be found, a $CD45^+CD235a/CD326/CD117^-$ population and a CD45+ population that showed a low level of fluorescence in the combined CD325a/CD326/CD117-PECy7 channel (Figure 47). This dimly PE-Cy7 fluorescent population showed an increased amount of autofluorescence in the FITC channel when compared with the $CD45^+CD235a/CD326/CD117^-$ population. In addition the majority of the population with higher FITC channel autofluorescence was positive for HLA-DR and positive for both CD163 and CD206 whereas the non-autofluorescent population was a mix of $HLA\text{-}DR^-$ and $HLA\text{-}DR^+$ cells with most of the HLA-DR positive population being $CD206^-CD163^-$. This indicated that the autofluorescent population consisted of a macrophage type population whereas the non-autofluorescent population did contain $HLA\text{-}DR^+$ cells but these cells lack expression of the macrophage markers CD163 and CD206.

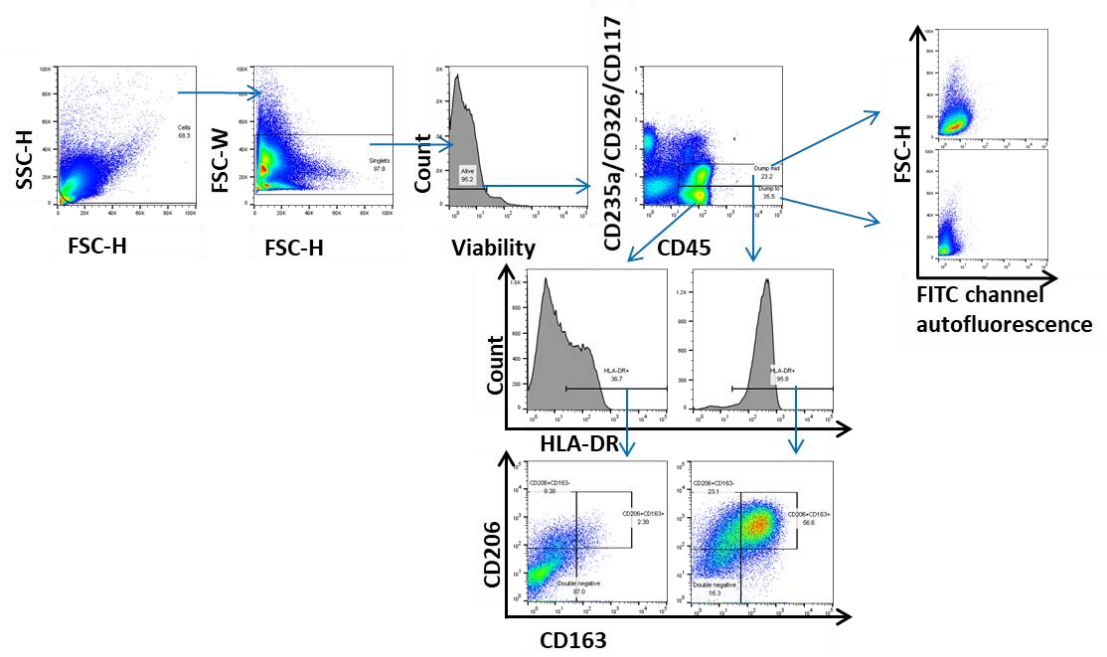


Figure 47: Expression of macrophage markers and FITC autofluorescence profile of the two CD45$^+$ populations visible. Increased autofluorescence was found in macrophage populations and in synovial digestions the more autofluorescent CD45$^+$ population demonstrated higher levels of staining for HLA-DR, CD163, and CD206 in keeping with this pattern.

Cell sorting the possible macrophage populations and generating cytospins can often aid in the identification of cell types by examining the morphological characteristics of the sorted cells. Using this approach CD45$^+$HLA-DR$^+$ cells split according to the amount of fluorescence in the PE-Cy7 dump channel and then further split according to CD206 and CD163 expression were sorted and used to generate cytospins (Figure 48, the dump low CD206$^+$CD163$^+$ population was lost due to equipment issues). The dump mid CD206$^+$CD163$^+$ cells showed morphological characteristics in keeping with those of macrophages, namely a ruffled cell membrane and oval shaped, sometimes pinched/kidney shaped nuclei. The dump mid CD206$^+$CD163$^-$ population also showed a similar morphology however the dump low CD206$^+$CD163$^-$ population demonstrated a different morphology with the presence of large vacuoles in some of the cells and other cells with smaller amounts of cytoplasm and dendritic

processes/pseudopodia. The dump low and dump mid double negative (DN) population both contained cells with a similar morphology to the double positive population, however they also contained more cells with a small amount of cytoplasm and more clear kidney shaped/pinched nuclei. These findings suggested that the $CD206^{+}CD163^{+}$ dump mid population was predominantly macrophage-like in appearance whereas the other populations were more heterogeneous, sometimes containing macrophage cells but in the presence of cells with various other morphological features.

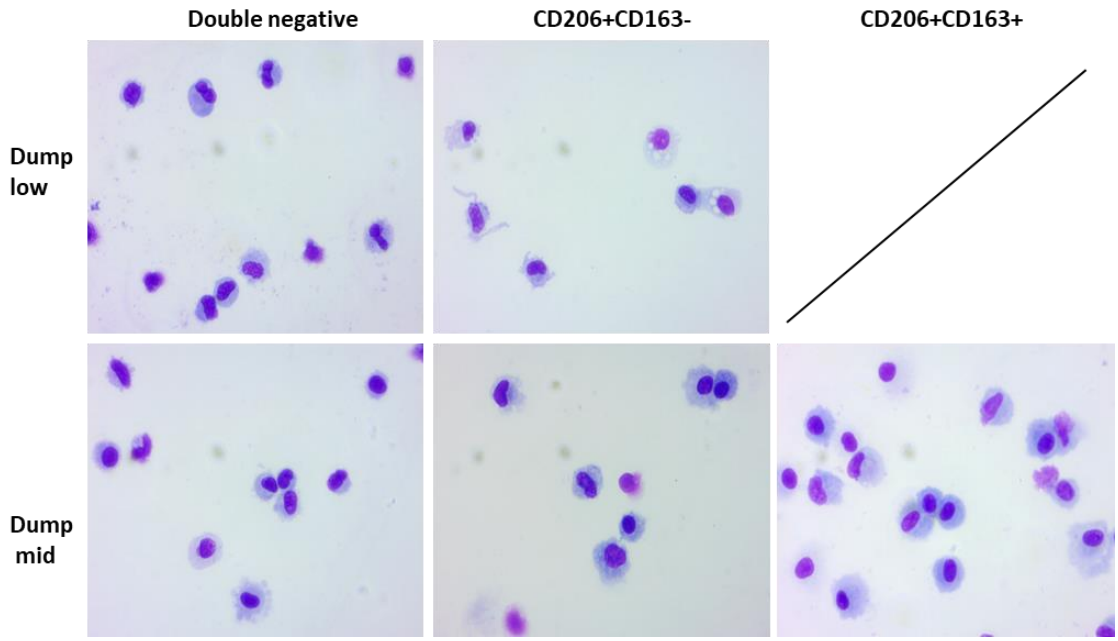


Figure 48: Cytospins of sorted putative synovial macrophage populations stained with Giemsa stain. The cells were sorted from the $CD45^{+}$HLA-$DR^{+}CD235a/CD326/CD117^{-}$ population (dump low) or the $CD45^{+}$HLA-$DR^{+}CD235a/CD326/CD117^{lo}$ population (dump mid) with further separation according to CD206 and CD163 expression. The $CD206^{+}CD163^{+}$ population appeared to have a macrophage like appearance (ruffled membrane; oval, pinched nucleus) and could have been slightly larger in size than the double negative population. The dump low $CD206^{+}CD163^{+}$ population was lost due to equipment failure.

As a final step to aid in the identification of the myeloid cells within the $CD45^+$ fraction of the digested synovial tissue cells, several experiments were completed in which the cells were stained with a different panel of markers. A summary of the results of these experiments can be found in Table 25. Using this approach four main populations were identified. Population 1 was negative for most markers and could be split into a large $CD11c^-$ population and a very small $CD11c^+$ population. Population 2 expressed low levels of HLA-DR but was positive for CD64 and also CD14 hinting at a myeloid identity. The macrophage markers CD163 and CD206 were absent from population 2. Population 3 had high expression of HLA-DR and was positive for CD64 and CD206 showing an increasing macrophage phenotype when compared to populations 1 and 2. Population 4 had slightly higher HLA-DR expression than population 3, was positive for CD64, CD206, CD163, GP38 and had high expression of CD14, CD1c, Mer, CD15, and CD197. Interestingly, expression of the putative tissue resident macrophage marker Mer followed a similar pattern to HLA-DR expression across the populations with no expression in population 1 and increasing expression in populations 2, 3, and 4. CD11c, a marker which is more highly expressed on dendritic cells was absent from all populations with the exception of a small subset of population 1 (Guilliams et al., 2016). CD197 (CCR7) is normally associated with activated dendritic cells and was expressed on populations 2, 3, and 4 yet CD11c was absent from these population (Dieu et al., 1998; Sallusto et al., 1998).

Marker	Population 1	Population 2	Population 3	Population 4
HLA-DR	-	lo	hi	hi+
CD64	-	+	+	+
CD206	-	-	+	+
CD163	-	-	-	+
GP38	+	-	-	+
CD14	-	int	int	hi
CD1c	int	-	int	hi
MerTK	-	lo	Int	hi
CD15	lo	lo	int	hi
CD197	-	int	int	hi
CD11c	-/+	-	-	-

Table 25: Expression of myeloid/macrophage markers in CD45$^+$ cells isolated from synovial tissue. Four main populations could be identified using the markers listed in the table which summarises the results of several different staining panels.

As part of the ongoing collaboration with partners in the AMP RA synovial biopsy tissue was cryopreserved and transferred to the lab of Michael Brenner (Boston, USA) for analysis using CyTOF, a technique combining mass spectrometry and flow cytometry technologies to allow the analysis of more markers than possible by traditional flow cytometry without the associated issues with spectral overlap (Bandura et al., 2009). Due to the increased dimensionality of the data dimensionality reduction techniques are used to visualise the relationships within the data. t-distributed stochastic neighbour embedding (tSNE) takes the output of PCA and facilitates the analysis of relationships between data without the limitations of the linear PCA approach (Van Der Maaten and Hinton, 2008). The data generated and provided by the collaborators identified clear clusters of T, B, myeloid, and stromal cells confirming the findings of the work above (Figure 49, Figure 50). CD14, CD64, and HLA-DR expression all showed a high degree of overlap with regards to the cluster of cells in which they were found (Figure 49). However CD11c staining in the CyTOF data appeared to be much higher than expected from the results above indicating some discordance between the two techniques. A possible explanation may be that the digestions completed by the

collaborators used the weaker Liberase TL mix as opposed to the Liberase TM used for the other marker work above leading one to hypothesise that CD11c may be cleaved from the cell surface by the harsher enzymatic mix. Alternatively, the antibody clone used by collaborators could have had a superior avidity.

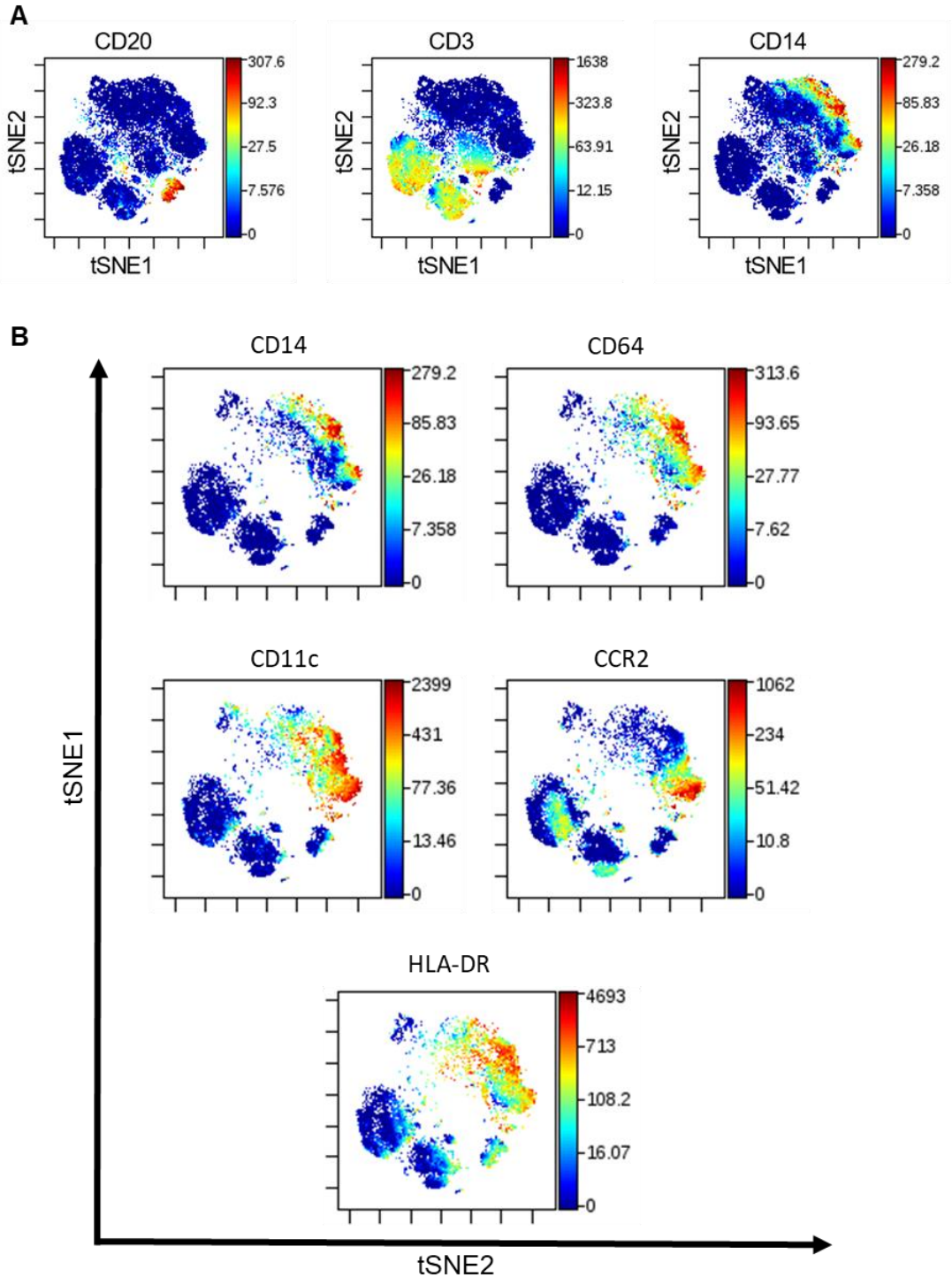


Figure 49: Representative coloured tSNE plots of leukocyte (A) and myeloid (B) marker expression on CD45$^+$ cells yielded by digestion of RA synovial biopsy material using the protocol developed in this project. Data were obtained using Cytometry by Time Of Flight (CyTOF) analysis by collaborators in the lab of Michael Brenner as part of the AMP collaboration. Colour shows normalised expression of the relative marker. tSNE = t-distributed stochastic neighbour embedding. Protocol development in collaboration with Brenner group, digestion and plots by Kevin Wei and adapted by the author.

Using the same CyTOF data but instead investigating the expression of markers of stromal populations provided some insight to the stromal populations (Figure 50). Vascular endothelial cadherin (VE-cadherin), a marker specific for endothelium, expression could be found in a cluster negative for FAP and GP38 indicating these markers do not label endothelial cells (Vestweber, 2008). However a small cluster of cadherin-11^{+}VE-cadherin^{+} cells could be found and all of the VE-cadherin expressing cells appeared to be CD34^{+} reinforcing that caution should be exercised to exclude endothelial cells from the analysis of synovial fibroblast populations. Regarding the VE-cadherin negative populations GP38 and FAP showed a considerable amount of overlap in the populations labelled, with two separate clusters identified using these markers; it could be argued that all GP38^{+} cells also appeared to be FAP^{+}. CD34 on the other hand only appeared to be expressed in one of the FAP^{+}GP38^{+} clusters, showing a more restricted pattern of expression.

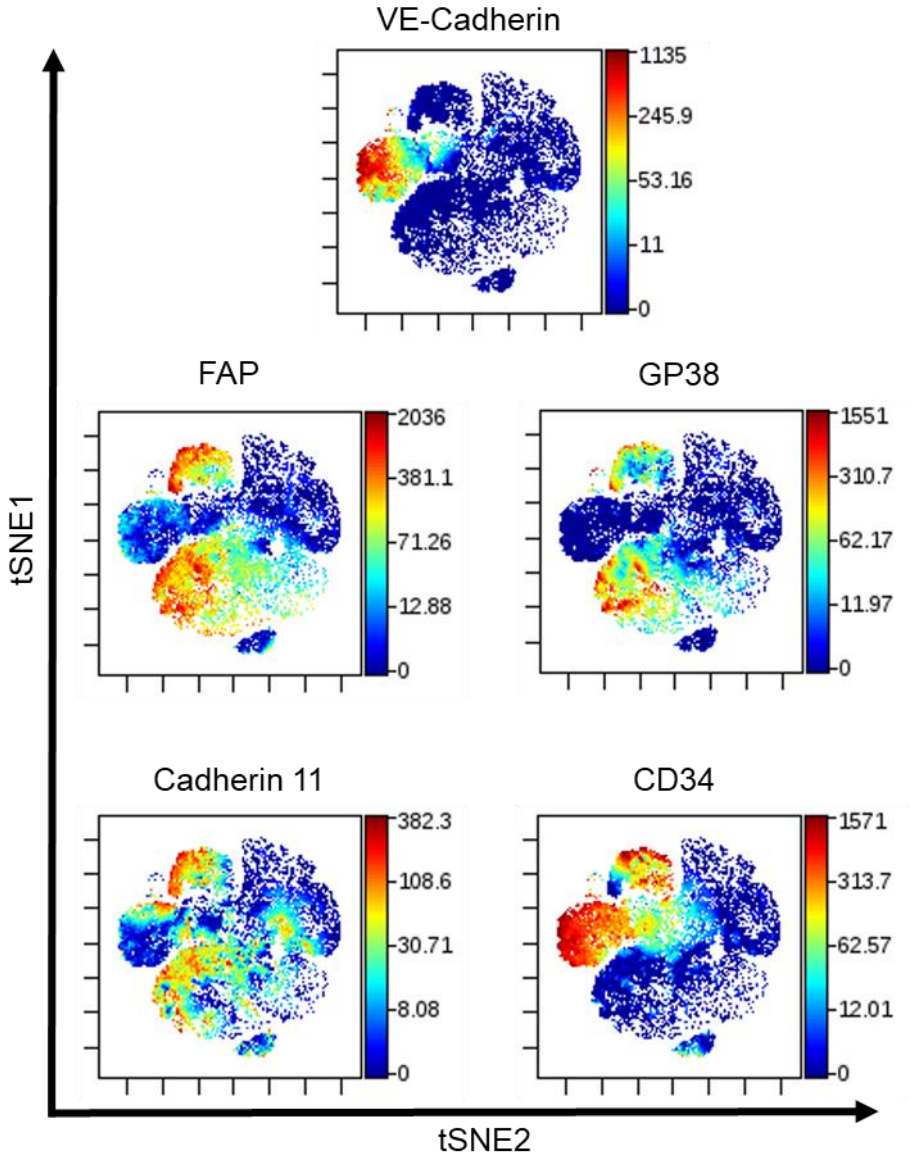


Figure 50: Representative coloured tSNE plots of stromal marker expression on CD45$^-$ cells yielded by digestion of RA synovial biopsy material. Data were obtained using Cytometry by Time Of Flight (CyTOF) analysis by collaborators in the lab of Michael Brenner as part of the AMP collaboration. Colour shows normalised expression of the relative marker. tSNE = t-distributed stochastic neighbour embedding. Protocol development in collaboration with Brenner group, digestion and plots by Kevin Wei and adapted by the author.

In summary three main stromal populations recently identified in single cell sequencing experiments could be identified in the cells yielded by digestion of synovial tissue. The work above identified possible macrophage populations identified by HLA-DR, CD206, and CD163 expression.

4.4 Discussion

The work in this chapter identified stromal populations present in the synovium that are modulated during different disease stages. With regards to stromal populations the surface marker GP38 was associated with lining layer fibroblasts, confirming the findings of other investigators (Bauer et al., 2006; Del Rey et al., 2014; Ekwall et al., 2011). In addition, the expression of GP38 has been found to be increased by proinflammatory factors indicating expression increases in response to inflammation (Croft et al., 2016). Staining of synovial tissue sections did not find significantly increased GP38 expression in VeRA. The fibroblast marker FAP was significantly increased in very early RA compared to uninflamed controls but not compared to other stages of RA and CD90 was significantly increased in resolving arthritis compared to normal samples. FAP expression is increased in the inflamed synovium in the murine collagen-induced arthritis model and is highly expressed by MMP1 expressing lining layer fibroblasts that lack CD90 in RA synovium (Bauer et al., 2006; Laverman et al., 2015). Additionally increased expression of mRNA encoding FAP was found in synovial fibroblasts cultured from patients with very early RA in section 2.3.2 of this thesis. The lack of an increase in $CD90^+$ expression in RA samples is of interest as it has been postulated that CD90+ stromal cells are the major source of IL-6 in cancer-associated fibroblasts and given the prevalence of this cytokine in RA it could be hypothesised that the expansion of CD90+ synovial fibroblasts is involved (Huynh et al., 2016).

Moving on to macrophage subsets, immunofluorescent staining of uninflamed and RA synovial tissue identified $CD68^+$ macrophage populations in lining and

sublining regions of the synovium. Using additional markers such as CD163 and HLA-DR two broad populations were demonstrated in the synovium, HLA-$DR^{hi}C163^{lo}$ and HLA-$DR^{lo}CD163^{hi}$ macrophages which match the findings in a mouse model of arthritis that identified these two populations as infiltrating and resident populations respectively (Misharin et al., 2014). In the murine K/BxN serum-transfer model of arthritis, this work suggested that resident macrophages in the synovium maintain an anti-inflammatory phenotype and aid in the control of arthritis, however an infiltrating macrophage population was required for normal resolution of the arthritis despite initially expressing both pro- and anti-inflammatory genes, with progression to a more regulatory phenotype as the arthritis progresses. Due to the low dynamic range of confocal microscopy, difficulty in quantifying data obtained using this technique, and lack of ability to isolate live cells for functional assays, it was decided to develop a unified digestion protocol in collaboration with partners in the NIH AMP which could be used to generate cell suspensions for flow cytometric analysis and cell sorting from synovial biopsy tissue.

Investigation of methods for disruption of synovial tissue found that mechanical dissociation was not sufficient to isolate stromal cells. Enzymatic digestion with Liberase TM performed well at isolating stromal and macrophage populations and data suggested that Liberase TL was not quite as efficacious as Liberase TM, although these findings were not statistically significant. Freezing of the tissue did not result in the total loss of stromal or $CD45^{+}$ populations or a reduction in viability, but did result in a reduction of the proportion CD45+ cells relative to the proportion of stromal cells.

It is important to mention that the number of replicates for the freezing work were low and should be increased to confirm or refute the findings. With this is mind an immediate theory regarding the change in proportion would be that the $CD45^+$ cells are killed by the freezing protocols. However, given that no change in the viability was apparent with the 10 minute freezing protocol this hypothesis seems unlikely. Another possibility could be that the freezing protocols disrupt the tissue aiding in the recovery of cells by digestion and thus leads to an increase in the number of stromal cells being isolated which would result in a reduction in the $CD45^+$ proportion.

Regarding the populations yielded by digestion three key populations recently identified as the major stromal subsets in the synovium by single cell RNA-sequencing were present in the cell suspensions (Mizoguchi et al., 2017). These populations were identified as $CD45^-CD325a/CD326/CD117^-CD31^-$ and then either $CD90^-GP38^+$, $CD90^+GP38^+CD34^+$, or $CD90^+GP38^+CD34^-$ and were present in various proportions in the digestions. In the case of the macrophages three putative macrophage populations were identified as $CD45^+HLA\text{-}DR^+$ and then separated as $CD206^-CD163^-$, $CD206^+CD163^-$, or $CD206^+CD163^+$. It was not possible to use the traditional macrophage marker CD68 due to the intracellular nature of the antigen which lead to highly varied results (Micklem et al., 1989; Parwaresch et al., 1986; Pulford et al., 1989). Increased autofluorescence in the FITC channel has previously been used to identify macrophages which is of interest as the three possible macrophage populations appeared to have differences in this autofluorescence with the $CD206^+CD163^+$ population having higher levels than the $CD206^-CD163^-$ population (Aubin, 1979; Pankow et al.,

1995; Umino et al., 1999). However this feature could be due to aspects of the immunofluorescent staining such as interactions between the fluorophores, a possibility that was not tested in this work. Although a full panel of macrophage markers was tested HLA-DR, CD206, CD163, and CD64 remained the best discriminators between three macrophage populations identified. During *in vitro* differentiation of human peripheral blood monocytes into macrophage populations the expression of CD206 and HLA-DR was found to increase on the surface of all macrophage subsets compared to monocytes but CD163 expression was absent from the pro-inflammatory M1 subset differentiated using GM-CSF (Rey-Giraud, Hafner and Ries, 2012). When comparing the multiple anti-inflammatory M2 subsets the M2a population, differentiated with M-CSF and IL-4, showed lower expression of CD163 and CD14 when compared to the M2c subset, differentiated with M-CSF and IL-10, and the M2c subset had decreased HLA-DR and CD206 levels. These findings mirror some of the data in this work indicating the macrophage populations identified in the synovial digestions may be similar to *in vitro* M2 subtypes, however some differences do exist such as no change in HLA-DR or CD206 levels between the populations which is probably due to the complex cytokine milieu surrounding the macrophages *in vivo*. Finally, assessment of the morphology of the *ex vivo* macrophage populations indicated the $CD206^{+}CD163^{+}$ cells are indeed macrophage populations but found a higher degree of heterogeneity in the $CD206^{-}CD163^{-}$ population hinting that this may represent a mixed population.

The analysis of the immunofluorescent staining of synovial tissue sections had several limitations. Confocal microscopy is a useful tool for obtaining data on the

geographic location of cells within a tissue but has a limited dynamic range. In the case of some markers with a broad range of expression this could result in the loss of data on low or high expressing cells biasing the analysis. Additionally the analysis method used in this work, to manually select reasonable quality regions of tissue, could clearly result in a bias in the analysis. The analysis of the data by several operators could also introduce bias however approaches were taken to limit this effect by comparing the results of comparisons of the same tissues across operators.

Regarding the digestion optimisation work several improvements could be made. The work was completed using synovial tissue from different clinical outcome groups obtained by arthroscopy or ultrasound-guided biopsy. The heterogeneity introduced by these factors would have confounded the comparisons possibly leading to incorrect conclusions on the optimal digestion technique. If the work was to be repeated it would be sensible to use tissue obtained using one method and preferably from one clinical outcome group. In addition, due to the nature of the collaborative work the freezing protocol was optimised on digestions using Liberase TL even though Liberase TM was chosen as the enzyme for use in the optimised protocol. Ideally this work should be repeated using the appropriate enzyme mix and digestion protocol to reduce spurious findings.

For the digest optimisation work data on the size of populations was all proportional meaning changes in the size of one or more populations would result in changes in others even if the cell count remained the same. To this end the experiments should be repeated using counting beads to obtain cell counts and

the tissue weights prior to digestion should be noted to allow the calculation of normalised cell counts which would clarify the results.

In summary the expression of stromal markers FAP and CD90 were significantly different in very early RA and resolving arthritis respectively compared to normal samples. Macrophage populations that may correspond to resident and infiltrating populations could also be identified. Through work with collaborators a unified digestion protocol was established which can be used to isolate synovial cellular populations. Through the investigation of several cell markers both stromal and possible synovial populations could be identified in the populations yielded by synovial digestion.

The increased expression of CD90 and FAP indicates that different fibroblasts subsets may be more active during different arthritic disorders. FAP, a marker identifying both lining and sublining fibroblasts, was higher in very early RA whereas CD90, a sublining specific marker, was increased in resolving arthritis. From this evidence it could be argued that the $CD90^{+}$ synvovial population is beneficial, perhaps possessing a pro-resolution phenotype whereas the FAP expressing population could drive RA pathogenesis. However this contrasts somewhat with the literature given that tumour CD90 expression is associated with a poor prognosis in gallbladder carcinoma (Zhang, D. H. et al., 2016). The identification of $CD206^{+}CD163^{+}$ and $CD206^{+}CD163^{-}$ macrophages contributes towards the knowledge on synovial macrophages. CD206 and CD163 have both previously been shown to be expressed by synovial macrophages and are thought to be expressed by macrophages with an anti-inflammatory phenotype

(Ambarus et al., 2012b; Fahy et al., 2014; Jablonski et al., 2015). In addition, the finding of MerTK expression on both of these populations could mean they represent tissue-resident synovial macrophages with the main function of maintaining homeostasis within the joint (Gautier et al., 2012).

As the work in this chapter developed a protocol allowing the identification and isolation of pure populations of synovial fibroblasts and macrophages it was decided to characterise the transcriptome of these populations both in treated and treatment naïve RA along with arthritis with a distinct aetiology.

5 *EX VIVO* CHARACTERISATION OF SYNOVIAL FIBROBLAST AND MACROPHAGE POPULATIONS ISOLATED FROM BIOPSY MATERIAL

5.1 Introduction

To further address the third hypothesis, that different fibroblast and macrophage subsets exist within the synovium with distinct functions, stromal and macrophage populations were isolated from synovial tissue using the digestion protocol established in chapter 3 and the transcriptome of the cells was interrogated using whole transcriptome RNA-sequencing with an aim of identifying populations, ascribing functions, and comparing differences between disease states. RNA-sequencing of cell populations provides data on the average gene expression for cells in that population and allows comparison between different populations to elucidate functional differences. RNA-sequencing was used rather than cDNA microarrays as RNA-sequencing can work with low amounts of RNA and also the dynamic range of RNA-sequencing is greater than microarrays and less subject to background noise (Hrdlickova, Toloue and Tian, 2017).

Clinical samples for the RNA-sequencing experiments were dependent on clinical sample availability from biopsy procedures, as the procedures to freeze, digest and recover cells had to be determined before sample collection could commence. Therefore the disease groups used were not the same as those described previously for the BEACON cohort. The three groups used were (1) patients with RA who have had symptoms for longer than 3 months but have not yet received any treatment, (2) patients with RA who have had symptoms for longer than 3 months and have had an inadequate response to methotrexate and are therefore progressing onto biologic therapy (phase 2 in Figure 1), and (3) a comparison group of patients who had developed arthritis as a result of receiving

the anti-programmed cell death protein 1 (PD1) checkpoint inhibitors pembrolizumab or nivolumab to treat advanced cancer. Patients receiving immune checkpoint inhibitors have been found to be at risk of developing a range of adverse effects resulting from immune system overactivity (Weber et al., 2013). For example patients receiving nivolumab, an anti-PD1 agent, suffered a range of adverse events such as pneumonitis, allergic rhinitis, diarrhoea, rash, puritis, vitiligo, pruritic rash, macular rash, alopecia, hypopigmentation, hypothyroidism, and hyperthyroidism (Topalian et al., 2012). Most of the adverse events developed around 7 weeks after treatment initiation, can be successfully treated, and do not appear to be persistent. However, development of arthritis occurred later with two patients showing symptoms around 11 and 14 months after treatment initiation (Calabrese et al., 2017; Chan et al., 2015). It is thought that the issues arise due to the removal of restraints upon T lymphocyte activation by the anti-PD1 treatment. However it is unclear if T cells recognising relevant auto-antigens were already present in the patients or developed subsequent to the treatment (Chan et al., 2015; Johnson et al., 2016).

The work in this chapter aimed to investigate the phenotype of the synovial fibroblast and macrophage subpopulations identified in chapter 4 as these populations were as yet not fully characterised in the literature. Specifically RNA-sequencing was used to characterise the transcriptome of these populations and address the following questions:

1. Do the synovial fibroblast and macrophage subpopulations possess distinct transcriptomes?
2. Does treatment of RA induce significant differences in the transcriptome

of these subpopulations?

3. Does the transcriptome of synovial subpopulations vary between RA and anti-PD1 induced arthritis?

5.2 Methods and materials

5.2.1 Biopsy material and clinical outcome groups

Synovial tissue from patients with RA and having experienced symptoms for a duration longer than 3 months but not yet receiving any DMARDs or biologics (untreated RA), patients with RA which have experienced symptoms for a duration longer than 3 months and have received either DMARDs alone or DMARDs in combination with biologics, and from patients who developed arthritis after receiving nivolumab or pembrolizumab to treat advanced cancer (anti-PD1) were obtained by ultrasound-guided biopsy and cryopreserved by freezing in 1ml Cryostor 10 (Sigma Aldrich) with 10 minutes incubation at 4^{o}C before transfer to -80^{o}C in a Mr Frosty (Nalgene). When required for batch processing, tissue was rapidly thawed in a 37^{o}C waterbath and digested using the protocol developed in section 8.3. Clinical details of patients from which biopsies were taken and samples generated are presented in Table 26.

ID	Group	Site	CCP +/-	RF +/-	Disease duration (Weeks)	Age	Sex	DMARD	Prednisolone	DAS28 ESR baseline	ESR (mm/ hr)	CRP (mg/l)
BX186	TreatedRA	Knee	+	+	158	43	F	MTX, Etanercept	None	4.3	2	0
BX189	TreatedRA	Knee	-	-	36	66	F	MTX, Etanercept	None	5.5	13	5
BX190	RA	Knee	+	-	52	39	M	None	None	6.9	9	58
BX199	RA	Knee	-	-	18	69	M	None	None	5.1	2	7
BX200	RA	Knee	+	+	18	57	M	None	None	5.4	21	10
BX210	PD1	Knee	-	-	33	51	M	Pembrolizumab	10mg	3.5	8	9
BX211	RA	Knee	+	+	79	51	M	None	None	5.8	21	59
BX215	TreatedRA	Knee	-	-	429	59	F	MTX, SSA	None	6.0	5	0
BX219	PD1	Knee	-	-	22	51	F	Pembrolizumab	None	5.9	49	71
BX224	PD1	Knee	-	-	10	74	F	Nivolumab	None	4.4	32	7
BX225	PD1	Knee	-	-	21	66	F	Pembrolizumab	6mg	4.3	42	9
BX228	TreatedRA	Knee	+	+	186	71	M	MTX, HCQ	None	4.9	20	4
BX229	TreatedRA	Knee	+	+	78	56	F	MTX, HCQ	None	6.3	9	11
BX230	RA	Ankle	-	+	43	62	M	None	9mg	5.7	12	4
BX231	RA	Wrist	-	-	8	69	M	None	None	3.9	16	13
BX233	TreatedRA	Knee	+	+	192	43	M	MTX, HCQ, LEF	None	6.6	31	16
BX235	PD1	Knee	-	-	22	76	F	Pembrolizumab	None	3.8	12	0

Table 26: Clinical characteristics of the patients from which synovial biopsies were collected and used for the RNA-sequencing experiments. NA = Not available, DAS28 = Disease activity 28 score, ESR = Erythrocyte sedimentation rate, CRP = C-reactive protein, CCP = Anti-cyclic citrullinated peptide, RF = Rheumatoid factor, MTX = Methotrexate, SSA = Sulfasalazine, HCQ = Hydroxychloroquine, LEF = Leflunomide.

5.2.2 RNA sequencing and data analysis

16 cryopreserved synovial tissues and 1 fresh synovial tissue, all obtained by ultrasound guided biopsy, were digested in three batches (2 batches of cryopreserved tissue and one batch with the single fresh tissue) according to the protocol in section 8.3, stained with the panel in Table 27, and six synovial populations were sorted using a MoFloAstrios EQ. RNA was isolated from the sorted cells as in 3.2.11 and RNA quantity and quality was assessed using RNA 6000 Pico Kits (Agilent Technologies) as specified in the manufacturer's instructions in combination with a 2100 Bioanalyzer (Agilent Technologies).

Target	Conjugate	Clone	Isotype	Supplier
CD31	PerCP	9G11	Mouse IgG1	R&D Systems
CD34	FITC	581	Mouse IgG1, κ	Biolegend
CD45	V450	2D1	Mouse IgG1, κ	BD Biosciences
CD64	BV605	10.1	Mouse IgG1, κ	Biolegend
CD90	PerCP-Cy5.5	eBio5E10	Mouse IgG1, κ	eBioscience
CD117	PE-Cy7	104D2	Mouse IgG1, κ	Biolegend
CD163	PECF594	GHI/61	Mouse IgG1, κ	BD Biosciences
CD206	APC	15-2	Mouse IgG1, κ	Biolegend
CD235a	PE-Cy7	HI264	Mouse IgG2a, κ	Biolegend
CD326	PE-Cy7	9C4	Mouse IgG2b, κ	Biolegend
GP38	PE	NZ-1.3	Rat IgG2a	eBioscience
HLA-DR	BV510	L243	Mouse IgG2a, κ	Biolegend
Live/Dead	Amine binding dye (APC-Cy7 emission)	N/A	N/A	Life Technologies

Table 27: Antibody panel used to stain and sort cells yielded by synovial digestion

Library preparation and RNA sequencing was completed by Virginie Petitjean in the lab of Marc Sultan (Novartis, Basel). Library preparation was completed using the SMARTer Stranded Total RNA-Seq Kit - Pico Input Mammalian (Takara Bio) using 250pg of RNA from each sample where available and for samples with less RNA than 250pg the total RNA available was used. 6 positive controls

recommended by the SEQC/MAQC-III Consortium (Consortium, 2014), intact SEQC A, intact SEQC B, intact SEQC C, degraded SEQC A, degraded SEQC B and degraded SEQC C were also used to prepare libraries in parallel with the other samples using 250pg of RNA as the input from each control. A negative control of nuclease-free water was also prepared in the same manner. All libraries were prepared with dual index labelling (8 base pair (bp) + 8 bp) to allow multiplexing of samples, were then pooled in equimolar ratios to produce 4 multiplexes with 28-29 libraries in each, and then each multiplex was sequenced in quadruplicate using 4 HiSeq Paired End v4 Flow Cells (read lengths of 2 x 76bp, Illumina) run on an HiSeq 2500 instrument (Illumina). Sequences were de-multiplexed allowing 2bp mismatches across the index reads which produced between 15 and 113 million reads per library (62 million mean).

The author of this thesis travelled to Novartis to view and be instructed on the analysis of the generated data. The sequencing data were prepared for analysis by Sabina Pfister (Novartis). FASTA files generated from the same samples in the replicate sequencing runs were merged using UNIX scripts generated by scripts written in the R programming language. Any possible contaminating adaptor and/or index sequences in the reads were trimmed using Cutadapt (Martin, 2011) and quality control was completed using the FastQC software (Babraham Bioinformatics). Reads were aligned to the transcriptome and feature counts were produced using Bowtie2 (Langmead and Salzberg, 2012) in the Exon Quantification Pipeline established at Novartis (Schuierer and Roma, 2016).

Analysis from this point onwards was completed by the author of this thesis with supervision by Sabina Pfister, Archana Sharma-Oates, and Boris Noyvert (based at Novartis and University of Birmingham). Count data were analysed using R version 3.3.3 in the integrated development environment R Studio version 1.0.136 using packages available from the Comprehensive R Archive Network (https://cran.r-project.org/) or from Bioconductor (https://www.bioconductor.org/) the full list of which is available in section 8.4 (Gentleman et al., 2004; Huber et al., 2015). Features for which weight-adjusted counts were at least one in at least 75% of samples in group (one population from one disease group) were kept and other features excluded. Samples were then transformed using mean-variance modelling at the observational level (VOOM) to allowing linear modelling and normalised using the Trimmed mean of M-values (TMM) method and quantile-normalisation both of which were performed using the edgeR and limma packages (Ritchie et al., 2015). Linear modelling was then used on the prepared dataset to test for differential expression of genes between comparisons of interest using the limma package(Robinson, Mccarthy and Smyth, 2010). Gene set enrichment and was performed using the limma package on features for which entrez IDs were available. Gene ontology (GO) category and Kyoto encyclopaedia of genes and genomes (KEGG) pathway enrichment were completed using the goseq package (Young et al., 2010). For all enrichment tests only features with ensembl IDs included in the ensembl primary assembly were used meaning that gene alleles, alternative sequence genes, and deprecated ensembl IDs were excluded from the analysis. The full R script for the analysis is shown in section 8.4.

5.3 Results

5.3.1 The proportions of stromal and macrophage populations show relationships with clinical correlates and clinical outcome group

To elucidate the phenotype and differences of synovial fibroblast and macrophage populations between disease groups six populations were sorted from treatment naïve patients with RA (referred to as RA in the RNA-sequencing experiments), from patients with RA who have had an inadequate response and may have progressed onto biologics (referred to as RA in the RNA-sequencing experiments), and from cancer patients who have been treated with the anti-PD1 checkpoint inhibitor pembroluzimab and have subsequently developed arthritis as a result of the treatment (called anti-PD1 or PD1 in the RNA-sequencing experiments). Synovial biopsy material from 6 RA patients, 6 treated RA patients, and 5 anti-PD1 patients was digested and three fibroblast populations, all $CD45^{-}$ $CD235a/CD326/CD117^{-}CD31^{-}$ and either $CD90^{-}GP38^{+}$, $CD90^{+}GP38^{+}CD34^{-}$, or $CD90^{+}GP38^{+}CD34^{+}$, and three putative macrophage populations, all $CD45^{+}HLA\text{-}DR^{+}$ and either $CD206^{-}CD163^{-}$, $CD206^{+}CD163^{-}$, or $CD206^{+}CD163^{+}$, were sorted and used for RNA-sequencing experiments (Figure 51).

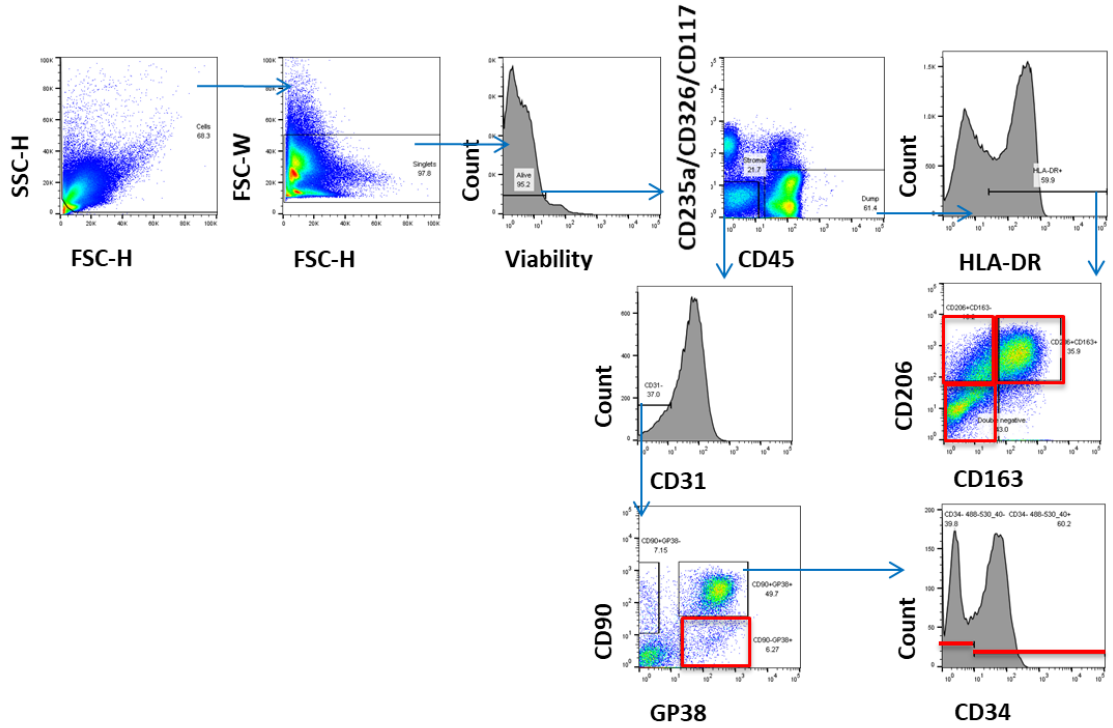


Figure 51: Representative plots of fibroblast and putative macrophage populations identified by flow cytometric analysis of cells yielded by digestion of synovial tissue. Gates highlighted in red indicate populations sorted for RNA-sequencing experiments.

When investigating the flow data obtained during cell sorting of these samples no significant differences in total cell viability were found between clinical outcome groups with the percentage of viable cells being approximately 90% in each disease (Figure 52). Comparing the percentage of the total viable cell population constituted by either stromal ($CD45^{-}CD235a/CD326/CD117^{-}$) or $CD45^{+}$ cells no significant differences were found in the proportions of these populations between disease stages. Comparing the proportions of the CD90-GP38+ and the CD90+GP38+ populations relevant to the total of the two populations across clinical outcome groups, no significant differences were found when tested using independent Kruskal-Wallis tests and Dunn's multiple comparisons tests for each population. Similarly, no significant differences were found using the same approach to test for differences in the proportion of CD34+ or CD34- cells in the

CD90+GP38+ stromal fraction, however only four replicates were available in the anti-PD1 group for this marker and this will have decreased the power of the test. No significant difference was found comparing the proportion of HLA-DR$^+$ cells in the CD45$^+$ cell fraction. Statistically significant differences did exist between the putative macrophage populations with the proportion of DN CD45$^+$HLA-DR$^+$ cells being higher in the RA samples than the anti-PD1 samples and the proportion of CD206$^+$CD163$^+$ cells being increased in the anti-PD1 patients compared to the RA patients (Figure 54).

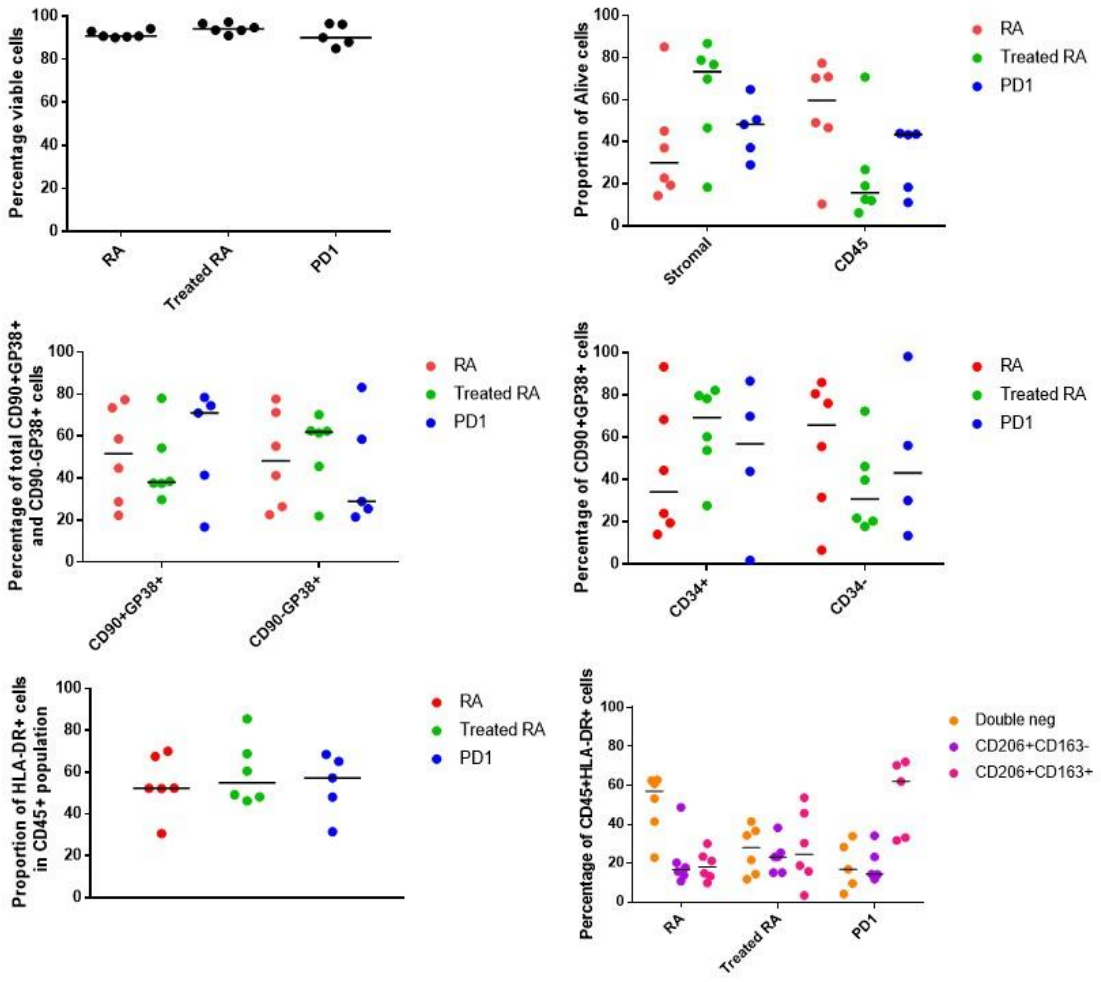


Figure 52: Flow cytometry data obtained from the digestion of synovial tissue from untreated RA patients (RA), treated RA patients with an inadequate response (treated RA), and cancer patients treated with anti-PD1 immune checkpoint inhibitor who have subsequently developed arthritis (PD1). Data on surface markers is presented as relative proportions and not cell numbers. With the exception of the proportion CD45+HLA-DR+ cells (expanded in Figure 54) no significant differences were detected when tested using Kruskal-Wallis tests with Dunn's multiple comparisons tests. Bars show the median.

Further investigation of the stromal populations uncovered relationships between the proportions of stromal populations and clinical measures (Figure 53). The proportion of the $CD90^{-}GP38^{+}$ population positively correlated with the ESR for the RA and treated RA patients. The percentage of $CD34^{-}$ cells in the $CD90^{+}GP38^{+}$ population in the RA and treated RA patients positively correlated with both ESR and the percentage of $CD45^{+}$ cells in the total viable cell population. When testing for correlations between the stromal and $CD45^{+}$

populations with CRP no significant correlations were found. However, the proportions of $CD90^{+}GP38^{+}$, $CD34^{+}$, and $CD45^{+}$ cells all correlated with synovial hypertrophy as measured by ultrasound. These findings indicated that the $CD90^{+}GP38^{+}$ and $CD90^{+}GP38^{+}CD34^{-}$ populations expand with inflammation and inflammatory infiltrate into the synovium. When the anti-PD1 samples were included in the analysis all correlations were weakened and correlations between stromal populations and either ESR or the $CD45^{+}$ proportion lost statistical significance. This might indicate spurious relationships within the smaller sample, or that the arthritis elicited by immune checkpoint inhibition has a different pathology to RA.

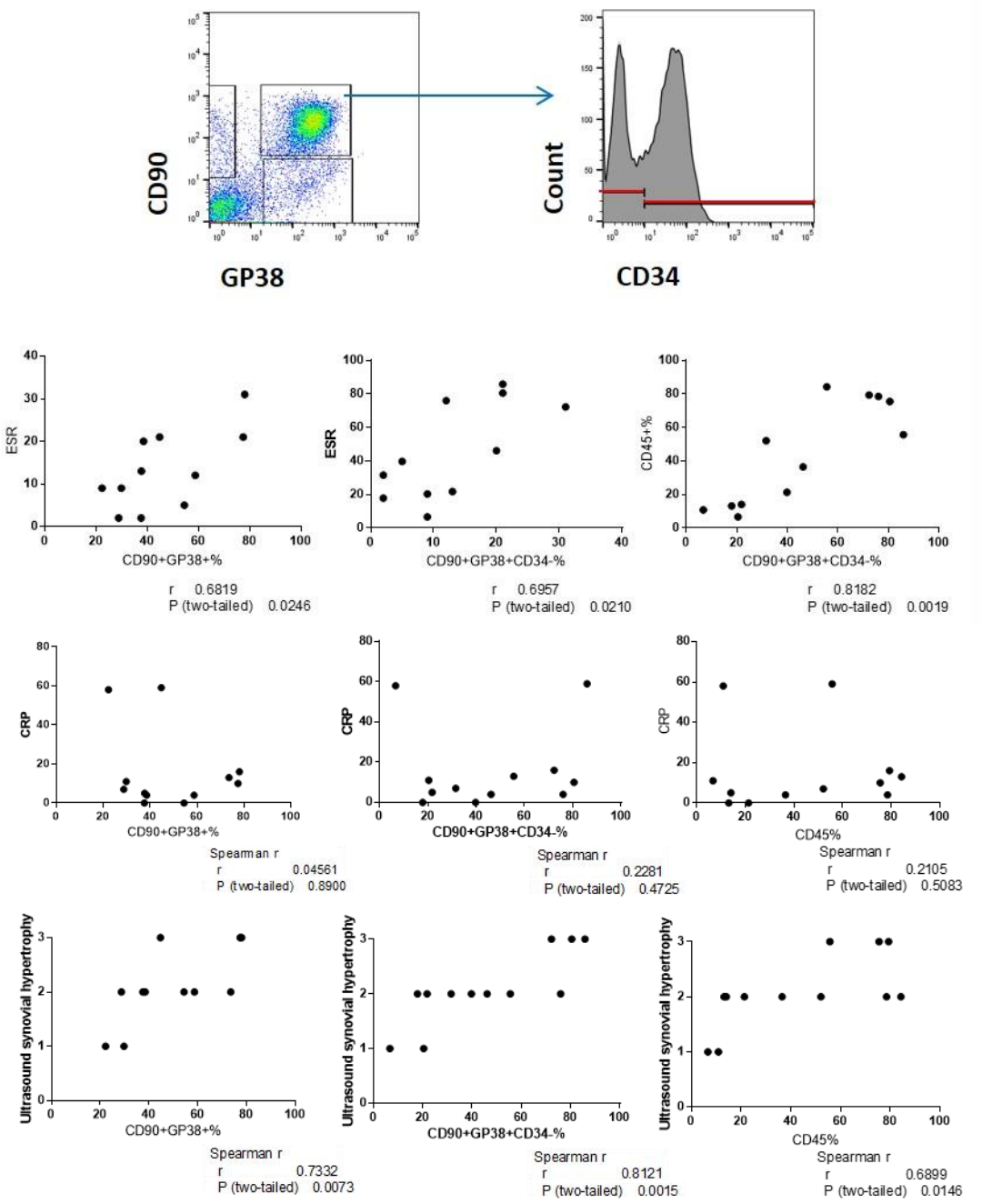


Figure 53: Correlations of the RA and treated RA stromal populations with $CD45^+$ cell proportion, erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP). The proportion of the $CD90^+GP38^+$ population relative to the total of the $CD90^+GP38^-$, $CD90^+GP38^+$ and $CD90^-GP38^+$ populations positively correlated with ESR but not CRP and the proportion of $CD90^+GP38^+CD34^-$ cells relative to the total $CD90^+GP38^+$ population positively correlated with ESR and the percentage of $CD45^+$ (relative to the total stromal and $CD45^+$ populations) cells but not CRP in the total viable cell population. The $CD90^+GP38^+$, $CD90^+GP38^+CD34^-$, and $CD45^+$ proportions all correlated with synovial hypertrophy as measured by ultrasound. Including the anti-PD1 samples in this analysis weakens all correlations. All tested using Spearman's correlation.

Performing a similar analysis on the macrophage populations did not show a clear relationship between ESR and population size (Figure 54). When analysing only the RA and treated RA samples or samples from all disease groups and testing for correlations between ESR and the $CD163^{+}CD206^{+}$ or the $CD163^{-}CD206^{-}$ populations no correlations were found, although the correlation between the $CD163^{+}CD206^{+}$ population in the RA and treated RA samples was close to significance. The population sizes did not correlate with ESR but did vary significantly between clinical outcome groups as a higher proportion of $CD163^{-}CD206^{-}$ cells were present in RA than in anti-PD1 patients and the reverse relationship with was found the $CD163^{+}CD206^{+}$ population as previously mentioned.

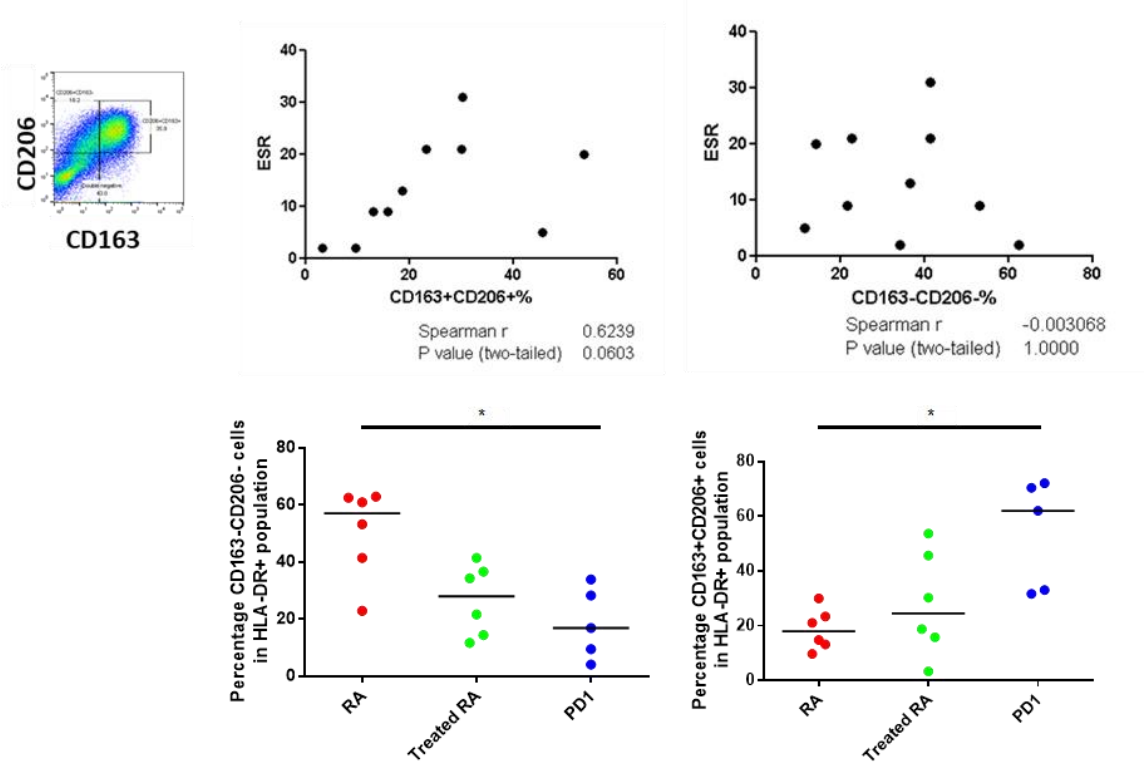


Figure 54: Comparison of the proportion of the three putative macrophage populations relative to the total CD45$^+$HLA-DR$^+$ population across the three disease groups. The proportions of the CD163$^+$CD206$^+$ or the CD163$^-$CD206$^-$ populations in the RA and treated RA samples did not correlate with the erythrocyte sedimentation rate (ESR, Spearman's correlation). The proportion of the two populations did significantly vary between the RA and anti-PD1 samples with a higher proportion of double negative cells in the RA samples and a higher proportion of the CD163$^+$CD206$^+$ cells in the anti-PD1 samples. Kruskal-Wallis test with Dunn's multiple comparison test, * = p value < 0.05.

5.3.2 Exclusion of poor quality RNA samples following sequencing

The synovial stromal and macrophage populations identified by flow cytometry were cell sorted into six populations from which RNA was then extracted (Figure 51). Analysis of the integrity and quantity of isolated RNA using a Bioanalyzer provided varied results, perhaps unsurprisingly given the variation in the number of sorted cells between populations and samples. Of the samples for which RNA Integrity Numbers (RINs) were obtained (81 samples) the median RIN was 6.4, 35 samples had RINs equal to or greater than 7, a typically used threshold for quality control of samples to be RNA sequenced, and 28 had a RIN score of 1

indicating either poor quality or an inability to accurately measure which may be the main reason in this case.

The number of sorted events, equivalent to sorted cells, varied between samples with the largest number of sorted events in one sample being 163,800 and the smallest 10 and the median number of sorted events from which RNA was extracted being 3010 (Table 28). The quantity of RNA used for library preparation was standardised to 250pg which 69 samples were able to provide, the remaining 32 samples had lower than 250pg of RNA available and in these samples the total available RNA was used to generate the libraries.

	Sorted events	Library RNA input (pg)
Max	163,800	250
3rd quartile	11,980	250
Mean	10,900	211
Median	3010	250
1st quartile	446	176
Min	10	16

Table 28: Descriptive statistics of the number of sorted events and RNA input into library preparations across all 101 samples sorted from synovial tissue digestions. The ceiling effect apparent in the Library RNA input data was due to 250pg of RNA from most samples being used for the library preparation with no sample using more than this quantity. However approximately one third of samples had less than 250pg of RNA available for the library preparation.

The libraries generated had varying sizes which was reflected in the distribution of non-normalised or transformed log2 counts per million (log2CPM) of the total features in each sample (Figure 55). As would be expected samples with a small library size had abnormal log2CPM boxplots with most values considered outliers and the median log2CPM residing around 0 indicating very few counts for any features were obtained for these samples. One would expect the samples with small library samples to be those derived from small numbers of cells or small

quantities of RNA. As seen in Figure 56 this relationship was true for these samples as both the concentration of RNA used for library preparations and the number of sorted events from each sample were significantly correlated with the size of libraries generated.

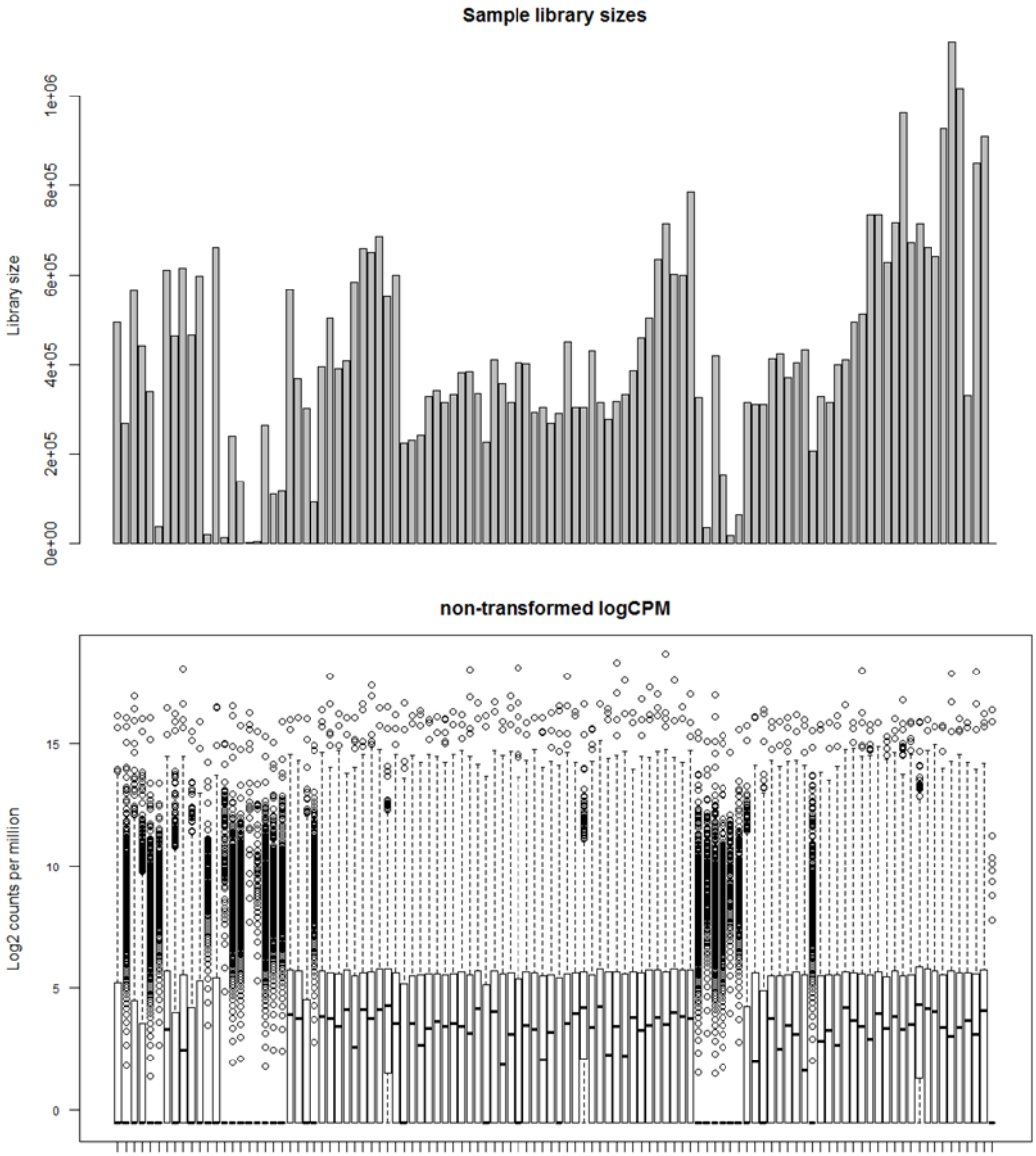


Figure 55: The size of the library for each RNA sample sequenced and the non-transformed/non-normalised log2 counts per million (log2CPM) of all features in each sample which passed filtering. Note the similarity between small library size and boxplots of log2CPM which were almost entirely classed as outliers. In the samples with reasonable library sizes the median log2CPM value showed substantial variation between samples highlighting the need for normalisation.

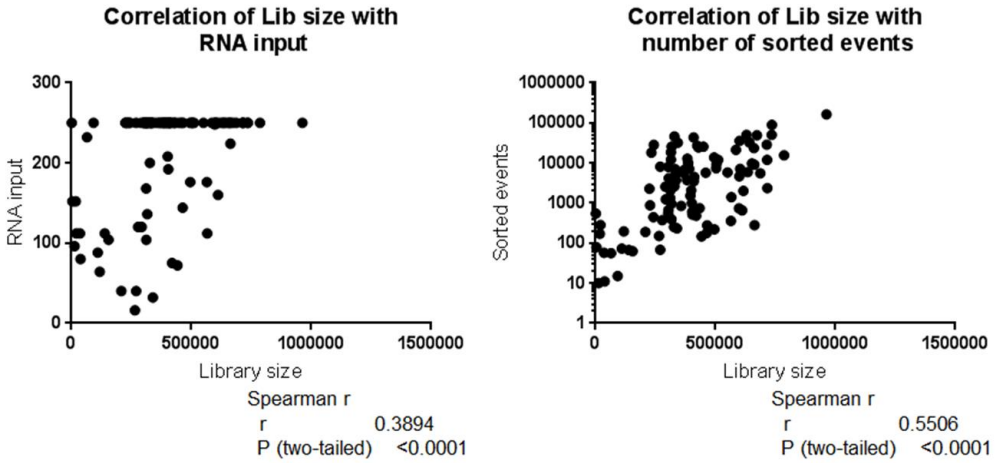


Figure 56: Library size is significantly correlated with the concentration of RNA used for the library preparation and also with the number of sorted events/cells from which RNA was isolated as assessed by Spearman's correlation.

Another technical aspect that must be considered with low-input RNA sequencing is the risk of over-representation of sequences introduced by the whole-transcriptome amplification rather than by true biological differences. The amplification of RNA by the polymerase chain reaction (PCR) is non-linear with more abundant sequences likely to be amplified more than less abundant species which can lead to biases in the results (Aird et al., 2011). The FastQC software provides an easy automated method to estimate duplication bias in sequencing data and indicates which sequences may be present at higher levels due to this duplication bias. FastQC found a high percentage of sequences that had been duplicated more than 10 times relative to unique sequences indicating a bias was present in the dataset (Figure 57). A BLAST search of the over-represented sequences identified the sequences as ribosomal RNA (Table 29, Table 30). The protocol for library preparation contained steps for the depletion of ribosomal RNA from samples and as such the presence of these sequences in the samples indicated that the depletion did not function correctly. The efficacy of the ribosomal RNA depletion was dependent on the integrity and fragmentation of

the RNA which could have complicated the removal step leading to the incomplete depletion found in the samples. However, as the over-represented sequences were of ribosomal origin this could have indicated that the bulk of sequences which were of interest to this work could have been impacted to a lesser degree by the duplication bias.

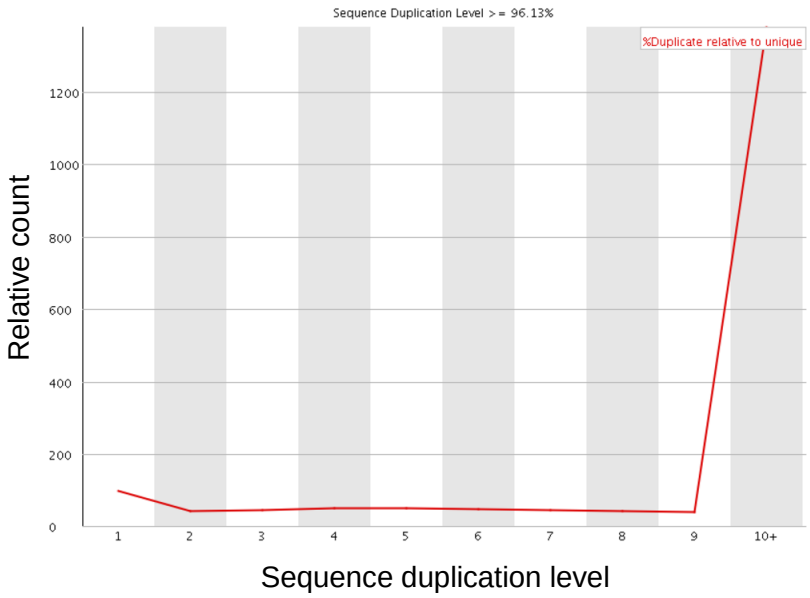


Figure 57: Adapted FastQC duplication plot demonstrating a high number of sequences which were duplicated in excess of 10 times.

Sequence	Count	Percentage of total reads
GTCTACGGCCATACCACCCTGAACGCGCCCGATCTCGTCTGATCTCGGAAGCTAAGCAGGGTCGGGCCTGGTT	214223	0.348337
GGTTTAGTGAGGCCCTCGGATCGGCCCCGCCGGGGTCGGCCCACGGCCCTGGCGGAGCGCTGAGAAGACGGTC	104242	0.169503
CGACTCTTAGCGGTGGATCACTCGGCTCGTGCGTCGATGAAGAACGCAGCTAGCTGCGAGAATTAATGTGAAT	97648	0.158781
AGATCGGAAGAGCACACGTCTGAACTCCAGTCACAGCGATAGATCTCGTATGCCGTCTTCTGCTTGAAAAAAA	87887	0.142909
GAGATCGGAAGAGCACACGTCTGAACTCCAGTCACAGCGATAGATCTCGTATGCCGTCTTCTGCTTGAAAAAA	81317	0.132225
GTAGCCCCGGGAGGAACCCGGGGCCGCAAGTGCGTTCGAAGTGTCGATGATCAATGTGTCCTGCAATTCACAT	72070	0.117189
TAGATCGGAAGAGCACACGTCTGAACTCCAGTCACAGCGATAGATCTCGTATGCCGTCTTCTGCTTGAAAAAA	68868	0.111983

Table 29: Overrepresented sequences detected by FastQC in the sample as in Figure 57. The sequence in bold was subjected to a BLAST search to assess the possible source of the overrepresented sequences.

Description	Identity %	Accession
Homo sapiens RNA, 45S pre-ribosomal 5 (RNA45S5), ribosomal RNA	100%	NR_046235.1
Homo sapiens RNA, 5.8S ribosomal 5 (RNA5-8S5), ribosomal RNA	100%	NR_003285.2

Table 30: Results of BLAST analysis of the bold sequence in Table 29. This indicated that ribosomal sequences are present in the sequenced reads.

Samples which could be classed as poor quality through assessment of library sizes and log2CPM boxplots were therefore excluded from further analysis. This left a mean of 4.32 samples in each experimental group (numbers in each group are stated in Table 31) with a median library size of 703,807 reads. Features with log2CPM values above 1 in at least 75% of samples in each group, i.e. Treated RA $CD90^{-}GP38^{+}$or anti-PD1 $CD163^{+}CD206^{+}$, were retained and all other features discarded due to the low level of expression or lack of expression in enough samples to facilitate analysis. This reduced the number of features being analysed from 58,510 to 17,188. The resulting samples were subjected to VOOM transformation, TMM normalisation, and quantile normalisation and the resulting

log2CPM values were plotted (Figure 58). The boxplots across samples displayed a similar distribution and median indicating the normalisation was successfully able to prepare the data for analysis.

	RA	Treated RA	Anti-PD1
CD34+	4	5	4
CD34-	3	4	4
CD90-GP38+	3	5	5
CD206+CD163+	5	3	4
CD206+CD163-	5	5	5
Double negative	4	4	4

Table 31: RNA sequencing samples remaining in each experimental group after the filtering out of poor quality samples

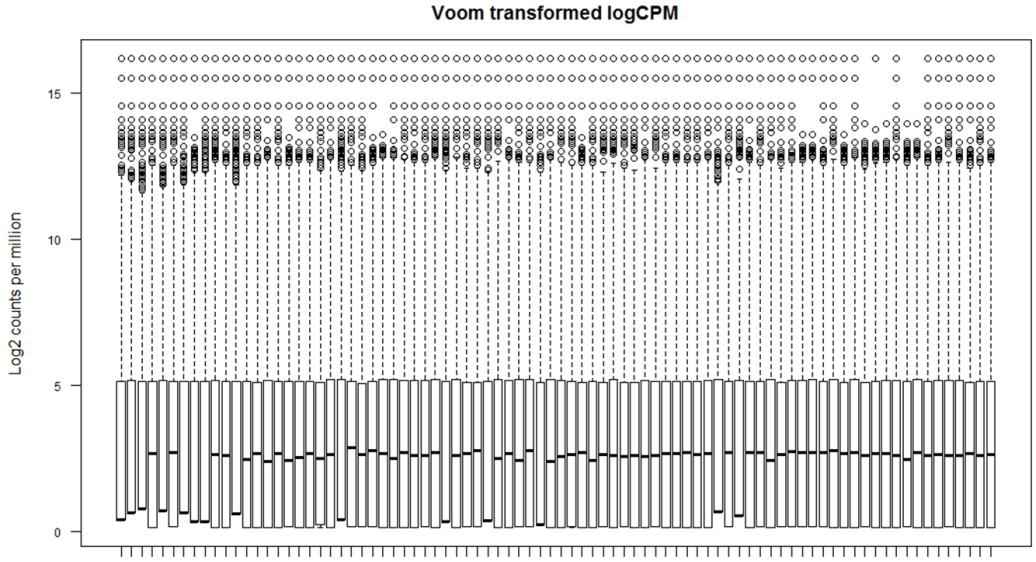


Figure 58: Boxplots of log2CPM values of all features which passed filtering per sample. The samples had been VOOM transformed, TMM normalised, and quantile normalised. The median log2CPM values and interquartile ranges were similar indicating successful normalisation.

After quality control and preparation of the dataset, the sources of variation within the data were assessed to ensure that variation from undesirable sources such as stochastic noise or the batch in which the tissue was digested was not the dominant source of variation. This was important, because such variation may mask variation of interest to the question being addressed. One method to address this issue is to use approaches that aim to reduce the many dimensions

of the data into a few easily interpretable dimensions (Hotelling, 1933; Kruskal, 1964). Approaches such as principal component analysis (PCA) or multidimensional scaling (MDS) are commonly used for this reason. Plots of the first two dimensions of a MDS analysis of the dataset are shown in Figure 59 with each plot coloured according to a different variable. The dominant source of variation was cell type with a clear separation between fibroblast and macrophage populations. There was some possible clustering on the basis of the sex of the tissue donor, however neither site of biopsy nor digest batch appeared to show clear clustering when examining the first two dimensions. Diagnosis also did not appear to show clear clustering in the first two dimensions indicating that the differences between the disease groupings did not contribute as much to the variation within in the data as the cell type did. This was a reassuring finding. Additionally, the control samples generated from tissues formed a clear cluster removed from either the fibroblast or macrophage clusters, an important internal control.

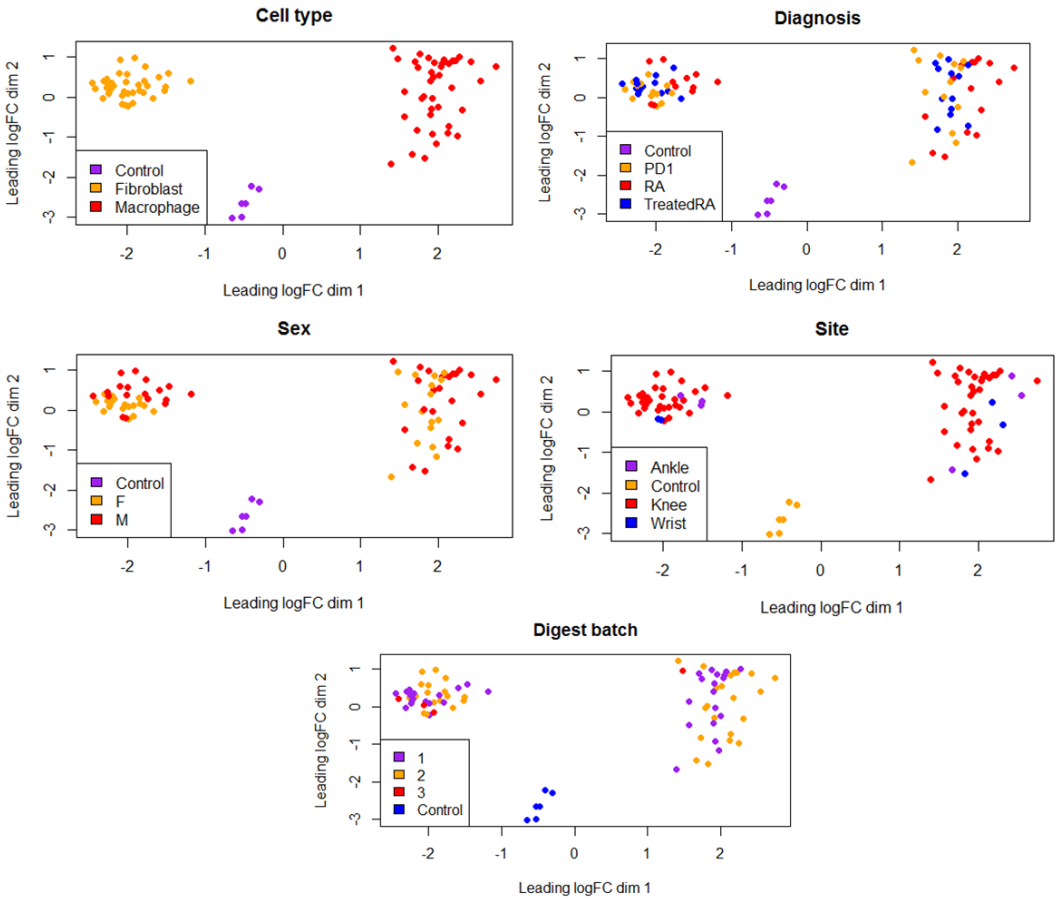


Figure 59: Plots of the first two dimensions of multidimensional scaling (MDS) analysis of the normalised RNA sequencing data coloured according to various variables. The plots indicated that the largest source of variation was cell type and that undesirable noise from sources such as the batch in which the tissue was digested prior to sorting and RNA isolation was minimal.

In addition to being able to identify clear fibroblast and macrophage clusters, additional clusters corresponding to macrophage subsets could also be identified (Figure 60). PCA separated the macrophage samples into two to three clusters that corresponded to the macrophage subsets with the $CD206^{+}CD163^{+}$ (206p163p) and CD206+CD163- (206p163n) showing a higher degree of overlap than either population did with the $CD206^{-}CD163^{-}$ population, although some of the $CD206^{+}CD163^{-}$ samples did appear to cluster with $CD206^{-}CD163^{-}$ samples. Interestingly, the fibroblast subsets did not appear to form clear clusters in the first or second dimensions/components of the unsupervised clustering methods

reflecting a lower degree of heterogeneity between the fibroblast subsets than the macrophage subsets.

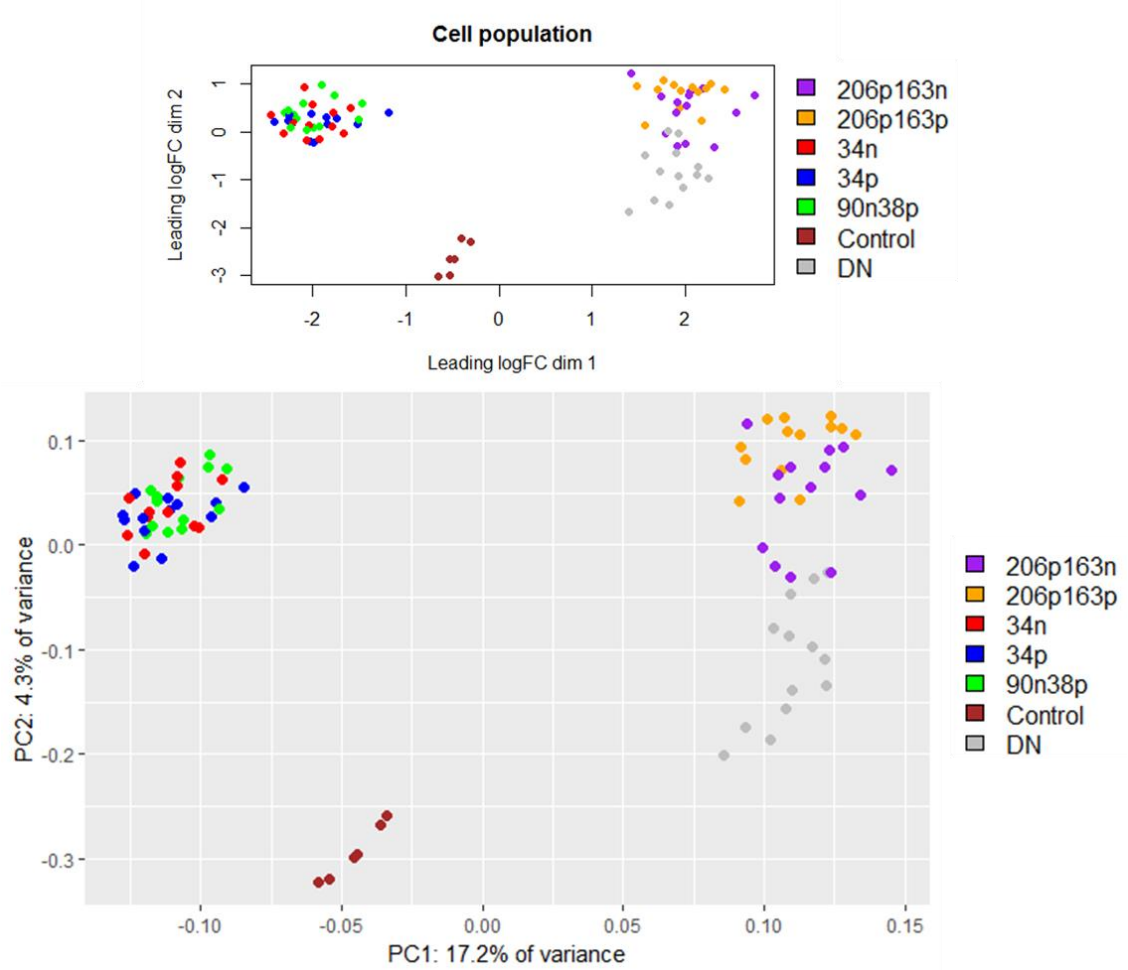


Figure 60: Multidimensional scaling analysis and principal component analysis show similar results highlighting the similarity between the two techniques. The macrophage subpopulations formed two, possibly three, clusters indicating that the subsets had differential gene expression. The sorted fibroblast populations showed a less clear picture. Macrophage populations are $CD206^{+}CD163^{-}$ (206p163n), $CD206^{+}CD163^{+}$ (206p163p), and $CD206^{-}CD163^{-}$ (double negative, DN). Fibroblast populations are $CD34^{+}$ (34p), $CD34^{-}$ (34n), and $CD90^{-}GP38^{+}$ (90n38p).

Methods such as PCA also allow assessment of how much of the total variation in the dataset is accounted for by different variables. The first principal component which clearly separated the fibroblast and macrophage subsets accounted for 17.2% of the variance within the total dataset with the second principal component, which separated the macrophage subsets, accounting for an additional 4.3% of the variance.

The findings so far suggested that in spite of the difficulties of working with such a low input of RNA the data remaining after cleaning and filtering were robust and the largest sources of variance were found between cell type and subsets as expected, and not from noise introduced by technical or stochastic elements. As a final quality control to determine the specificity of the sorting procedure isolating pure populations of cells, the expression of genes for the surface markers used for the sorting strategy were compared between the sorted populations. Figure 61 shows the expression of four of the markers used for the sorting of fibroblast subsets in each sorted population (all clinical outcome groups pooled). As expected the expression of *PECAM1* (CD31) was lower in the fibroblast populations than the macrophages as the stromal populations were sorted on the basis of low CD31 expression whereas the macrophages were not selected using CD31 expression, and are known to express surface CD31 protein (Mckenney, Weiss and Folpe, 2001). Expression of *THY1* (CD90) was also as expected from the flow data as the $CD34^-$ and $CD34^+$ samples (34n or 34p respectively), which were both $CD90^+$, had higher *THY1* expression than the $CD90^-GP38^+$ (90n38p) subset. *PDPN* (GP38) expression was fairly uniform across the fibroblast groups as expected and finally *CD34* expression was clearly highest in the $CD34^+$ (34p) population providing further confirmation that the sorting and sequencing procedures were successful.

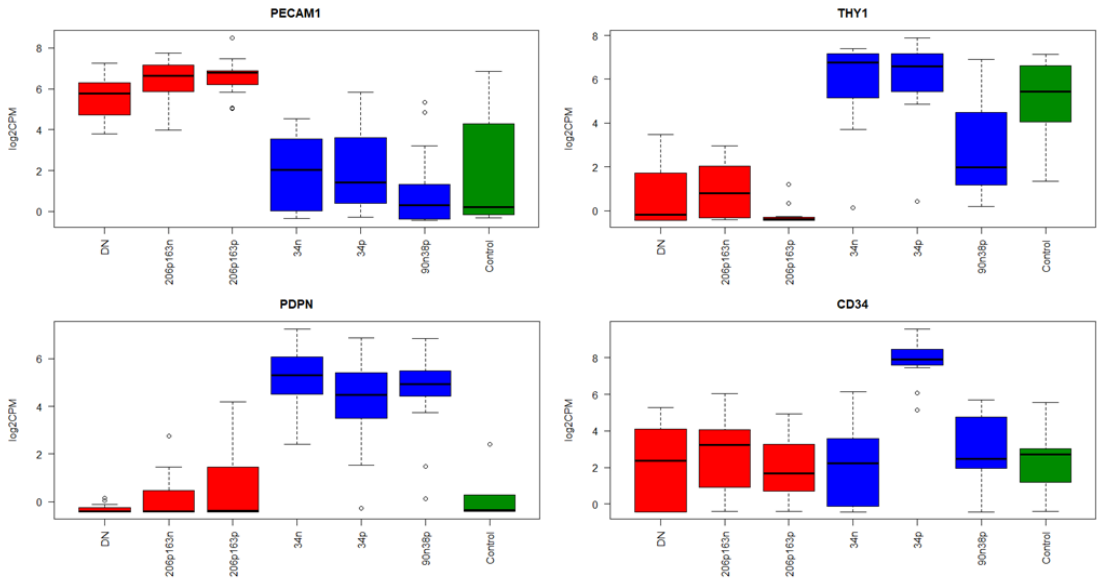


Figure 61: Log2CPM of genes *PECAM1*, *THY1*, *PDPN*, and *CD34*, which encode the proteins used for the stromal sorting strategy , CD31, CD90, GP38, and CD34 respectively. Macrophage populations are coloured red, fibroblast blue, and the sequencing controls green.

Completing the same process but for the markers used for the sorting of macrophage populations also yielded results in keeping with the flow cytometry data (Figure 62). *PTPRC* (CD45) was expressed in all three macrophage populations but at low levels in the fibroblast, *HLA-DRA* followed a similar pattern but was reduced in the DN macrophage population, a finding confirmed by the reduction in HLA-DR surface protein levels as detected by flow cytometry. *CD163* showed the highest levels in the $CD206^{+}CD163^{+}$ (206p163p) population with lower levels in the $CD206^{+}CD163^{-}$ (206p163n) population albeit still higher than the DN population indicating although CD163 surface protein was not detectable in this population the transcript was still expressed. Finally the levels of *MRC1* (CD206) were higher in the $CD206^{+}$ populations but expressed at levels similar to those found in the fibroblast subsets in the DN population.

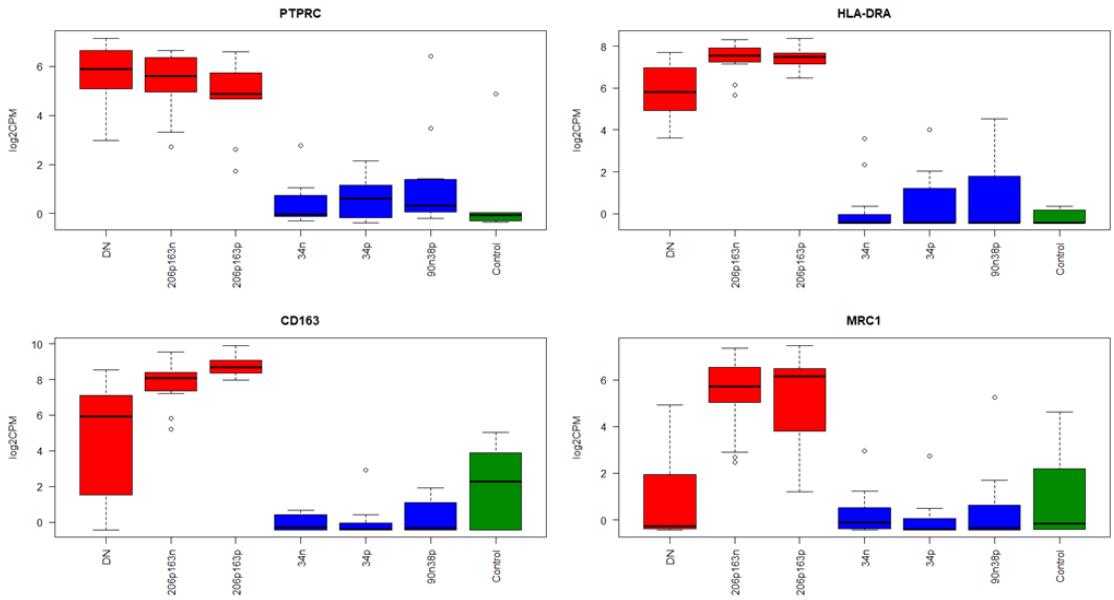


Figure 62: Log2CPM of genes *PTPRC*, *HLA-DRA*, *CD163*, and *MRC1* which encode the proteins used for the stromal sorting strategy , CD45, HLA-DR alpha subunit, CD163, and CD206 respectively. Macrophage populations are coloured red, fibroblast blue, and the sequencing controls green.

Taking the evidence presented in this section it was clear that RNA sequencing from low numbers of cells/low RNA input could be challenging. In this dataset these difficulties resulted in the loss of approximately one third of the samples submitted for sequencing. However, the remaining two thirds of the samples were of sufficient quality to continue the analysis pipeline. When considering the contribution of noise to the variation in the data it was apparent that noise introduced by technical and stochastic elements did not mask the variance of interest contributed by the biology of the samples. Additionally, ‘sanity’ checks of the expression of transcripts encoding the surface proteins used for the sorting procedure confirmed that the sorted populations express the panel of markers expected providing confidence that the sorting and sequencing procedure performed well for these samples. The prepared dataset was then taken forwards for analysis.

5.3.3 Macrophage populations and a mixed lymphocyte population can be identified from the RNA sequencing data

As the largest sources of variation highlighted by PCA of the dataset were between cell type it was prudent to create fibroblast or macrophage subsets of the data and run PCA on these subsets as other sources of variation could have become more apparent due to the increased relative proportion of the variance for which they account. When PCA was completed on the macrophage samples the difference between macrophage subsets became more apparent particularly when investigating the second and third principal components which accounted for 7.23% and 3.97% of the variation in this subsetted dataset (Figure 63). The 95% confidence intervals of the macrophage subsets make it apparent that the DN subset formed a clear cluster separate from either of the other subsets. However, whilst the $CD206^{+}CD163^{+}$ and $CD206^{+}CD163^{-}$ subsets did appear to cluster separately the clusters were close together with a large overlap of confidence intervals. Importantly when assessing the first 10 principal components for an influence of unwanted variation such as a batch effect dependent on when the tissue was digested no clear pattern could be found.

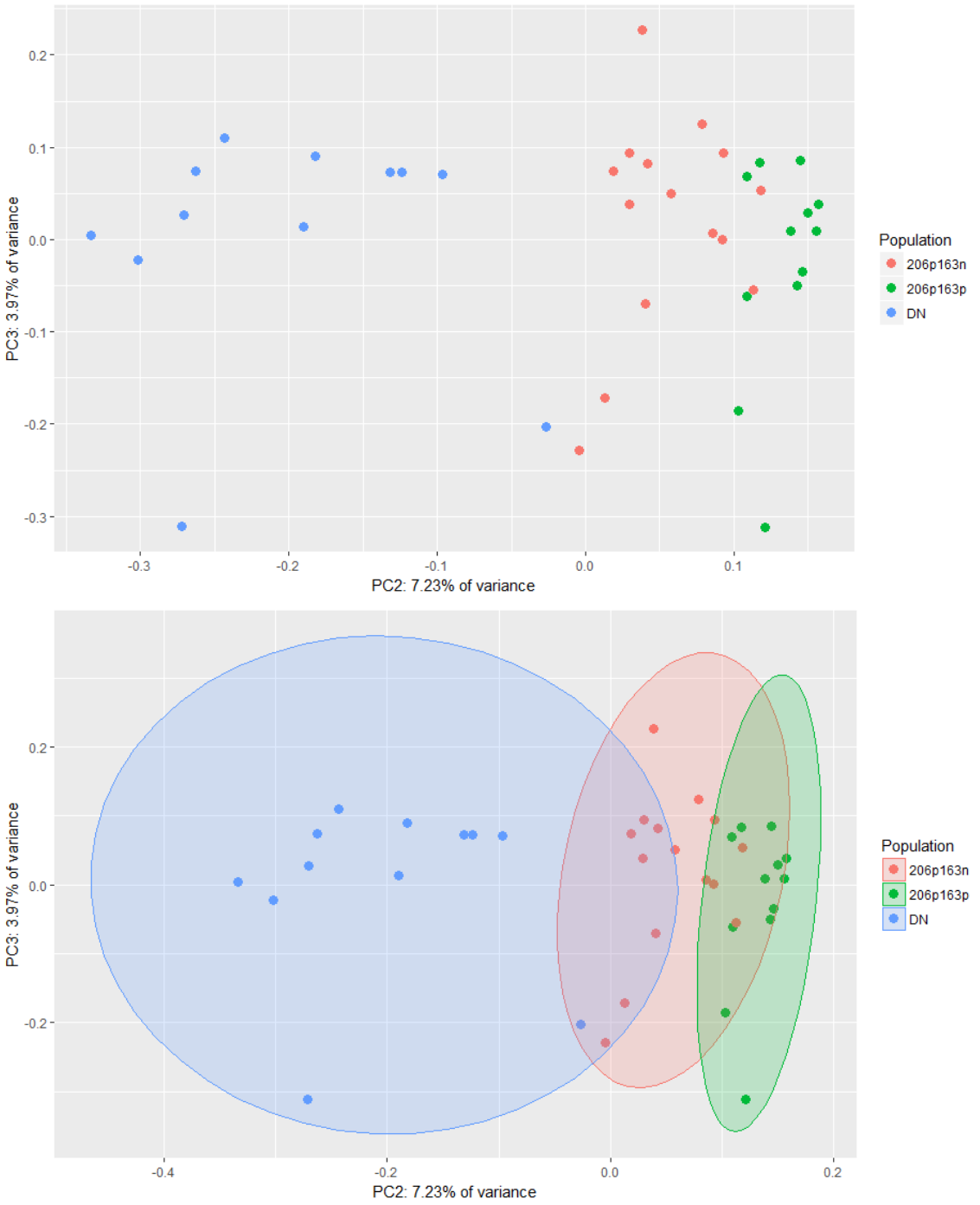


Figure 63: Principal component analysis of the RNA sequencing data generated from the macrophage samples only. The data is coloured according to the macrophage subset from which the sample was generated. The same plot is shown twice, once with just the data points and once with the 95% confidence interval ellipses for each group. The separation between the subsets was apparent with the double negative population showing higher variation and further distance from the other clusters

Hierarchical clustering is another unsupervised machine learning technique that can be used to explore relationships between the samples in the dataset. When using hierarchical clustering a range of distance measures such as Euclidean or Manhattan can be used along with different clustering techniques which may allow other relationships between the data to become apparent (Jaskowiak, Campello and Costa, 2014). Figure 64 shows the results of complete linkage hierarchical clustering of the macrophage dataset using the Pearson's correlation distance measure based on the top 400 most variable genes between macrophage subsets. As with PCA the DN macrophage subset formed a clear cluster irrespective of the clinical outcome group from which the sample was generated. The remaining two subsets were more mixed with clinical outcome appearing to have a larger impact, particularly with clustering of the $CD206^{+}CD163^{+}$ and $CD206^{+}CD163^{-}$ subsets within the RA samples. The $CD206^{+}CD163^{+}$ and $CD206^{+}CD163^{-}$ subsets from anti-PD1 and treated RA patients were intermingled, however each diagnosis often had the two subsets in the same pairs indicating that little variation between these samples existed or that donor specific variation could have been responsible for a significant amount of the variation in the dataset. Indeed when the clustering was re-annotated using the patient identifiers then it was apparent that the pairs were often from the same donor (Figure 65).

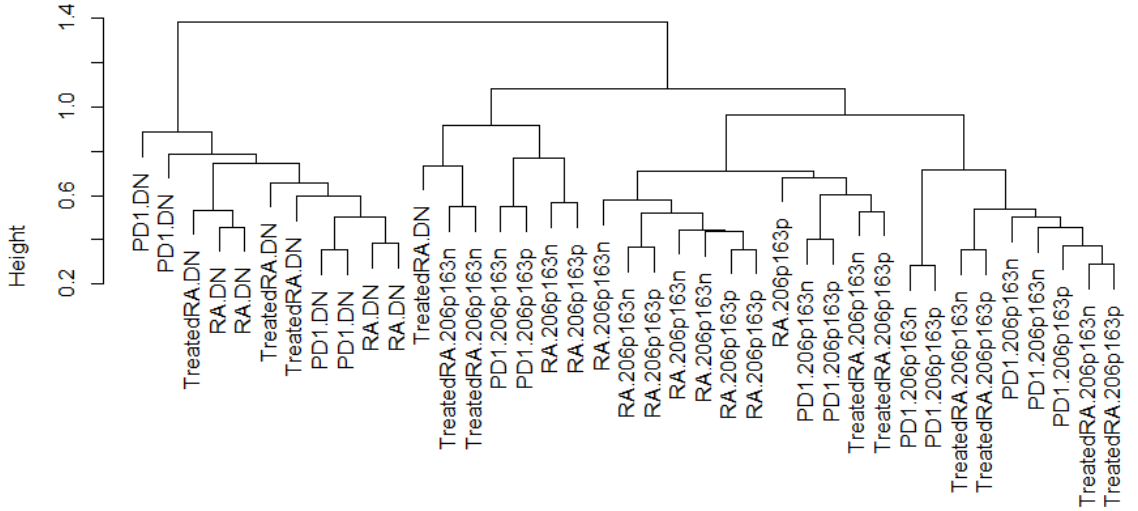


Figure 64: Complete linkage hierarchical clustering using Pearson's correlation distance of the top 400 most variable genes of the macrophage data subset. The double negative macrophage subset (DN) clustered separately from the other samples regardless of the clinical outcome group from which they were isolated. The remaining two subsets were mixed with a cluster of $CD206^{+}CD163^{+}$ (206p163p) and $CD206^{+}CD163^{-}$ (206p163n) RA samples and the same subsets from treated RA and anti-PD1 samples mixed among each other.

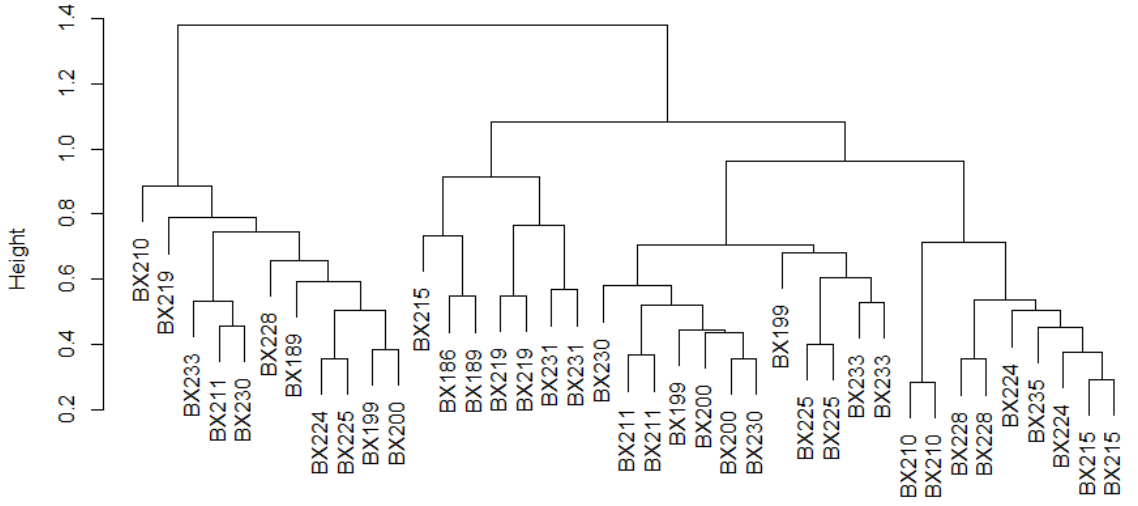


Figure 65: Complete linkage hierarchical clustering using Pearson's correlation distance of the top 400 most variable genes of the macrophage data subset as in Figure 64 but using patient labels as opposed to diagnosis and population labels. It was apparent that $CD206^{+}CD163^{+}$ and $CD206^{+}CD163^{-}$ subsets from the same donor formed pairs within some of the clusters indicating little difference existed between these subsets or an influence of donor specific variation.

The multivariate analysis approaches so far indicated that there were significant differences between the gene expression profiles of the macrophage subsets, particularly between the DN subset and either of the $CD206^+$ subsets. A commonly used approach to detect genes expressed at significantly different levels between experimental groups is to use univariate analysis techniques for each gene and to adjust the P values obtained to avoid a high number of false-positives as a result of repeated comparisons. The approach taken with this dataset was to fit a linear model to the expression value of each gene and use the Benjamini-Hochberg method to correct for multiple comparisons (Benjamini and Hochberg, 1995). The number of statistically significant up- or down-regulated genes between the $CD206^+CD163^+$ and $CD206^+CD163^-$ subsets were relatively small whereas the differences between the $CD206^+CD163^+$ or the $CD206^+CD163^-$ and the DN subset were larger with the same pattern found regardless of the disease group in which the subsets are being compared (Table 32). The number of differentially expressed (DE) genes was largest between the $CD206^+CD163^+$ and the DN subsets with a mean of 428.67 DE genes. The differences between the $CD206^+CD163^-$ and DN subsets were less with 250 DE genes detected and the smallest between the $CD206^+CD163^+$ and the $CD206^+CD163^-$ subsets with only 9.3 DE genes.

	RA	Treated RA	Anti-PD1
CD206+CD163+ vs. CD206+CD163-	3	13	12
CD206+CD163+ vs. DN	537	214	535
CD206+CD163- vs. DN	360	95	295

Table 32: Number of statistically significant differentially expressed genes, both up- and down-regulated, between macrophage subsets in the different disease groupings. Tested using linear modelling, significance is classed as Benjamini-Hochberg adjusted p value <0.05.

In order to investigate the differences between subsets the RA samples will be considered first. As predicted by the results of the unsupervised clustering techniques the differences between the $CD206^+$ subsets were less than between the $CD206^+$ subsets and the DN population with only 3 significantly DE genes detected in the RA samples all of which were downregulated in the $CD206^+CD163^+$ subset compared to the $CD206^+CD163^-$ subset. Of these genes only one encoded a known protein with the remaining two being a non-coding, antisense transcript or encoding an as yet uncharacterised protein making interpretation of the biological significance of the DE genes difficult (Table 33). *VPS37B*, the single characterised protein encoding gene, encodes Vacuolar protein sorting-associated protein 37B which is a component of ESCRT-I, a complex involved in the generation of multivesicular bodies and in the budding and pathogenesis of HIV-1 infection (Katzmann et al., 2003; Shields et al., 2009; Stuchell et al., 2004). The investigation of the $CD163^+$ and $CD163^-$ macrophage subsets found little difference in gene expression despite the apparent clustering by PCA which indicated the data could have been under-powered to detect many statistically significant DE genes and as such an increase in the sample size would be required to investigate this further.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000139722	VPS37B	-4.931897735	0.00581004
ENSG00000273188	(Non-coding, antisense)	-4.963454639	0.008992451
ENSG00000173366	(Protein encoding, uncharacterised)	-4.687845138	0.032134005

Table 33: Significantly differentially expressed genes in the $CD206^+CD163^+$ versus the $CD206^+CD163^-$ subsets from RA samples. Assessed using linear modelling and corrected for multiple comparisons using the Benjamini-Hochberg method.

The differences between the $CD206^{+}CD163^{+}$ and the DN subsets and also between the $CD206^{+}CD163^{-}$ and DN subsets were more profound with 537 and 360 significantly DE genes detected respectively in the RA samples. However despite the different number of genes detected only 3 genes were detected as differentially regulated between these comparisons indicating that the profiles of DE genes between either $CD206^{+}$ subset and the DN subset were similar again reflecting the small difference between the $CD206^{+}CD163^{+}$ and the $CD206^{+}CD163^{-}$ subsets. Examining the DE genes which were upregulated in the $CD206^{+}CD163^{+}$ subset confirmed the macrophage phenotype of these cells with the expression of complement components such as *C1QB*, *C2*, *C1QA*, the LPS co-receptor *CD14*, the collectin *COLEC12,* the scavenger receptor *MARCO*, *SIGLEC1*, the postulated tissue resident macrophage marker *MERTK*, and the anti-inflammatory mediator *TGFB1* (Table 34) (Elomaa et al., 1998; Gautier et al., 2012; Hartung and Hadding, 1983; Jaguin et al., 2013; Xue et al., 2014a). The combination of scavenger receptor and *TGFB1* expression may indicate this subset had a regulatory phenotype.

Ensembl ID	HGNC symbol	Log 2 fold change	Adjusted P value
ENSG00000198865	CCDC152	7.940412	3.02E-07
ENSG00000277517		7.528202	1.16E-09
ENSG00000204706	MAMDC2-AS1	7.449745	3.63E-09
ENSG00000136235	GPNMB	7.334419	1.28E-08
ENSG00000107643	MAPK8	6.719737	7.57E-07
ENSG00000177575	CD163	6.592754	4.65E-06
ENSG00000173369	C1QB	6.318715	2.47E-05
ENSG00000279660		6.304964	4.86E-05
ENSG00000155659	VSIG4	6.177894	4.22E-09
ENSG00000049247	UTS2	6.111618	1.66E-06
ENSG00000153071	DAB2	6.06767	4.05E-06
ENSG00000236467	KCNMA1-AS1	5.967043	0.000228
ENSG00000158270	COLEC12	5.965191	2.62E-06
ENSG00000110077	MS4A6A	5.939815	2.00E-05
ENSG00000170458	CD14	5.887226	1.76E-07
ENSG00000118855	MFSD1	5.773929	8.64E-05
ENSG00000168612	ZSWIM1	5.772565	8.84E-05
ENSG00000236581	STARD13-AS	5.669603	3.88E-05
ENSG00000267199		5.658803	0.00133
ENSG00000093010	COMT	5.646451	2.56E-05
ENSG00000159189	C1QC	5.644668	9.43E-06
ENSG00000280222		5.598552	9.92E-05
ENSG00000278206		5.580433	0.000367
ENSG00000239665		5.565917	0.000161
ENSG00000197093	GAL3ST4	5.41126	0.000249
ENSG00000074964	ARHGEF10L	5.406619	0.000181
ENSG00000274877	NA	5.38718	6.84E-05
ENSG00000255874	LINC00346	5.343534	0.000189
ENSG00000141338	ABCA8	5.317973	0.000807
ENSG00000127980	PEX1	5.303836	0.000221
ENSG00000125818	PSMF1	5.254467	0.000509
ENSG00000092068	SLC7A8	5.231992	0.000197
ENSG00000021355	SERPINB1	5.206477	0.006276
ENSG00000229036	VDAC1P8	5.204373	0.00093
ENSG00000271605	MILR1	5.171967	0.001723
ENSG00000279433		5.00273	6.56E-05
ENSG00000100600	LGMN	4.990873	0.002326
ENSG00000268364	SMC5-AS1	4.964322	0.030741
ENSG00000238133	MAP3K20-AS1	4.957368	0.004526

Table 34: Top 40 significantly upregulated genes in the $CD206^{+}CD163^{+}$ subset versus the double negative subset in RA patients sorted by log2 fold change. P values were adjusted using the Benjamini-Hochberg method.

It is often informative to use gene set enrichment analysis (GSEA) to facilitate the investigation of the functions of genes in known pathways as opposed to considering genes on a one-by-one basis. The Broad Institute molecular signatures database contains many gene sets designed for use in GSEA. Replicate genes in DE gene lists would result in false-positives during the enrichment tests as each occurrence of a gene is counted as a member of a gene set. To prevent this all GSEA and GO category or KEGG pathway analysis completed was run on a dataset in which only single occurrences of either primary assembly or non-primary assembly genes, with a preference for primary assembly genes, were included and which had been re-tested for differential gene expression before enrichment analysis. Using collection 2 of the molecular signatures database for GSEA, which consists of curated gene sets such as REACTOME pathways, the DE genes between the $CD206^{+}CD163^{+}$ and DN subsets in RA patients showed enrichment in 27 gene sets at a false discovery rate of less than 0.05. Further evidence as to the macrophage phenotype of the $CD206^{+}CD163^{+}$ subset was provided in the form of enrichment of genes in the gene sets REACTOME creation of C4 and C2 activators and REACTOME trafficking and processing of endosomal TLR pathways although the number of DE genes in these gene sets was small (Table 35).

Using the molecular signatures database collection 7, which focuses on immunological gene sets generated from published microarray data, also found enrichment in a single gene set for the $CD206^{+}CD163^{+}$ subset (Table 36) (Subramanian et al., 2005). The enriched gene set in the $CD206^{+}CD163^{+}$ subset, GSE24634 Treg vs Tconv post day10 IL4 conversion DN (genes which are

downregulated in CD25+ regulatory T cells compared to CD25- T cells after 10 days of IL-4 treatment of both cell types), did not help to elucidate the phenotype of the macrophage subset however a large number of DE genes were in the gene set and showed a considerable enrichment.

Gene set	Number of DE genes in set	Direction of enrichment	FDR
BIOCARTA_TCYTOTOXIC_PATHWAY	12	Down	4.201710e-05
BIOCARTA_CTL_PATHWAY	12	Down	2.943158e-04
REACTOME_CREATION_OF_C4_AND_C2_ACTIVATORS	5	Up	3.594883e-03
BIOCARTA_TCAPOPTOSIS_PATHWAY	9	Down	5.832759e-03
BIOCARTA_TCRA_PATHWAY	11	Down	1.061436e-02
BIOCARTA_IL17_PATHWAY	11	Down	1.061436e-02
BIOCARTA_CLASSIC_PATHWAY	10	Up	1.589149e-02
BIOCARTA_THELPER_PATHWAY	12	Down	1.598335e-02
REACTOME_TRAFFICKING_AND_PROCESSING_OF_ENDOSOMAL_TLR	10	Up	2.624318e-02
BIOCARTA_NO2IL12_PATHWAY	13	Down	2.792765e-02

Table 35: Results of gene set enrichment in the differentially expressed genes between the CD206^{+}CD163^{+} and double negative subsets in RA patients using the Broad institute MSigDB C2 curated databases. Results displayed are REACTOME and BIOCARTA pathways which have a false discovery rate (FDR) less than 0.05%.

Gene set	Number of DE genes in set	Direction of enrichment	FDR
GSE24634_TREG_VS_TCONV_POST_DAY10_IL4_CONVERSION_DN	167	Up	0.007136543

Table 36: Results of gene set enrichment in the differentially expressed genes between the CD206^{+}CD163^{+} and double negative subsets in RA patients using the Broad institute MSigDB C7 curated databases. Results displayed are pathways which have a false discovery rate (FDR) less than 0.05%.

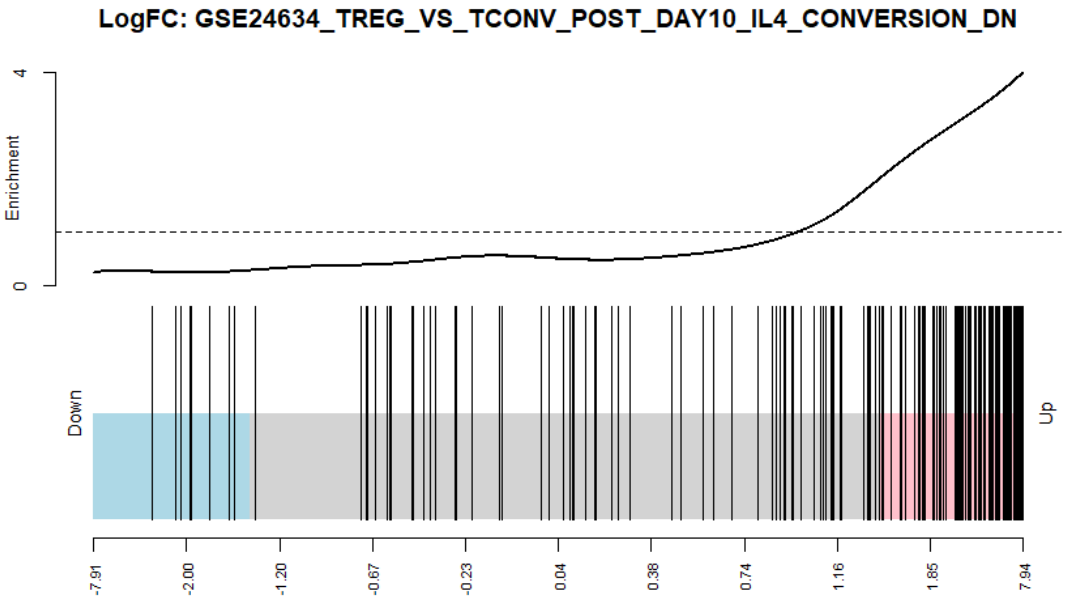


Figure 66: Barcode plot of the gene-set GSE24634 Treg vs Tconv post day10 IL4 conversion DN showing the log2 fold change of genes in the gene set and the relative enrichment score of this gene set in the CD206^{+}CD163^{+} compared to the double negative subset. The x-axis represents log2 fold-change.

A known issue with the detection of differentially expressed genes that may affect GSEA and gene ontology (GO) term enrichment in particular in RNA sequencing data is that the length of genes affects the number of reads generated with longer genes generating more reads (Oshlack and Wakefield, 2009). Using the goseq R package one can estimate bias from gene length and adjust the results of GO term enrichment and KEGG pathway enrichment to account for gene length bias. Using this approached the CD206^{+}CD163^{+} subset was found to be enriched for genes in 51 GO terms for biological processes (Benjamin-Hochberg corrected p value less than 0.05). These GO terms included immune response, complement activation, and myeloid leukocyte activation indicating a myeloid phagocyte phenotype in this subset as expected from the flow cytometry data, however terms such as neutrophil activation and immunoglobulin humoral immune response were also enriched possibly due to the DE genes being members of several gene sets (Table 37). Using the same approach but instead testing for

enrichment in KEGG pathways found one significantly enriched pathway, complement and coagulation cascades (Benjamini-Hochberg adjusted p value 0.01091599).

GO category	Number of DE genes in category	Number of genes in category	GO term	Adjusted p value
GO:0002376	60	1016	immune system process	2.442099e-08
GO:0002443	33	340	leukocyte mediated immunity	2.442099e-08
GO:0006955	49	714	immune response	2.442099e-08
GO:0002252	38	466	immune effector process	5.494062e-08
GO:0045055	33	373	regulated exocytosis	1.343676e-07
GO:0016192	51	867	vesicle-mediated transport	6.335515e-07
GO:0006887	33	409	exocytosis	1.317868e-06
GO:0043299	27	282	leukocyte degranulation	1.629336e-06
GO:0002275	27	284	myeloid cell activation involved in immune response	1.685625e-06
GO:0002444	27	285	myeloid leukocyte mediated immunity	1.685625e-06
GO:0002274	28	311	myeloid leukocyte activation	2.216773e-06
GO:0002283	25	262	neutrophil activation involved in immune response	4.833798e-06
GO:0043312	25	262	neutrophil degranulation	4.833798e-06
GO:0042119	25	263	neutrophil activation	4.833798e-06
GO:0002446	25	266	neutrophil mediated immunity	5.852180e-06
GO:0036230	25	267	granulocyte activation	5.852180e-06
GO:0032940	38	583	secretion by cell	8.136843e-06
GO:0046903	39	616	secretion	9.616950e-06
GO:0002263	27	319	cell activation involved in immune response	9.616950e-06
GO:0002366	27	319	leukocyte activation involved in immune response	9.616950e-06
GO:0001775	36	538	cell activation	1.248697e-05
GO:1902578	59	1248	single-organism localization	2.464135e-05
GO:0044765	57	1186	single-organism transport	2.598041e-05
GO:0045321	33	481	leukocyte activation	2.877428e-05
GO:0051179	87	2317	localization	8.526574e-05
GO:0006810	75	1884	transport	1.894815e-04

GO:0006898	15	138	receptor-mediated endocytosis	4.050583e-04
GO:0050896	98	2856	response to stimulus	4.580411e-04
GO:0051234	75	1933	establishment of localization	5.095624e-04
GO:0072376	8	34	protein activation cascade	6.856729e-04
GO:0006956	7	27	complement activation	1.454987e-03
GO:0002682	30	506	regulation of immune system process	1.903319e-03
GO:0006897	21	312	endocytosis	5.548695e-03
GO:0006958	5	14	complement activation, classical pathway	5.636344e-03
GO:0032501	78	2227	multicellular organismal process	8.096762e-03
GO:0045087	20	285	innate immune response	8.096762e-03
GO:0006952	29	517	defense response	8.096762e-03
GO:0006950	55	1343	response to stress	8.096762e-03
GO:0002455	5	16	humoral immune response mediated by circulating immunoglobulin	1.059326e-02
GO:0050776	22	344	regulation of immune response	1.168618e-02
GO:0002576	8	56	platelet degranulation	2.006748e-02
GO:0071402	4	10	cellular response to lipoprotein particle stimulus	2.179581e-02
GO:0006959	8	56	humoral immune response	2.620322e-02
GO:0040011	31	645	locomotion	2.751136e-02
GO:0050866	8	60	negative regulation of cell activation	2.904154e-02
GO:0043277	5	21	apoptotic cell clearance	3.167408e-02
GO:0048870	28	562	cell motility	3.263271e-02
GO:0051674	28	562	localization of cell	3.263271e-02
GO:0044707	71	2055	single-multicellular organism process	3.284563e-02
GO:0051239	42	1011	regulation of multicellular organismal process	3.594518e-02
GO:0055094	4	12	response to lipoprotein particle	4.038486e-02

Table 37: GO biological process categories enriched for in the upregulated genes in the $CD206^{+}CD163^{+}$ subset versus the double negative subset in RA patients. Benjamin-Hochberg adjusted p value cutoff <0.05.

In addition to investigating the genes that are upregulated in the $CD206^{+}CD163^{+}$ subset compared to the DN it was also of interest to investigate genes which are downregulated and hence expressed at higher levels in the DN subset than the

double positive. In the differential expression analysis comparing the two subsets in RA, 318 of the 537 significantly DE genes were found at higher levels in the DN subset than in the double positive, however the genes expressed were difficult to interpret as the NK cell marker *NKG7* was highly expressed compared to the $CD206^{+}CD163^{+}$ subset as were T cell markers such as *CD8A* (log2 fold (log2FC) change of 6.50874), *CD3G* (log2FC 3.499673276), *CD3E* (log2FC 4.692273633), and *CD3D* (log2FC 5.001018657) and B cell genes such as *JCHAIN* (log2FC 5.335226216), *IGHG3* (log2FC 5.160072371), and *IGHM* (log2FC 4.752540244) (Table 38). The lack of obvious macrophage associated genes and expression of mixed lymphocyte genes indicated that the DN population may consist of several $CD45^{+}$HLA-DR^{+} cell types rather than a single population. The results of the previous GSEA (Table 35) found several BIOCARTA pathways which were enriched in the DN subset rather than the $CD206^{+}CD163^{+}$ subset many of which are related to T cell biology such as the BIOCARTA T cytotoxic pathway (Figure 67), BIOCARTA T helper pathway, the BIOCARTA IL-17 (signalling) pathway, and the BIOCARTA TCRA pathway providing further evidence that the DN population contained a T lymphocyte population.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000009694	TENM1	-7.9085	2.19E-11
ENSG00000240535		-7.87094	3.32E-13
ENSG00000267369		-7.52547	7.47E-14
ENSG00000206028		-7.51196	5.35E-05
ENSG00000105374	NKG7	-7.4614	4.83E-16
ENSG00000215039	CD27-AS1	-7.1479	6.40E-11
ENSG00000257957	QRSL1P3	-7.04593	6.39E-07
ENSG00000132464	ENAM	-7.01797	4.16E-10
ENSG00000177272	KCNA3	-7.0163	2.70E-06
ENSG00000245164	LINC00861	-6.85897	7.28E-08
ENSG00000211895	IGHA1	-6.84059	8.91E-06

ENSG00000276849	TRBC2	-6.78075	6.49E-10
ENSG00000267784		-6.71705	6.59E-06
ENSG00000107742	SPOCK2	-6.67499	3.63E-09
ENSG00000162892	IL24	-6.66409	2.66E-08
ENSG00000213262	VDAC2P3	-6.63303	1.09E-08
ENSG00000280348		-6.54324	1.36E-06
ENSG00000153563	CD8A	-6.50874	7.82E-10
ENSG00000279149		-6.39622	6.14E-10
ENSG00000259326		-6.38794	6.92E-09
ENSG00000280014		-6.36654	4.18E-11
ENSG00000270574		-6.35042	8.85E-06
ENSG00000183508	FAM46C	-6.32771	2.66E-08
ENSG00000267046	E2F3P1	-6.32226	7.47E-14
ENSG00000261054		-6.31559	0.000203
ENSG00000182463	TSHZ2	-6.13248	3.17E-06
ENSG00000266413	NA	-5.98241	4.52E-09
ENSG00000173598	NUDT4	-5.98143	5.49E-06
ENSG00000261346		-5.97113	2.69E-06
ENSG00000120875	DUSP4	-5.94401	3.83E-06
ENSG00000235300		-5.9211	8.78E-09
ENSG00000237286		-5.90329	0.000862
ENSG00000115661	STK16	-5.90014	7.85E-05
ENSG00000231858		-5.86973	3.04E-06
ENSG00000270956		-5.72055	2.07E-07
ENSG00000227217		-5.69132	2.66E-10
ENSG00000231799	PA2G4P6	-5.64528	3.39E-06
ENSG00000245648		-5.64314	0.000259
ENSG00000250751		-5.63939	0.000203

Table 38: Top 40 significantly downregulated genes in the CD206^{+}CD163^{+} subset versus the double negative subset in RA patients sorted by log2 fold change. P values were adjusted using the Benjamini-Hochberg method.

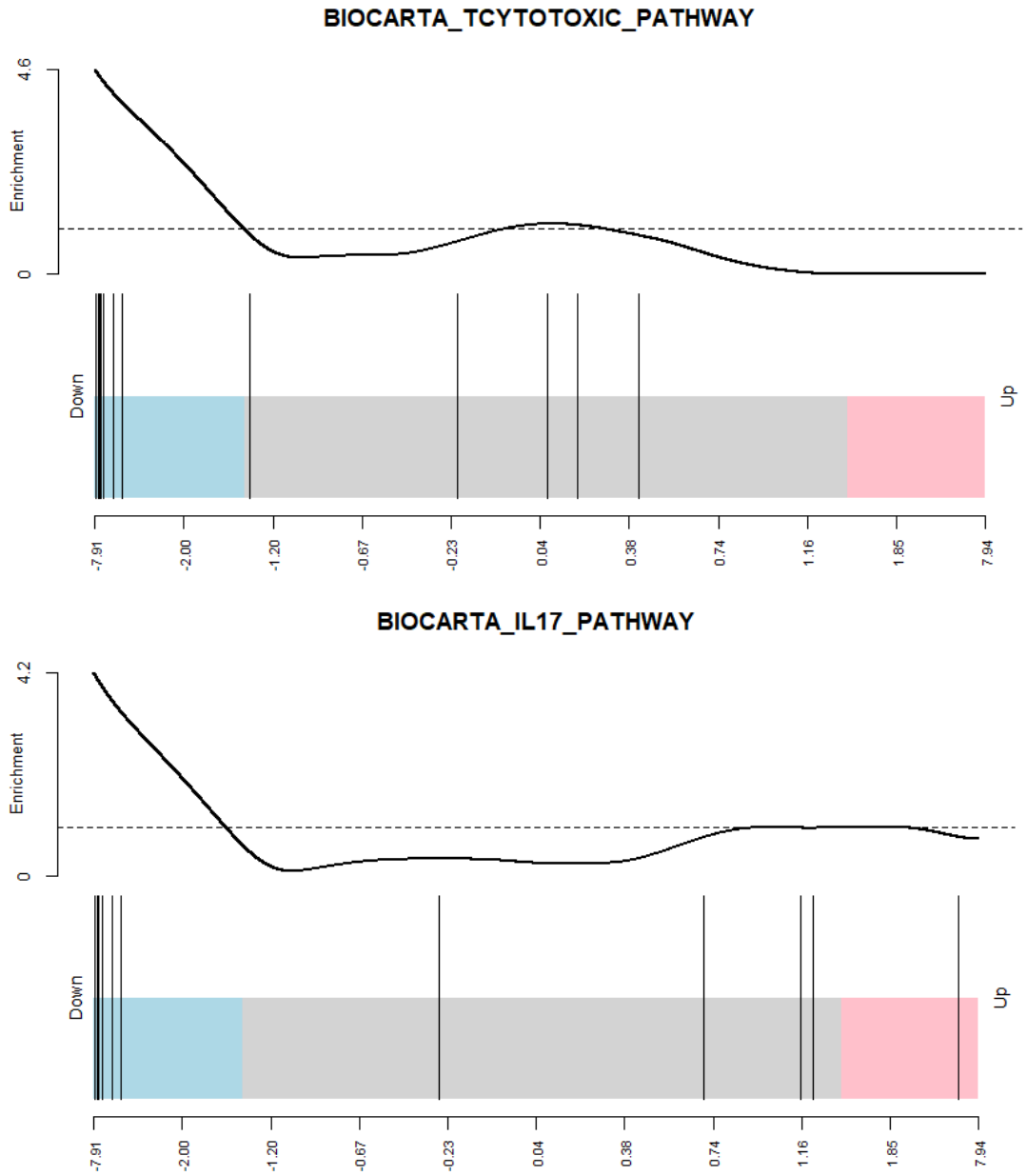


Figure 67: Barcode plots showing the log2 fold change of genes in the gene sets BIOCARTA Tcytotoxic pathway and GSE22886 IL17 pathway and the relative enrichment score of these gene sets in the CD206^{+}CD163^{+} compared to the double negative subset. The x-axis represents log2 fold-change.

Examining enrichment in GO biological process categories as before but substituting the list of DE genes for those higher in the DN subset revealed 65 significantly enriched GO categories (Table 39). Of the enriched GO categories 12 were specific for aspects of T cell biology with no other specific immune cell type explicitly stated in the remaining GO category titles. The remaining GO

categories were aspects of leukocyte and lymphocyte biology such as lymphocyte costimulation, lymphocyte activation, and leukocyte cell-cell adhesion. Using the same approach but instead testing for enrichment in KEGG pathways found four pathways enriched for in the DN upregulated genes, the T cell receptor signalling pathway (adjusted p value 0.0001147660), Natural Killer cell mediated cytotoxicity (adjusted p value 0.0004553556), primary immunodeficiency (adjusted p value 0.0004553556), and hematopoietic cell lineage (adjusted p value 0.0024522829).

GO category	Number of DE genes in category	Number of genes in category	GO term	Adjusted p value
GO:0042110	26	181	T cell activation	7.307012e-07
GO:0031294	10	22	lymphocyte costimulation	2.242512e-06
GO:0031295	10	22	T cell costimulation	2.242512e-06
GO:0046649	28	239	lymphocyte activation	4.607952e-06
GO:0006955	45	552	immune response	6.519748e-06
GO:0007159	19	123	leukocyte cell-cell adhesion	2.033754e-05
GO:1903037	18	113	regulation of leukocyte cell-cell adhesion	3.000701e-05
GO:1903039	15	79	positive regulation of leukocyte cell-cell adhesion	3.881554e-05
GO:0030217	16	96	T cell differentiation	5.451514e-05
GO:0050863	17	109	regulation of T cell activation	7.349036e-05
GO:0022409	15	85	positive regulation of cell-cell adhesion	7.349036e-05
GO:0050870	14	75	positive regulation of T cell activation	1.026148e-04
GO:0051249	19	143	regulation of lymphocyte activation	1.355637e-04
GO:0022407	18	130	regulation of cell-cell adhesion	1.355637e-04
GO:0045321	31	353	leukocyte activation	1.730469e-04
GO:0030098	17	124	lymphocyte differentiation	1.993927e-04
GO:0002376	52	800	immune system process	2.182536e-04
GO:0045785	17	125	positive regulation of cell adhesion	2.955877e-04
GO:0002696	15	99	positive regulation of	3.640167e-04

			leukocyte activation	
GO:0002694	19	157	regulation of leukocyte activation	3.640167e-04
GO:0050867	15	102	positive regulation of cell activation	4.930454e-04
GO:0050865	19	164	regulation of cell activation	6.216705e-04
GO:0002682	33	416	regulation of immune system process	6.216705e-04
GO:0051251	14	93	positive regulation of lymphocyte activation	7.993189e-04
GO:0001775	31	394	cell activation	1.076730e-03
GO:0016337	19	172	single organismal cell-cell adhesion	1.076730e-03
GO:0050776	25	277	regulation of immune response	1.441857e-03
GO:0002684	26	298	positive regulation of immune system process	1.441857e-03
GO:0002521	18	164	leukocyte differentiation	1.441857e-03
GO:0006952	31	398	defense response	1.486300e-03
GO:0045058	7	23	T cell selection	1.854097e-03
GO:0032944	11	63	regulation of mononuclear cell proliferation	1.916041e-03
GO:0046631	10	54	alpha-beta T cell activation	1.985408e-03
GO:0098602	19	185	single organism cell adhesion	2.512251e-03
GO:0070663	11	65	regulation of leukocyte proliferation	2.751957e-03
GO:0030155	20	208	regulation of cell adhesion	3.797050e-03
GO:0030097	22	248	hemopoiesis	3.797050e-03
GO:0045087	21	235	innate immune response	6.978579e-03
GO:0032943	12	86	mononuclear cell proliferation	6.978579e-03
GO:0042129	9	48	regulation of T cell proliferation	7.065498e-03
GO:0098609	20	222	cell-cell adhesion	7.073226e-03
GO:0045061	5	12	thymic T cell selection	7.073226e-03
GO:0048534	22	260	hematopoietic or lymphoid organ development	7.246462e-03
GO:0050670	10	62	regulation of lymphocyte proliferation	7.934753e-03
GO:0050851	12	89	antigen receptor-mediated signaling pathway	8.168410e-03
GO:0070661	12	88	leukocyte proliferation	8.293147e-03
GO:0050852	11	76	T cell receptor signaling pathway	8.856201e-03
GO:0046632	8	43	alpha-beta T cell differentiation	1.054236e-02
GO:0042098	10	64	T cell proliferation	1.054236e-02

GO:0032609	8	42	interferon-gamma production	1.410477e-02
GO:0002757	16	160	immune response-activating signal transduction	1.533023e-02
GO:0002250	14	128	adaptive immune response	1.538536e-02
GO:0002520	22	276	immune system development	1.625446e-02
GO:0018108	12	102	peptidyl-tyrosine phosphorylation	1.692419e-02
GO:0018212	12	102	peptidyl-tyrosine modification	1.692419e-02
GO:0050778	19	218	positive regulation of immune response	2.015974e-02
GO:0046651	11	85	lymphocyte proliferation	2.160840e-02
GO:0032946	8	46	positive regulation of mononuclear cell proliferation	2.212807e-02
GO:0002253	16	169	activation of immune response	2.550103e-02
GO:0070665	8	47	positive regulation of leukocyte proliferation	2.550103e-02
GO:0002429	13	119	immune response-activating cell surface receptor signaling pathway	2.550103e-02
GO:0042102	7	35	positive regulation of T cell proliferation	2.618942e-02
GO:0002764	16	171	immune response-regulating signaling pathway	2.750859e-02
GO:0032729	6	27	positive regulation of interferon-gamma production	4.677034e-02
GO:0046641	3	5	positive regulation of alpha-beta T cell proliferation	4.874571e-02

Table 39: GO biological process categories enriched for in the downregulated genes in the CD206^{+}CD163^{+} subset versus the double negative subset in RA patients. Benjamin-Hochberg adjusted p value cutoff <0.05.

The results of this analysis indicated that the CD206^{+}CD163^{+} subset was a macrophage population with a possible scavenger phenotype as opposed to an overtly pro-inflammatory phenotype as evidence by the enrichment in GO categories such as apoptotic cell clearance. In contrast the identity of the DN subset was unclear with significantly higher expression of T cell, B cell, and NK

cell genes in this subset compared to the $CD206^{+}CD163^{+}$ subset. When testing for enrichment in gene-sets, GO categories, and KEGG pathways, T cell related pathways were abundant indicating that the DN population consisted predominantly of T lymphocytes.

5.3.4 Stromal subsets demonstrate differential gene expression

Using the approach demonstrated in the previous section the differences between fibroblast subsets appeared less profound than those between the macrophage subsets. Principal component analysis found some separation between fibroblast subsets when examining the second and third components however the clusters overlap between 95% confidence interval ellipses (Figure 68). However the $CD34^{+}$ and $CD90^{-}GP38^{+}$ subsets showed a larger degree of separation than that of either subset with the $CD34^{-}$ samples. The fibroblast subset groups also clustered together loosely indicating higher heterogeneity between samples than was found between the macrophage samples. A feature that was not seen in the macrophage data was the presence of an experimental batch effect dependent on the digestion batch of samples with possible clusters visible (Figure 69). Due to careful experimental design this batch effect was not confounded with the disease groupings and so the effects on the analysis could be minimised with batch correction techniques if required.

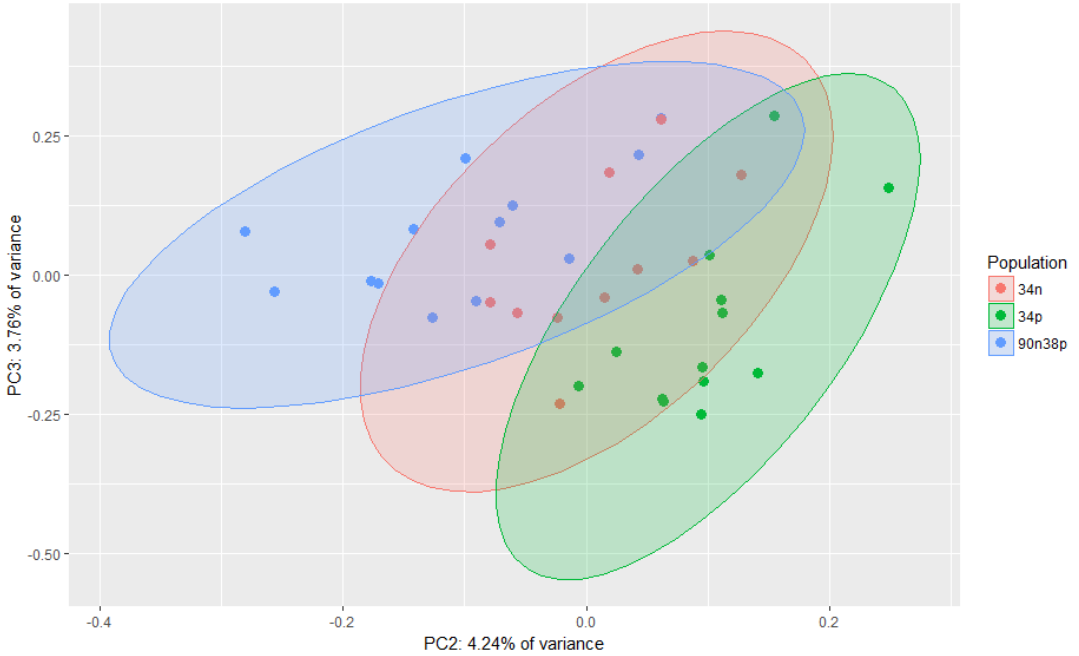


Figure 68: Principal component analysis of the RNA sequencing data generated from the fibroblast samples only. The data is coloured according to the fibroblast subset from which the sample was generated. Ellipses show the 95% confidence interval for each group. Although clusters could be defined the samples within each group clustered loosely and a large amount of overlap was present between ellipses indicating high variation within groups and only small differences between clusters.

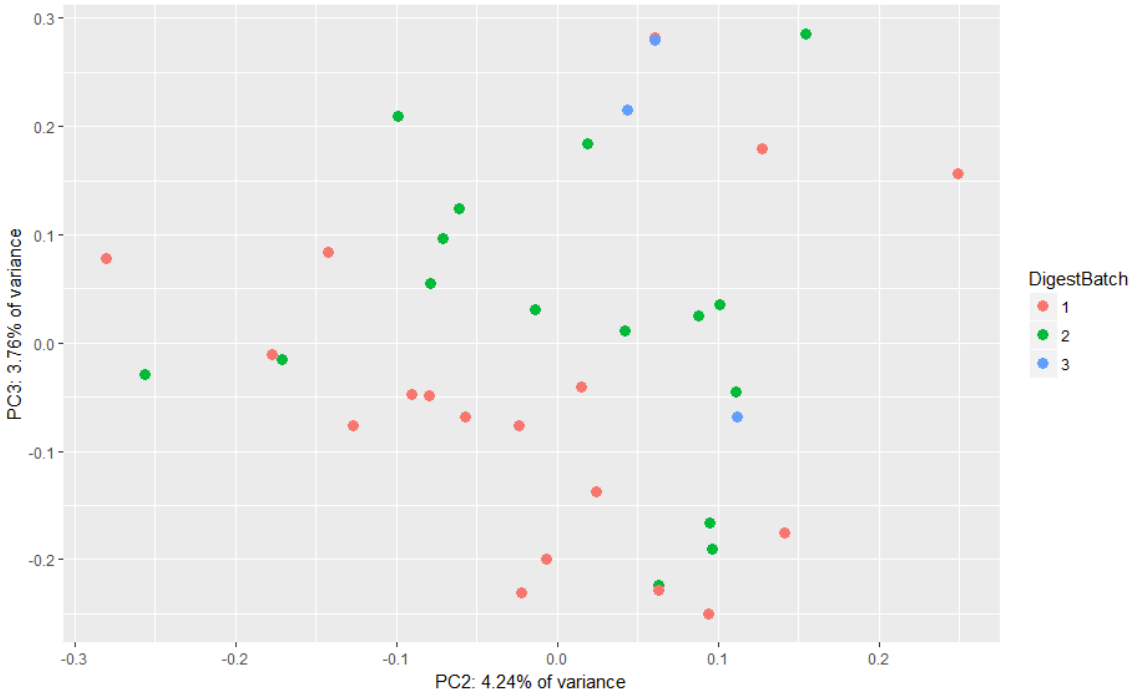


Figure 69: Principal component analysis of the RNA sequencing data generated from the fibroblast samples only. The data is coloured according to the experimental batch in which the biopsy was digested and the sample generated. The presence of a small batch effect could be seen between the experimental batches.

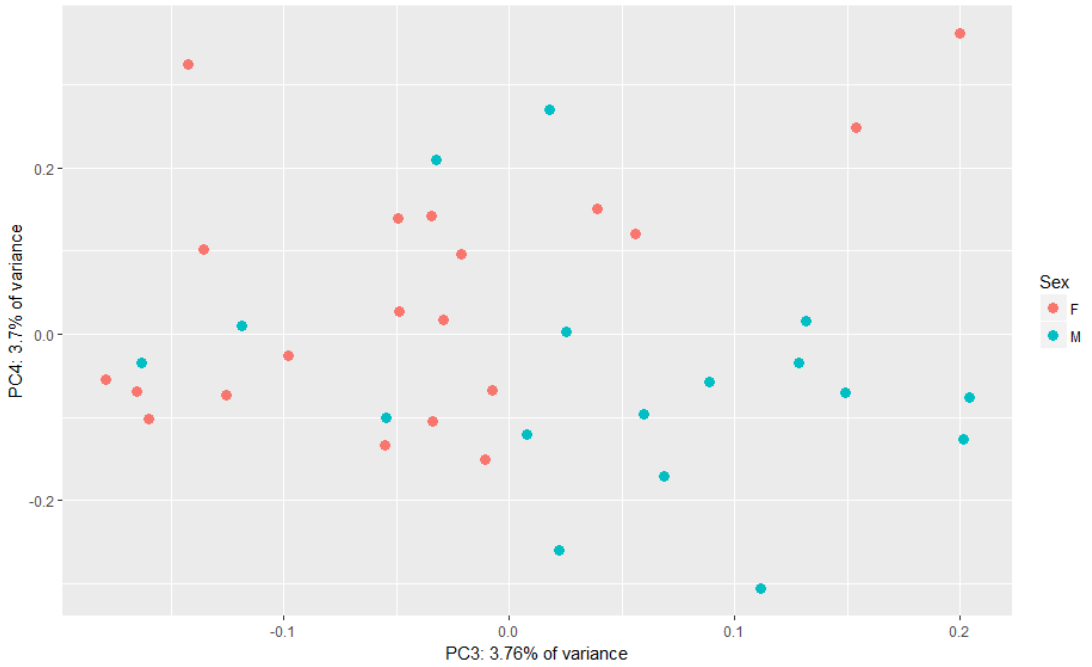


Figure 70: Principal component analysis of the RNA sequencing data generated from the fibroblast samples only. The data is coloured according to the sex of the donor from which the sample was generated. There was some segregation according to sex apparent, although the contribution to the overall variance in the dataset was less than that of the fibroblast subset which the samples belong to.

Hierarchical clustering of the fibroblast subsets using Pearson's correlation distance between samples confirmed the findings of the PCA (Figure 71). Three broad clusters could be seen; a predominantly $CD34^+$ cluster, a mixed $CD34^+$ and $CD34^-$ cluster, and a predominantly $CD90^-GP38^+$ cluster. These clusters consisted of samples from all clinical outcome groups, indicating a higher contribution of the fibroblast subset to the variation in the data than the clinical outcome group, although some clustering according to clinical outcome could be found within the three subset clusters. As with the macrophage dataset, in some cases samples could be seen clustering with other samples from the same donor which could be explained by high homogeneity between subset samples, which seems unlikely given the ability of PCA and hierarchical clustering to find subset clusters, or more likely by a contribution of donor-specific variation to the dataset (Figure 72).

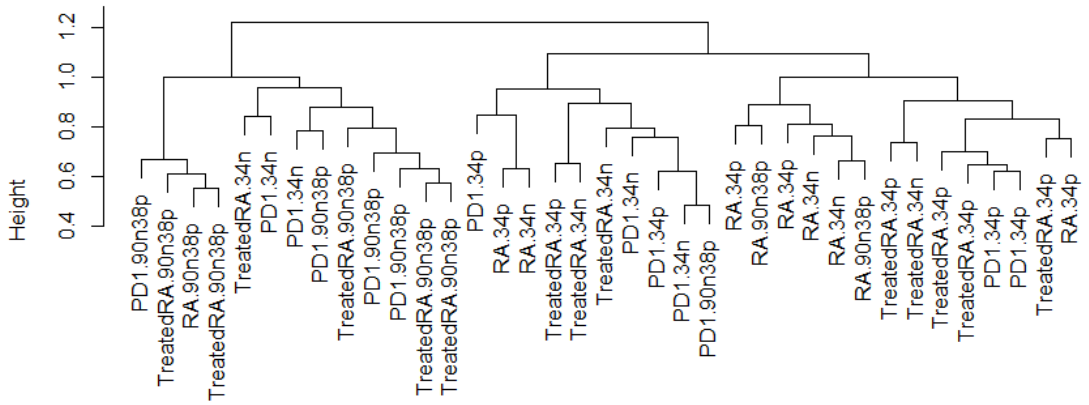


Figure 71: Complete linkage hierarchical clustering using Pearson's correlation distance of the top 400 most variable genes of the fibroblast data subset. Three general clusters could be found; one mostly consisting of the CD34+ sample subset, one consisting of a mixed set of subset samples but broadly containing $CD34^{+}$ and $CD34^{-}$ samples, and one mostly consisting of the $CD90^{-}GP38^{+}$ subset.

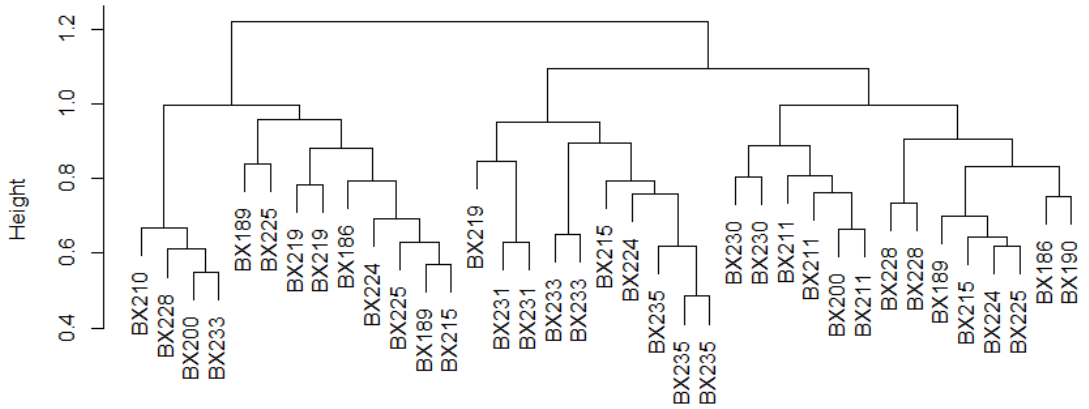


Figure 72: Complete linkage hierarchical clustering using Pearson's correlation distance of the top 400 most variable genes of the fibroblast data subset as in Figure 71 but using patient labels as opposed to diagnosis and population labels. As with the macrophage samples some samples clustered with other samples from the same donor indicating either small differences between these subsets or significant inter-donor variability.

The number of differentially expressed genes between fibroblast subsets were smaller than those between the macrophage subsets with the greatest number of DE genes found between the $CD34^{+}$ and the $CD90^{-}GP38^{+}$ subsets in all clinical outcome groups, as expected from the PCA (Table 40, mean of 108 DE genes). In the RA and anti-PD1 subset comparisons the $CD34^{+}$ and $CD34^{-}$ subsets had the second largest number of DE genes, 60 and 23 respectively, with the lowest

numbers found between the $CD34^-$ and the $CD90^-GP38^+$ subsets. The treated RA samples were different in that the number of DE genes between the $CD34^-$ and $CD90^-GP38^+$ subsets, 19, was larger than that between the $CD34^+$ and the $CD34^-$ subsets, 12.

	RA	Treated RA	Anti-PD1
CD34+ vs. CD34-	60	12	23
CD34+ vs. CD90-GP38+	98	148	78
CD34- vs. CD90-GP38+	16	19	10

Table 40: The number of significantly DE genes, both up- and down-regulated, between fibroblast subsets in each disease group. Detected using linear modelling, significance is <0.05 Benjamini-Hochberg adjusted p-value.

To simplify the analysis of the fibroblast subset data only the comparisons between RA subsets were considered, although the comparisons between subsets in the other disease states are obviously also of interest. When examining DE genes between the $CD34^+$ and the $CD34^-$ subsets in RA *MMP1* and *S100A10* were highly upregulated in the $CD34^-$ subset compared to the $CD34^+$ (Table 41). The $CD34^+$ subset expressed significantly higher levels of the genes *FUT2* and *FUT11* which both encode fucosyltransferases, a family of enzymes of which members have been implicated in the proliferation of synovial fibroblasts, angiogenesis, and in other inflammatory disorders such as Crohn's disease and Behçet's disease (Isozaki et al., 2014a; Isozaki et al., 2014b; Mcgovern et al., 2010; Tsou et al., 2013; Xavier et al., 2015).

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000157570	TSPAN18	6.29528	0.00153
ENSG00000223949	ROR1-AS1	5.59664	0.00295
ENSG00000176920	FUT2	5.51965	0.03444
ENSG00000198522	GPN1	5.41003	0.04565
ENSG00000066422	ZBTB11	5.4	0.01885

ENSG00000270300	PHACTR2P1	5.29583	0.04565
ENSG00000196968	FUT11	5.02206	0.01885
ENSG00000224468		4.98012	4.18E-06
ENSG00000144476	ACKR3	4.90388	0.03444
ENSG00000104611	SH2D4A	4.83396	0.04416
ENSG00000135437	RDH5	4.79762	0.04529
ENSG00000136816	TOR1B	4.62085	0.03943
ENSG00000166226	CCT2	4.57854	0.03196
ENSG00000140848	CPNE2	4.52478	0.04565
ENSG00000170473	PYM1	4.51695	0.03674
ENSG00000149646	CNBD2	4.17163	0.02571
ENSG00000237413		3.98405	0.02827
ENSG00000229702	NA	3.82248	0.00153
ENSG00000244171	PBX2P1	3.08492	0.04416
ENSG00000239823		2.86532	0.04219
ENSG00000254102		-3.22256	0.04219
ENSG00000130707	ASS1	-4.36006	0.03444
ENSG00000105856	HBP1	-4.49571	0.03444
ENSG00000160991	ORAI2	-4.5482	0.03444
ENSG00000131116	ZNF428	-4.55606	0.03943
ENSG00000275871		-4.62341	0.04219
ENSG00000246174	KCTD21-AS1	-4.6509	0.03943
ENSG00000249176		-4.82239	0.02779
ENSG00000166579	NDEL1	-4.97747	0.01956
ENSG00000226706		-4.98732	0.04565
ENSG00000258648	UBE2CP1	-5.05178	0.04565
ENSG00000261222		-5.16108	0.01885
ENSG00000066654	THUMPD1	-5.16355	0.03943
ENSG00000197747	S100A10	-5.181	0.01885
ENSG00000259627		-5.18938	0.03666
ENSG00000255164		-5.2016	0.03157
ENSG00000269480		-5.22547	0.01885
ENSG00000251593	MSNP1	-5.24033	0.02276
ENSG00000197982	C1orf122	-5.29129	0.03302
ENSG00000267139		-5.30723	0.01514
ENSG00000267439		-5.43019	0.0192
ENSG00000249971		-5.43441	0.01885
ENSG00000226152		-5.483	0.00153
ENSG00000228490		-5.483	0.00153
ENSG00000144747	TMF1	-5.51962	0.02453
ENSG00000229781		-5.5442	0.04435
ENSG00000229044		-5.57602	0.03786
ENSG00000166171	DPCD	-5.62691	0.01527
ENSG00000100036	SLC35E4	-5.6297	0.01148
ENSG00000279649		-5.70367	0.01956
ENSG00000213839	TMX2P1	-5.7482	0.01956
ENSG00000255959		-5.79059	0.03943
ENSG00000279364		-5.88565	0.01527

ENSG00000226736		-5.95942	6.46E-05
ENSG00000231197		-5.95942	6.46E-05
ENSG00000028839	TBPL1	-6.0775	0.00762
ENSG00000224721		-6.30773	0.0003
ENSG00000267776		-6.58257	0.00153
ENSG00000255282	WTAPP1	-6.73487	0.01106
ENSG00000196611	MMP1	-7.13899	0.00013

Table 41: Significantly differentially expressed genes between the CD34$^+$ subset and the CD34$^-$ subset in RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

Investigating the genes which were DE between the CD34$^+$ and the CD90$^-$GP38$^+$ subsets found some genes in common with those DE between the CD34$^+$ and CD34$^-$ subset such as *MMP1* and *NDEL1* with higher expression of both in the CD90$^-$GP38$^+$ subset again indicating a downregulation of these genes in the CD34$^+$ subset (Table 42). The CD34$^+$ subset expressed higher levels of *IGF1*. Interestingly, the levels of *CAV1*, the gene encoding caveolin-1, were significantly lower in CD34$^+$ than CD90$^-$GP38$^+$ synovial fibroblasts and *CTSK*, encoding cathepsin K, was also detected as DE between the two subsets with higher expression in the CD90$^-$GP38$^+$ subset which could be important given the degradative nature of cathepsins.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000017427	IGF1	7.657472	3.14E-07
ENSG00000181800	CELF2-AS1	7.278109	1.72E-06
ENSG00000159596	TMEM69	6.134911	0.005201
ENSG00000277401	TJP1	6.052997	0.000327
ENSG00000183562		5.9672	0.006424
ENSG00000255565		5.922744	0.010009
ENSG00000254122	PCDHGB7	5.77401	0.022256
ENSG00000254192		5.719499	0.002031
ENSG00000234338		5.623439	0.016066
ENSG00000213222		5.610919	0.001388
ENSG00000215277	RNF212B	5.420674	0.005201
ENSG00000270300	PHACTR2P1	5.418524	0.025784
ENSG00000248254		5.212413	0.015669

ENSG00000104067	TJP1	5.211505	0.002092
ENSG00000011465	DCN	5.1122	0.010487
ENSG00000151748	SAV1	5.102822	0.014444
ENSG00000196396	PTPN1	5.077432	0.012993
ENSG00000076053	RBM7	4.961398	0.005201
ENSG00000260111		4.929137	0.044998
ENSG00000205323	SARNP	4.920461	0.044059
ENSG00000255390		4.917876	0.034117
ENSG00000230552		4.826452	0.040382
ENSG00000223949	ROR1-AS1	4.810049	0.009115
ENSG00000116729	WLS	4.745186	0.022256
ENSG00000233682		4.743805	0.003225
ENSG00000224468		4.743153	5.27E-06
ENSG00000107175	CREB3	4.693984	0.016405
ENSG00000143387	CTSK	4.676193	0.044699
ENSG00000235033		4.646065	0.024828
ENSG00000166025	AMOTL1	4.632639	0.042987
ENSG00000136816	TOR1B	4.618524	0.026483
ENSG00000232284	GNG12-AS1	4.615269	0.044059
ENSG00000254418	SPON1-AS1	4.513102	0.03054
ENSG00000258114		4.487902	0.017735
ENSG00000270607		4.460797	0.025784
ENSG00000108010	GLRX3	4.45456	0.044699
ENSG00000269950		4.418677	0.025457
ENSG00000174348	PODN	4.344878	0.022256
ENSG00000272235		4.309307	0.032152
ENSG00000196586	MYO6	4.300711	0.043001
ENSG00000149646	CNBD2	4.298336	0.013373
ENSG00000134574	DDB2	4.258815	0.033452
ENSG00000242888		4.233996	0.049981
ENSG00000277288	C10orf142	4.224141	0.014444
ENSG00000229702	NA	4.147683	0.000145
ENSG00000229996		4.022207	0.030313
ENSG00000234557		3.894989	0.045816
ENSG00000237413		3.88366	0.024828
ENSG00000273656		3.638532	0.025457
ENSG00000254634	SMG1P6	3.541354	0.032152
ENSG00000244171	PBX2P1	3.082589	0.030313
ENSG00000138135	CH25H	3.038581	0.036013
ENSG00000239823		2.862988	0.029142
ENSG00000079462	PAFAH1B3	-3.85193	0.048117
ENSG00000267563		-4.1701	0.040382
ENSG00000253523		-4.28097	0.044059
ENSG00000112640	PPP2R5D	-4.29059	0.044699
ENSG00000105974	CAV1	-4.36605	0.025457
ENSG00000226627	SHANK2-AS1	-4.44459	0.026059
ENSG00000104325	DECR1	-4.4609	0.031188

ENSG00000246174	KCTD21-AS1	-4.49705	0.049643
ENSG00000232406		-4.64444	0.026059
ENSG00000107736	CDH23	-4.64875	0.034117
ENSG00000226152		-4.7475	0.022256
ENSG00000228490		-4.7475	0.022256
ENSG00000197046	SIGLEC15	-4.85683	0.034117
ENSG00000039319	ZFYVE16	-4.92921	0.024828
ENSG00000111276	CDKN1B	-5.00785	0.025457
ENSG00000271454		-5.00822	0.017735
ENSG00000089693	MLF2	-5.022	0.044059
ENSG00000107164	FUBP3	-5.06468	0.039549
ENSG00000196182	STK40	-5.14848	0.023132
ENSG00000155265	GOLGA7B	-5.14905	0.036279
ENSG00000078795	PKD2L2	-5.15751	0.037115
ENSG00000109323	MANBA	-5.15951	0.026059
ENSG00000134532	SOX5	-5.18668	0.022256
ENSG00000261546		-5.18779	0.025457
ENSG00000166579	NDEL1	-5.22231	0.012916
ENSG00000267139		-5.27995	0.015669
ENSG00000229127		-5.29722	0.017735
ENSG00000142634	EFHD2	-5.3097	0.026098
ENSG00000272078		-5.47999	0.010487
ENSG00000267439		-5.51407	0.015669
ENSG00000140157	NIPA2	-5.53436	0.030313
ENSG00000101126	ADNP	-5.55212	0.025457
ENSG00000196611	MMP1	-5.60814	0.042925
ENSG00000119760	SUPT7L	-5.68738	0.002851
ENSG00000272072		-5.71974	0.029142
ENSG00000226736		-5.80288	0.000169
ENSG00000231197		-5.80288	0.000169
ENSG00000177409	SAMD9L	-5.83618	0.007714
ENSG00000257150	PGAM1P5	-5.96012	0.00127
ENSG00000232682		-6.26042	0.014922
ENSG00000279083		-6.32014	0.005538
ENSG00000165480	SKA3	-6.33485	0.013783
ENSG00000104154	SLC30A4	-6.91767	0.00292
ENSG00000028839	TBPL1	-7.47584	7.38E-05
ENSG00000224721		-7.56903	3.14E-07

Table 42: Significantly differentially expressed genes between the $CD34^+$ subset and the $CD90^-GP38^+$ subset in RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

The number of DE genes between the $CD34^-$ and the $CD90^-GP38^-$ subsets was relatively small, as a result of direct similarity, or as a result of higher heterogeneity between subjects in these subsets influencing the results of the

linear model. Of the genes detected *CDKN1B* was significantly upregulated in the $CD90^{-}GP38^{+}$ subset which is of interest as this gene has been shown to inhibit fibroblast proliferation and was expressed at lower levels in the peripheral blood of RA patients who have not been treated with methotrexate compared to healthy controls and methotrexate treatment increased *CDKN1B* RNA levels in the peripheral blood of RA patients (Spurlock et al., 2012; Yang et al., 2013) (Table 43). Of the genes expressed at higher levels in the $CD34^{-}$ subset *CDSN* is thought to be a risk allele for the development of RA (Sun et al., 2013).

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000263873		6.457526	0.049351
ENSG00000126861	OMG	6.24403	0.047604
ENSG00000107223	EDF1	5.832043	0.047604
ENSG00000279333		5.789325	0.049351
ENSG00000100036	SLC35E4	5.608836	0.047604
ENSG00000263823		5.566868	0.049351
ENSG00000213222		5.503838	0.047604
ENSG00000275871		5.174837	0.047604
ENSG00000270412		5.001844	0.049351
ENSG00000204539	CDSN	4.358396	0.047604
ENSG00000219790	OSTCP6	-4.53676	0.049351
ENSG00000108639	SYNGR2	-5.34299	0.049351
ENSG00000111276	CDKN1B	-5.44997	0.047604
ENSG00000142634	EFHD2	-5.74409	0.049351
ENSG00000163875	MEAF6	-6.27948	0.047604
ENSG00000279083		-6.32197	0.047604

Table 43: Significantly differentially expressed genes between the $CD34^{-}$ subset and the $CD90^{-}GP38^{-}$ subset in RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

As with the macrophage subsets it was of interest to test if the DE genes between fibroblast subsets were found in certain gene sets and if any enrichment in certain biological functions could be ascribed to the fibroblast subsets. GSEA and KEGG pathway or GO term enrichment was run on all of the DE genes between comparisons to determine if any enrichment was present, but no enrichment was

found. The detection of several genes which have previously been shown to be DE between RA synovial fibroblasts/synovial tissue and control samples indicates that the findings of these studies may be due to the expansion of or upregulation of these genes in specific subsets rather than a general upregulation in all fibroblast populations. This opens up avenues for targeted treatment of specific fibroblast subsets depending on the role they play in the disease.

5.3.5 Stromal and macrophage populations show differences between clinical outcome group

The previous two sections focused on the differences between fibroblast and macrophage subsets within untreated RA samples. It is of great interest to investigate how these subsets vary between diseases as this could possibly allow deconvolution of the complex interactions occurring in these pathologies.

5.3.5.1 Macrophage subsets and clinical outcome groups

Clusters representing the clinical outcome groups of the samples were found when examining the fourth and fifth principal components of the macrophage dataset (Figure 73). Interestingly the RA and anti-PD1 samples formed clear spatially separated clusters with some overlap of the 95% confidence intervals, however the treated RA samples showed much more variation across the principal components. The treated RA samples could be seen both in and out the RA and anti-PD1 clusters whereas the samples in these clusters were almost completely separated from each other.

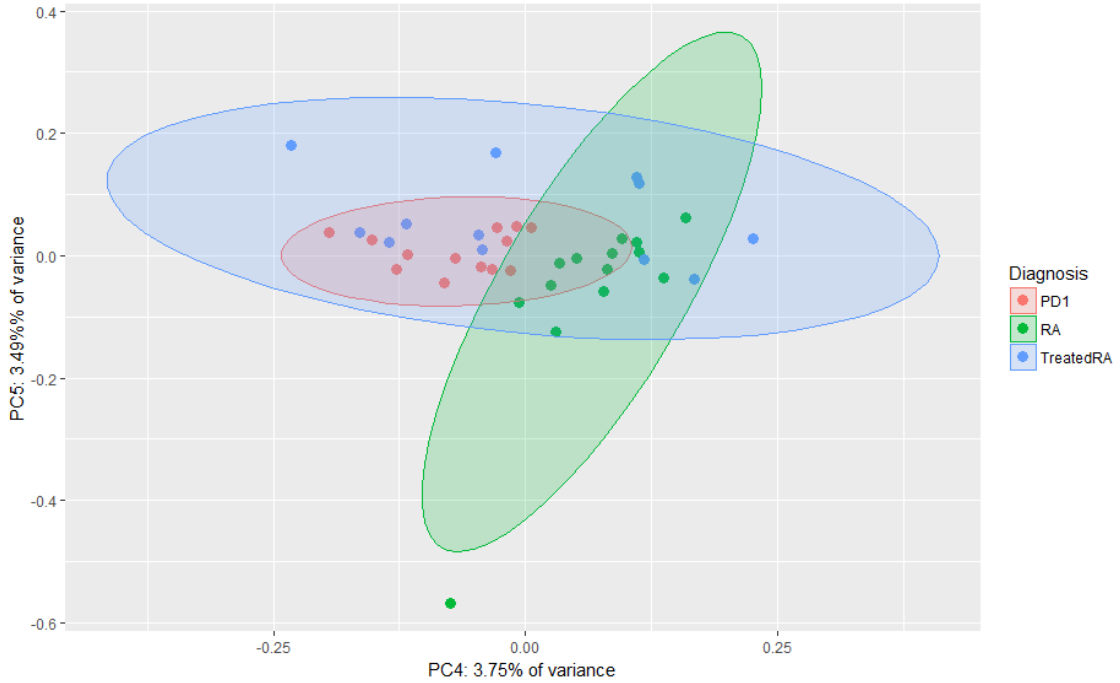


Figure 73: Principal component analysis of the RNA sequencing data generated from the macrophage samples only. The data are coloured according to the clinical outcome group from which each sample was generated. Ellipses are 95% confidence interval ellipses for each group. The RA and anti-PD1 groups formed clear spatial separated clusters with a small amount of overlap in 95% confidence interval ellipses whereas the treated RA cluster was much more variable with the samples spread in and out of the other clusters.

5.3.5.1.1 CD206+CD163+ population

When considering the number of DE genes between the same macrophage subset in different disease groups it was apparent that the differences between them are less than those seen between subsets from the same clinical outcome group (Table 44). The $CD206^{+}CD163^{+}$ subset showed the smallest differences in all disease comparisons (mean of 5 DE genes) whereas the DN subset showed the greatest (mean of 43.67 DE genes) and the $CD206^{+}CD163^{-}$ subset an intermediate number of DE genes (mean of 15.67). The fact that the largest numbers of DE genes were detected in the DN subset could be a reflection of the unclear origin of this subset with multiple possible sources of variation being included.

	CD206+CD163+	CD206+CD163-	Double negative
RA vs Treated RA	3	10	82
RA vs Anti-PD1	6	11	10
Anti-PD1 vs Treated RA	6	26	39

Table 44: The number of significantly DE genes, both up- and down-regulated, between the same macrophage subset in different disease groups. Detected using linear modelling, significance is <0.05 Benjamini-Hochberg adjusted p-value.

Of the 3 downregulated genes in the RA $CD206^{+}CD163^{+}$ subset compared to the same subset in treated RA patients, two genes were uncharacterised and the remaining gene, *TMEM218*, encodes transmembrane protein 218 (Table 45). The comparison of the same subset between the RA and the anti-PD1 samples showed four genes upregulated, of which one gene was uncharacterised, and two genes downregulated in the RA samples (Table 46). The final comparison of the double positive macrophage subset, anti-PD1 patients versus treated RA patients, found six significantly DE genes at higher levels in the treated RA patients three of which were uncharacterised (Table 47). No GO term or KEGG pathway enrichment was found in any of the comparisons.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000279758		-4.53573	0.031289
ENSG00000262412		-5.61484	1.63E-05
ENSG00000150433	TMEM218	-6.74268	5.21E-05

Table 45: Significantly differentially expressed genes in the $CD206^{+}CD163^{+}$ subset in the RA patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000104154	SLC30A4	5.793852	0.029687
ENSG00000131503	ANKHD1	5.643881	0.00324
ENSG00000165660	FAM175B	4.747566	0.01203
ENSG00000228863		4.56843	0.022915
ENSG00000260027	HOXB7	-3.84983	0.022915
ENSG00000166250	CLMP	-4.52859	0.008838

Table 46: Significantly differentially expressed genes in the CD206^{+}CD163^{+} subset in the RA patients compared to the anti-PD1 patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000279758		-4.53695	0.028925
ENSG00000262412		-5.12532	0.003807
ENSG00000140265	ZSCAN29	-5.53273	0.022914
ENSG00000233485		-5.55673	0.006375
ENSG00000131503	ANKHD1	-5.6599	0.006375
ENSG00000225796	MTND4P23	-6.14558	0.000547

Table 47: Significantly differentially expressed genes in the CD206^{+}CD163^{+} subset in the anti-PD1 patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

5.3.5.1.2 CD206^{+}CD163^{-} population

The number of DE genes between the CD206^{+}CD163^{-} subset in different disease groupings was slightly higher than that of the CD206^{+}CD163^{+} subset in the same comparisons. Between the single positive subset in RA patients and treated RA patients there were eight genes upregulated in the RA samples and two upregulated in the treated RA (Table 48). No GO category or KEGG pathway enrichment was found.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000211895	IGHA1	5.674698	0.008564
ENSG00000230325		5.612619	0.00842
ENSG00000204922	UQCC3	4.849422	0.014902
ENSG00000277874	NA	3.915545	0.046671
ENSG00000171860	C3AR1	3.83263	0.019804
ENSG00000277336	CCL3L1	3.262052	0.008564
ENSG00000277796	CCL3L1	3.261793	0.008564
ENSG00000277768	CCL3L1	3.185079	0.014902
ENSG00000104408	EIF3E	-4.28629	0.014889
ENSG00000236824	BCYRN1	-4.86716	0.017267

Table 48: Significantly differentially expressed genes in the CD206^{+}CD163^{-} subset in the RA patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

Continuing with the comparisons of the CD206^{+}CD163^{-} subset across diseases found 6 genes upregulated in the RA samples compared to the anti-PD1 samples and 5 downregulated with no detectable GO category or KEGG pathway enrichment (Table 49).

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000232062	HLA-DQA1	4.76201	0.048233
ENSG00000105851	PIK3CG	4.444169	0.007808
ENSG00000231286	HLA-DQB1	4.334007	0.048233
ENSG00000166501	PRKCB	4.299499	0.048233
ENSG00000170088	TMEM192	4.255366	0.048233
ENSG00000163734	CXCL3	3.943253	0.033567
ENSG00000225496		-3.28804	0.007808
ENSG00000251632	LINC02172	-3.66572	0.048233
ENSG00000163815	CLEC3B	-4.78911	0.013582
ENSG00000028839	TBPL1	-4.87909	0.042892
ENSG00000235033		-5.51636	0.000997

Table 49: Significantly differentially expressed genes in the CD206^{+}CD163^{-} subset in the RA patients compared to the anti-PD1 patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

The comparison of the CD206^{+}CD163^{-} subset between the anti-PD1 samples and the treated RA samples detected 20 genes upregulated and six genes downregulated in the anti-PD1 subset again with no GO category or KEGG pathway enrichment (Table 50).

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000137880	GCHFR	5.973088	0.000206
ENSG00000278879		5.325296	0.018367
ENSG00000198756	COLGALT2	5.253338	0.048806
ENSG00000277287		5.182031	0.020276
ENSG00000168612	ZSWIM1	5.087848	0.036794
ENSG00000235033		4.994644	0.013461
ENSG00000126895	AVPR2	4.963598	0.011044
ENSG00000241043	GVQW1	4.960956	0.043906
ENSG00000119729	RHOQ	4.678253	0.037364
ENSG00000142046	TMEM91	4.539908	0.043906
ENSG00000230733		4.515823	0.045574
ENSG00000179598	PLD6	4.479462	0.036794
ENSG00000246174	KCTD21-AS1	4.422779	0.013461
ENSG00000245522		4.107623	0.048806
ENSG00000255441		3.993933	0.048806
ENSG00000183798	EMILIN3	3.892043	0.048806
ENSG00000158869	FCER1G	3.857084	0.036794
ENSG00000137393	RNF144B	3.396949	0.048806
ENSG00000064601	CTSA	3.053456	0.04918
ENSG00000225496		3.025847	0.018367
ENSG00000279687		-3.22937	0.048806
ENSG00000105373	NOP53	-4.34447	0.036794
ENSG00000145907	G3BP1	-4.36424	0.048806
ENSG00000240350		-4.54109	0.048806
ENSG00000117115	PADI2	-4.60171	0.025721
ENSG00000237989	LINC01679	-4.96256	0.000206

Table 50: Significantly differentially expressed genes in the CD206^{+}CD163^{-} subset in the anti-PD1 patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

5.3.5.1.3 CD206^{-}CD163^{-} population

Moving on to the DN subset compared between RA and treated RA 62 genes were found to be upregulated and 20 down in the RA samples with 16 GO categories enriched for in the upregulated genes (Table 51, Table 52). As in the comparison of CD206^{+}CD163^{-} subset between RA and treated RA patients *IGHA1* was the top upregulated gene in the RA samples which could have several interpretations including (a) similarities between the CD206^{+}CD163^{-} subset and the DN subset, or (b) the possibility that the markers used for the sorting strategy

did not isolate pure populations. Using the results of the GO category enrichment to guide the interpretation of the DE genes revealed that several GO categories for NK cytotoxicity and cell cytotoxicity were enriched for in the genes upregulated in the RA samples compared to the treated RA. The presence of several solute carrier genes (e.g. *SLC25A47*) in the DE gene lists suggested that the treatment of RA could modulate metabolic pathways within this subset.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000211895	IGHA1	6.974043	0.001877
ENSG00000213262	VDAC2P3	6.553248	9.63E-06
ENSG00000280348		6.076781	0.001447
ENSG00000250295	RDH10-AS1	6.023984	0.026024
ENSG00000231858		5.982128	0.000378
ENSG00000167987	VPS37C	5.536398	0.000319
ENSG00000132464	ENAM	5.518767	0.004158
ENSG00000101084	C20orf24	5.439185	3.70E-07
ENSG00000241043	GVQW1	5.428825	0.04139
ENSG00000186020	ZNF529	5.416118	0.027787
ENSG00000279232		5.256478	0.026024
ENSG00000267046	E2F3P1	5.196253	7.43E-06
ENSG00000011590	ZBTB32	5.101131	0.04139
ENSG00000140107	SLC25A47	5.026169	0.04356
ENSG00000120875	DUSP4	5.005842	0.014427
ENSG00000109943	CRTAM	4.922803	0.000458
ENSG00000164402	SEPT8	4.898939	0.03063
ENSG00000257284		4.811731	0.000228
ENSG00000116685	KIAA2013	4.808939	0.043038
ENSG00000270956		4.776992	0.004412
ENSG00000265008		4.708816	0.037796
ENSG00000076924	XAB2	4.626947	0.029587
ENSG00000237798		4.62062	0.038523
ENSG00000228274		4.616442	0.043962
ENSG00000213445	SIPA1	4.610195	0.017539
ENSG00000253980		4.540726	0.017539
ENSG00000111276	CDKN1B	4.527756	0.043758
ENSG00000089692	LAG3	4.436409	0.000228
ENSG00000106952	TNFSF8	4.402801	0.036652
ENSG00000105639	JAK3	4.375774	0.018687
ENSG00000258869	LINC02312	4.339381	0.001447

ENSG00000174255	ZNF80	4.331375	0.036652
ENSG00000272765	NDUFA6	4.25523	0.020706
ENSG00000229757		4.176823	4.21E-05
ENSG00000184983	NDUFA6	4.148222	0.038523
ENSG00000235300		4.116512	0.034411
ENSG00000211959	IGHV4-39	4.10428	0.04139
ENSG00000277365	NDUFA6	4.057473	0.008208
ENSG00000273397	NDUFA6	4.057473	0.008208
ENSG00000179256	SMCO3	3.986879	0.014427
ENSG00000100453	GZMB	3.938109	0.033589
ENSG00000241511		3.916206	0.034998
ENSG00000267394		3.886087	0.037796
ENSG00000111371	SLC38A1	3.810471	0.04139
ENSG00000231621		3.791735	0.017591
ENSG00000276410	HIST1H2BB	3.778707	0.001445
ENSG00000213809	KLRK1	3.755903	0.043758
ENSG00000267289		3.726481	0.01997
ENSG00000256705		3.695339	0.012903
ENSG00000280014		3.639581	0.038523
ENSG00000250214		3.573027	0.0072
ENSG00000261188		3.532596	0.004158
ENSG00000249713		3.460803	0.045842
ENSG00000213049	HNRNPA1P34	3.318779	0.001357
ENSG00000253690		3.313296	0.007675
ENSG00000276173	IGHA2	3.118938	0.025499
ENSG00000138684	IL21	3.004527	0.003952
ENSG00000211890	IGHA2	2.963678	0.024684
ENSG00000267349		2.903838	0.026122
ENSG00000227777		2.87414	0.010244
ENSG00000266075	RN7SL574P	2.821571	0.008843
ENSG00000260810		2.724308	0.034998
ENSG00000224397	SMIM25	-3.74895	0.037796
ENSG00000079819	EPB41L2	-3.88235	0.038982
ENSG00000125355	TMEM255A	-3.93978	0.04139
ENSG00000230729		-3.97282	0.023515
ENSG00000171700	RGS19	-4.00507	0.043962
ENSG00000154237	LRRK1	-4.07276	0.026024
ENSG00000150782	IL18	-4.102	0.035053
ENSG00000259687	LINC01220	-4.10503	0.021092
ENSG00000115266	APC2	-4.33038	0.037796
ENSG00000184014	DENND5A	-4.35715	0.04791
ENSG00000169682	SPNS1	-4.41657	0.014427
ENSG00000256803		-4.44183	0.020771
ENSG00000197208	SLC22A4	-4.55919	0.023515
ENSG00000079691	CARMIL1	-4.63509	0.04139
ENSG00000182685	BRICD5	-4.66069	0.014427
ENSG00000101166	PRELID3B	-4.84251	0.014427
ENSG00000091106	NLRC4	-5.1076	0.012261

ENSG00000260918		-5.16476	0.004412
ENSG00000237166	LINC01792	-5.18256	0.000272
ENSG00000073756	PTGS2	-5.23397	0.004158

Table 51: Significantly differentially expressed genes in the double negative subset in the RA patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

GO category	Number of DE genes in category	Number of genes in category	GO term	Adjusted p value
GO:0002228	5	24	natural killer cell mediated immunity	0.0004472571
GO:0045954	4	11	positive regulation of natural killer cell mediated cytotoxicity	0.0006305874
GO:0002717	4	12	positive regulation of natural killer cell mediated immunity	0.0007599665
GO:0001909	5	32	leukocyte mediated cytotoxicity	0.0009850166
GO:0042269	4	15	regulation of natural killer cell mediated cytotoxicity	0.0012826535
GO:0002715	4	16	regulation of natural killer cell mediated immunity	0.0012826535
GO:0001906	5	36	cell killing	0.0012826535
GO:0001912	4	16	positive regulation of leukocyte mediated cytotoxicity	0.0014491686
GO:0031343	4	18	positive regulation of cell killing	0.0025256920
GO:0001910	4	21	regulation of leukocyte mediated cytotoxicity	0.0034940440
GO:0031341	4	23	regulation of cell killing	0.0053411435
GO:0002703	5	61	regulation of leukocyte mediated immunity	0.0089071018
GO:0002708	4	33	positive regulation of lymphocyte mediated immunity	0.0150655636
GO:0002705	4	36	positive regulation of leukocyte mediated immunity	0.0191917170
GO:0002449	5	87	lymphocyte mediated immunity	0.0398879453

Table 52: GO biological process categories enriched for in the upregulated genes in the double negative subset in RA patients versus treated RA patients. Benjamin-Hochberg adjusted p value cutoff <0.05.

In the comparison of DE genes between the DN subset in RA versus anti-PD1 arthritis eight genes were detected as upregulated and four as downregulated in the RA samples with no GO or KEGG enrichment when Ensembl IDs not from the primary genome assembly were excluded (Table 53). Of interest is that the genes *NDUFA6*, *IL21*, *RN7SL574P*, and *HIST1H2BB* were all upregulated in the RA samples compared to the anti-PD1 samples.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000071073	MGAT4A	4.736999	0.012503
ENSG00000272765	NDUFA6	4.206793	0.041975
ENSG00000277365	NDUFA6	4.05435	0.012503
ENSG00000273397	NDUFA6	4.05435	0.012503
ENSG00000276410	HIST1H2BB	3.81668	0.004999
ENSG00000150045	KLRF1	3.67763	0.012503
ENSG00000138684	IL21	2.892027	0.012503
ENSG00000266075	RN7SL574P	2.824856	0.015266
ENSG00000237011		-2.43368	0.049778
ENSG00000232759		-3.13133	0.045849
ENSG00000227191	TRGC2	-3.71764	0.049778
ENSG00000065534	MYLK	-5.14206	0.041975

Table 53: Significantly differentially expressed genes in the double negative subset in the RA patients compared to the anti-PD1 patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

In the final comparison of the macrophage subsets, the DN subset in anti-PD1 arthritis versus treated RA, 31 genes were detected as upregulated and 8 as downregulated in the anti-PD1 samples with no GO or KEGG enrichment (Table 54). Several genes known to be involved in T cell function and regulation were found in the upregulated genes such as *JAK3*, *LAG3*, *TRGC2*, and *TRGV8*, however genes such as *JAK3* and *LAG3* are also expressed by other cell types such as NK cells (Huard et al., 1994; Johnston et al., 1994; Kisielow et al., 2005; Leonard and O'shea, 1998; Triebel et al., 1990; Workman et al., 2009). With regards to the genes higher in the treated RA samples, *HAVCR2* could be of

interest due to its role as an immune checkpoint receptor expressed in T cells and cells of the monocyte/macrophage lineage (Anderson, 2014; Hou et al., 2016; Wang et al., 2017; Yi et al., 2017).

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000279232		6.946044	0.00024
ENSG00000076770	MBNL3	5.620846	0.032983
ENSG00000259224	SLC35G6	5.188346	0.049746
ENSG00000103260	METRN	5.02167	0.038415
ENSG00000267046	E2F3P1	5.017611	9.29E-05
ENSG00000111716	LDHB	4.977277	0.049746
ENSG00000140650	PMM2	4.924903	0.012571
ENSG00000255220	DDX18P5	4.922501	0.026005
ENSG00000271774		4.860723	0.048751
ENSG00000117868	ESYT2	4.82757	0.049746
ENSG00000235300		4.759529	0.011411
ENSG00000139626	ITGB7	4.716255	0.034727
ENSG00000236521	NPAP1P4	4.686774	0.000461
ENSG00000174255	ZNF80	4.581415	0.034727
ENSG00000147138	GPR174	4.563711	0.002007
ENSG00000121989	ACVR2A	4.538956	0.047257
ENSG00000223965	ZNF587P1	4.524749	0.030015
ENSG00000047579	DTNBP1	4.422963	0.023003
ENSG00000105639	JAK3	4.360885	0.037392
ENSG00000237914	SIRPG-AS1	4.314272	0.037392
ENSG00000177494	ZBED2	4.239018	0.012063
ENSG00000258268		4.210758	0.013209
ENSG00000255518		4.083287	0.000584
ENSG00000227191	TRGC2	3.993991	0.009495
ENSG00000258869	LINC02312	3.8239	0.024947
ENSG00000253690		3.647497	0.004206
ENSG00000267349		3.586344	0.004206
ENSG00000089692	LAG3	3.37021	0.037392
ENSG00000259078	PTBP1P	2.764817	0.041476
ENSG00000237011		2.43046	0.034727
ENSG00000211696	TRGV8	2.278121	0.049746
ENSG00000224397	SMIM25	-3.98236	0.013898
ENSG00000230729		-4.04974	0.02388
ENSG00000259687	LINC01220	-4.11006	0.022933
ENSG00000139146	FAM60A	-4.15515	0.049746

ENSG00000135077	HAVCR2	-4.32968	0.009495
ENSG00000118689	FOXO3	-4.44538	0.037392
ENSG00000118263	KLF7	-4.71123	0.049746
ENSG00000125355	TMEM255A	-4.72355	0.002007

Table 54: Significantly differentially expressed genes in the double negative subset in the anti-PD1 patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

5.3.5.2 Fibroblast subsets and clinical outcome groups

The differences between fibroblast subsets across clinical outcome groups were of a similar magnitude to those found in the macrophage. The number of differentially expressed genes in each comparison can be found in Table 55 with the comparison with the largest number of DE genes detected varying depending on the fibroblast subset under investigation. $CD34^+$ fibroblasts showed the most DE genes between RA and anti-PD1 samples, $CD34^-$ between anti-PD1 and treated RA samples (although the all comparisons had similar numbers of DE genes in this subset), and the $CD90^-GP38^+$ subset between the RA and treated RA samples. PCA did not find any clear separation of samples according to disease grouping in any of the principal components indicating that the variance imparted by the disease upon the fibroblast subsets is less distinct than the impact upon the macrophage subsets.

	CD34+	CD34-	CD90-GP38+
RA vs Treated RA	42	16	42
RA vs Anti-PD1	63	20	14
Anti-PD1 vs Treated RA	12	21	7

Table 55: Differentially expressed genes between the same fibroblast population between different disease stages

5.3.5.2.1 $CD34^+$ population

Beginning with the comparison of the $CD34^+$ fibroblast subset between RA and treated RA samples an interesting pattern could be seen in the upregulated genes (Table 56). Both *WARS*, encoding tryptophanyl-tRNA synthetase, and *CH25H*,

encoding cholesterol 25-hydroxylase, were upregulated in the RA samples and are interferon-inducible genes indicating cytokines in the interferon family may play an important role in modulating fibroblast gene expression in RA (Buwitt, Flohr and Bottger, 1992; Chen et al., 2014; Fleckner, Rasmussen and Justesen, 1991; Park and Scott, 2010). Of interest is the higher expression of *TGFB3-AS1* in the treated RA $CD34^+$ cells as antisense transcripts have been found to negatively regulate the translation of the complementary sense transcript of the gene from which they are transcribed through the pairing of the two transcripts (Mizuno, Chou and Inouye, 1984; Weiss, Davidkova and Zhou, 1999). No enrichment of GO categories or KEGG pathways was found in the eight up- or 34 down-regulated genes.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000140105	WARS	5.710392	0.010438
ENSG00000151929	BAG3	5.239385	0.018512
ENSG00000224468		4.808992	6.00E-06
ENSG00000227262	HCG4B	4.484109	0.007721
ENSG00000166226	CCT2	4.442182	0.03524
ENSG00000236708		4.309059	0.012061
ENSG00000229702	NA	3.890329	0.000473
ENSG00000138135	CH25H	3.041684	0.048838
ENSG00000239882	CYP21A1P	-3.18092	0.018456
ENSG00000231759	CYP21A1P	-3.25631	0.007721
ENSG00000227597	WASF1P1	-3.50633	0.021188
ENSG00000131116	ZNF428	-4.08139	0.018512
ENSG00000272140		-4.0906	0.040608
ENSG00000246627	CACNA1C-AS1	-4.33133	0.048227
ENSG00000267139		-4.34565	0.021188
ENSG00000227056	RPL6P2	-4.44317	0.003144
ENSG00000205056	LINC02397	-4.4467	0.01848
ENSG00000267226		-4.45933	0.03524
ENSG00000248138		-4.48368	0.012061
ENSG00000105583	WDR83OS	-4.51596	0.049096
ENSG00000235397	EPN2-AS1	-4.60567	0.018533
ENSG00000143473	KCNH1	-4.66878	0.007546
ENSG00000212443	SNORA53	-4.67212	0.03524
ENSG00000228170		-4.70031	0.043562

ENSG00000197747	S100A10	-4.73344	0.00552
ENSG00000103550	KNOP1	-4.80978	0.040608
ENSG00000214688	C10orf105	-4.85536	0.040608
ENSG00000028116	VRK2	-4.86701	0.007341
ENSG00000123240	OPTN	-4.90493	0.00646
ENSG00000235079	ZRANB2-AS1	-4.9142	0.049096
ENSG00000180658	OR2A4	-4.94809	0.000829
ENSG00000267439		-4.95693	0.006247
ENSG00000159200	RCAN1	-4.98667	0.00646
ENSG00000089234	BRAP	-5.32195	0.006794
ENSG00000243896	OR2A7	-5.32928	0.000338
ENSG00000258876	TGFB3-AS1	-5.41749	0.002008
ENSG00000255959		-5.45498	0.01848
ENSG00000224721		-5.49143	0.000219
ENSG00000166974	MAPRE2	-5.57152	0.000219
ENSG00000237070		-5.72333	0.00044
ENSG00000267776		-5.75136	0.002012
ENSG00000152217	SETBP1	-6.30935	0.000174

Table 56: Significantly differentially expressed genes in the CD34^{+} subset in the RA patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

Comparing the CD34^{+} subset between RA samples and anti-PD1 patients 15 genes were found to be upregulated in the RA samples and 48 in the anti-PD1 samples with no GO/KEGG enrichment in either group. Differential expression of genes related to the regulation of extracellular matrix and synovial fluid was apparent (Table 57). A key finding from the DE genes in this comparison that is of importance to the entire RNA-sequencing dataset and its analysis was the higher expression of *ZFY*; encoding zinc finger protein, y linked; in the RA samples and the higher expression of X-inactive specific transcript (*XIST*), encoding X inactive specific transcript, in the anti-PD1 samples. *ZFY* is located on the Y chromosome and hence is a male sex specific gene whereas *XIST* is located on the X chromosome and is critical for inactivation of one of the X chromosomes in females and hence can be considered a female sex specific gene (Herzing et al., 1997). When comparing the expression of *ZFY* and *XIST* in

both macrophage and fibroblast samples segregated according to sex this pattern could clearly be seen in this dataset (Figure 74). The detected differential expression of these genes in the $CD34^{+}$ RA versus anti-PD1 comparison indicated an imbalance in the sex of donors in each disease grouping contributed to the variance in the dataset and hence affected analysis of the data. Of the five anti-PD1 donors remaining after the sequencing dataset had been cleaned and filtered, four were from females whereas of the 11 RA donors three were female and eight were male meaning any comparison of any subset between these disease groups was confounded by sex specific differences (Table 58). The balance of sexes in the treated RA samples was better (three female and two male) but the comparison with other disease states will still have been impacted due to the relative imbalance in the other states.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000215277	RNF212B	6.09392	0.000362
ENSG00000223949	ROR1-AS1	5.243877	0.001214
ENSG00000162231	NXF1	5.077689	0.003668
ENSG00000224468		4.977171	1.13E-06
ENSG00000147457	CHMP7	4.722635	0.03345
ENSG00000236708		4.63318	0.004379
ENSG00000141644	MBD1	4.582734	0.025356
ENSG00000273445		4.424191	0.036043
ENSG00000272430		4.395145	0.01035
ENSG00000229702	NA	4.147064	0.000198
ENSG00000171502	COL24A1	4.109887	0.036043
ENSG00000149646	CNBD2	3.838423	0.036043
ENSG00000273656		3.556563	0.030381
ENSG00000067646	ZFY	3.34045	0.031522
ENSG00000244171	PBX2P1	3.08197	0.030381
ENSG00000006042	TMEM98	-2.8245	0.024013
ENSG00000257022		-3.78552	0.006615
ENSG00000231197		-3.97092	0.034925
ENSG00000226736		-4.00911	0.031587
ENSG00000271049		-4.23902	0.045844
ENSG00000112640	PPP2R5D	-4.28978	0.018287
ENSG00000266094	RASSF5	-4.31797	0.036043

ENSG00000267287		-4.42468	0.006615
ENSG00000101182	PSMA7	-4.53958	0.041507
ENSG00000107281	NPDC1	-4.56466	0.03345
ENSG00000139239		-4.57776	0.01035
ENSG00000131116	ZNF428	-4.58696	0.01035
ENSG00000261546		-4.66385	0.031587
ENSG00000237070		-4.69816	0.048001
ENSG00000246774	ZCCHC23	-4.74936	0.000198
ENSG00000106511	MEOX2	-4.75829	0.001885
ENSG00000184162	NR2C2AP	-4.77045	0.019001
ENSG00000214650		-4.77789	0.031326
ENSG00000268894	PLCE1-AS1	-4.78033	0.031326
ENSG00000205056	LINC02397	-4.79083	0.018287
ENSG00000255301		-4.79552	0.024013
ENSG00000183535	COL18A1-AS1	-4.81706	0.027628
ENSG00000106665	CLIP2	-4.84624	0.036043
ENSG00000267139		-4.84634	0.01233
ENSG00000231606		-4.85901	0.000878
ENSG00000068912	ERLEC1	-4.86135	0.027628
ENSG00000259627		-5.04022	0.023695
ENSG00000237990	CNTN4-AS1	-5.07734	0.006615
ENSG00000278899		-5.17621	0.031326
ENSG00000214688	C10orf105	-5.29493	0.018287
ENSG00000277277		-5.29545	0.00281
ENSG00000222018	C21orf140	-5.29545	0.00281
ENSG00000279649		-5.31598	0.023962
ENSG00000250902	SMAD1-AS1	-5.51104	0.00281
ENSG00000152217	SETBP1	-5.53329	0.006221
ENSG00000160716	CHRNB2	-5.54799	0.006887
ENSG00000115474	KCNJ13	-5.66035	0.018287
ENSG00000170961	HAS2	-5.72222	0.000198
ENSG00000260892		-5.75064	0.036043
ENSG00000078795	PKD2L2	-5.82622	0.001104
ENSG00000112851	ERBIN	-5.82641	0.00281
ENSG00000248138		-5.86026	0.000198
ENSG00000212443	SNORA53	-5.97818	0.001214
ENSG00000279989		-6.04146	0.00281
ENSG00000227844	NA	-6.05468	0.000198
ENSG00000279364		-6.19595	0.001085
ENSG00000267776		-6.29024	0.000915
ENSG00000229807	XIST	-7.78495	0.005775

Table 57: Significantly differentially expressed genes in the CD34^{+} subset in the RA patients compared to the anti-PD1 patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

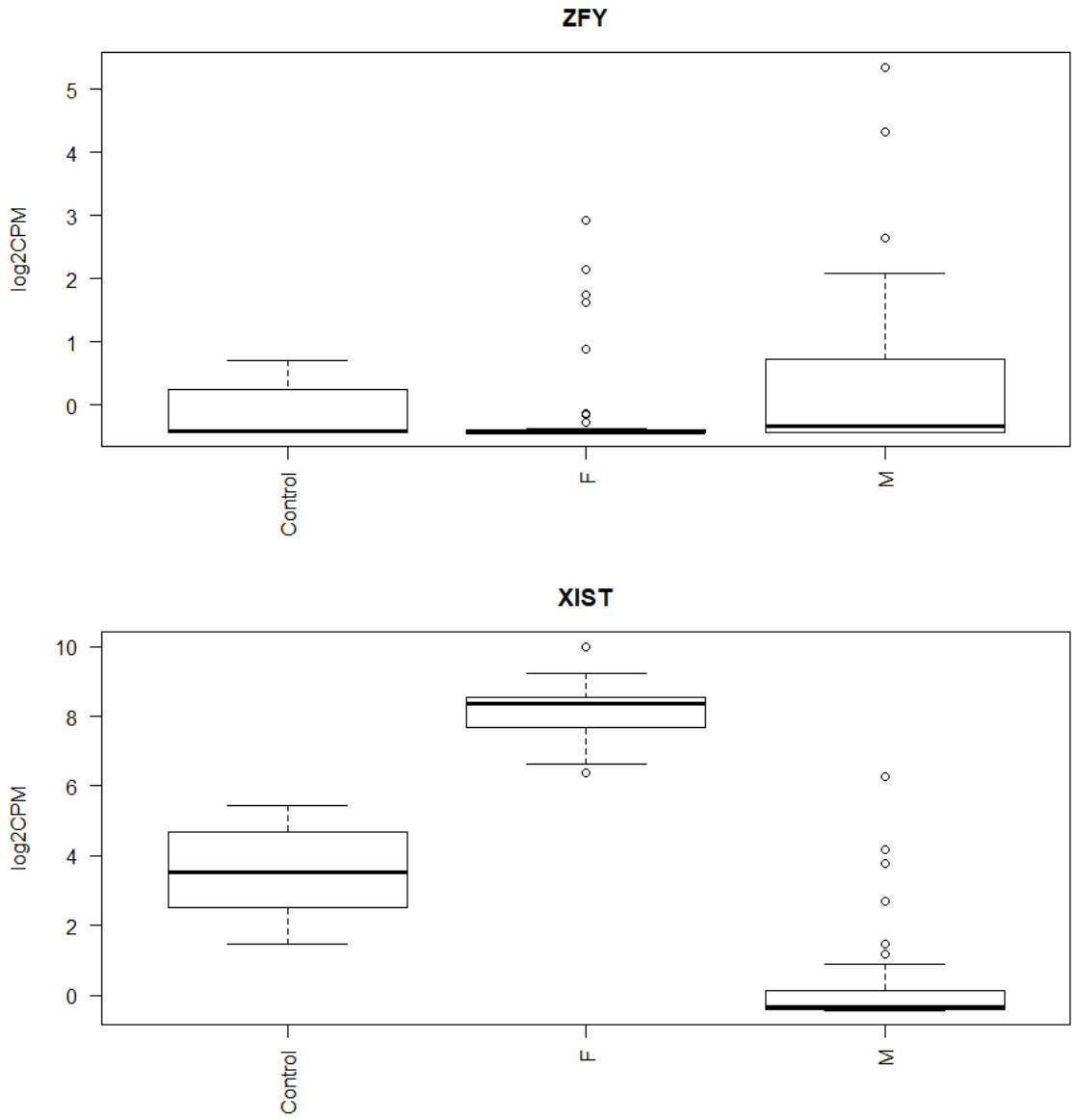


Figure 74: Log2 counts per million of the *ZFY* and *XIST* genes in the macrophage, fibroblast, and sequencing control samples grouped according to biological sex. Clear differences in expression depending on the sex of the patient were found in the expression of these sex specific genes as expected. The sequencing controls showed an intermediate level of expression possibly reflecting a mixed donor source used to generate the controls.

	Female	Male
RA	3	8
Treated RA	3	2
Anti-PD1	4	1

Table 58: Number of patients of each sex by disease group from which biopsies were collected

Of the four genes upregulated and the eight genes downregulated in the comparison of the CD34$^+$ subset between anti-PD1 and treated RA samples three genes were uncharacterised (Table 59). The remaining genes gave an unclear view of the differences between disease states.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000150995	ITPR1	5.315765	0.034378
ENSG00000246774	ZCCHC23	4.709045	0.001332
ENSG00000267287		4.640226	0.009334
ENSG00000257022		3.770372	0.028522
ENSG00000128951	DUT	-3.88263	0.028522
ENSG00000240163		-4.34741	0.046176
ENSG00000159445	THEM4	-4.39501	0.028522
ENSG00000227056	RPL6P2	-4.44205	0.009334
ENSG00000182197	EXT1	-4.79869	0.028522
ENSG00000154188	ANGPT1	-4.86355	0.009334
ENSG00000162231	NXF1	-4.89618	0.009334
ENSG00000100439	ABHD4	-4.93659	0.009334

Table 59: Significantly differentially expressed genes in the CD34$^+$ subset in the anti-PD1 patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

5.3.5.2.2 CD34$^-$ population

The comparison of the CD34$^-$ subset between RA and treated RA samples also found a smaller number of DE genes, four genes upregulated in RA samples and 12 upregulated in treated RA samples of which seven genes were uncharacterised (Table 60). As with many of the previous comparisons several pseudogenes were detected such as *TSPY26P* and *RPSAP44* and one antisense transcript, *MTUS2-AS1*, was also detected. No GO/KEGG enrichment was found in the DE genes and the remaining genes did not appear to have a clear disease-related role. In the comparison of the same fibroblast subset between RA and anti-PD1 samples 12 genes were detected as upregulated and eight as downregulated (Table 61). *XIST* was once again detected as expressed

at higher levels in the anti-PD1 patients reflecting the sex imbalance between the two disease groups. The final comparison of CD34$^-$ fibroblasts, anti-PD1 samples versus treated RA, found 10 genes upregulated and 11 downregulated with no GO/KEGG enrichment and a similar lack of apparent immunological genes (Table 62). The pseudogene *TSPY26P* was of interest as it was expressed at higher levels in the treated RA samples than in the anti-PD1 and also at higher levels than the RA samples indicating expression could have been elevated as a result of the treatment regime.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000275871		5.176225	0.029009
ENSG00000261222		5.134074	0.040076
ENSG00000071282	LMCD1	4.934517	0.033687
ENSG00000248367		4.016469	0.041791
ENSG00000235217	TSPY26P	-3.59175	0.024403
ENSG00000204186	ZDBF2	-4.70051	0.027551
ENSG00000179141	MTUS2-AS1	-4.75304	0.029009
ENSG00000169826	CSGALNACT2	-4.88309	0.044592
ENSG00000220773	RPSAP44	-5.18702	0.000924
ENSG00000163959	SLC51A	-5.26855	0.029009
ENSG00000280239		-5.36272	0.033687
ENSG00000179909	ZNF154	-5.60247	0.027551
ENSG00000267610		-5.64299	6.92E-05
ENSG00000248881		-5.69882	0.029009
ENSG00000250424		-5.73807	6.92E-05
ENSG00000166546	BEAN1	-6.19486	0.029009

Table 60: Significantly differentially expressed genes in the CD34$^-$ subset in the RA patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000259201		6.6854	0.01393
ENSG00000196611	MMP1	6.37126	0.01393
ENSG00000246422		6.1071	0.02152
ENSG00000119661	DNAL1	5.839298	0.01393
ENSG00000167291	TBC1D16	5.63874	0.01393
ENSG00000231340	ACTG1P10	5.593199	0.01393
ENSG00000213222		5.502951	0.01393
ENSG00000166848	TERF2IP	5.27065	0.01393

ENSG00000270412		5.235594	0.014671
ENSG00000261222		5.159689	0.02539
ENSG00000184863	RBM33	4.662555	0.01393
ENSG00000248367		4.014194	0.027866
ENSG00000125841	NRSN2	-4.58679	0.01393
ENSG00000272381		-4.61862	0.02539
ENSG00000179029	TMEM107	-4.90325	0.035133
ENSG00000196968	FUT11	-5.16166	0.016876
ENSG00000233728		-5.26956	0.01393
ENSG00000279291	NA	-5.9103	0.01393
ENSG00000232630	PRPS1P2	-5.91216	0.017643
ENSG00000229807	XIST	-7.64655	0.027866

Table 61: Significantly differentially expressed genes in the CD34$^-$ subset in the RA patients compared to the anti-PD1 patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000232630	PRPS1P2	5.838366	0.016858
ENSG00000101558	VAPA	5.628758	0.003034
ENSG00000235784	HNRNPA1P29	5.477071	0.002234
ENSG00000164708	PGAM2	5.409667	0.012965
ENSG00000233728		5.268614	0.011854
ENSG00000101109	STK4	5.255165	0.007151
ENSG00000129562	DAD1	4.599511	0.033735
ENSG00000205108	FAM205A	4.345098	0.012004
ENSG00000170265	ZNF282	4.26553	0.020164
ENSG00000279891		3.651538	0.044487
ENSG00000235217	TSPY26P	-3.59335	0.007151
ENSG00000231084		-3.61279	0.033735
ENSG00000020129	NCDN	-4.18743	0.035143
ENSG00000243230		-4.27197	0.007151
ENSG00000264714		-4.67001	0.016858
ENSG00000204186	ZDBF2	-4.99886	0.012004
ENSG00000267610		-5.05299	0.000681
ENSG00000263179	HNRNPCP4	-5.09903	0.033735
ENSG00000246731		-5.21062	0.035424
ENSG00000178988	MRFAP1L1	-6.01414	0.000241
ENSG00000254898		-6.399	0.000798

Table 62: Significantly differentially expressed genes in the CD34$^-$ subset in the anti-PD1 patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

5.3.5.2.3 CD90^{-}GP38^{+} population

Investigation of the final subset, CD90^{-}GP38^{+}, in RA compared to anti-PD1 found five upregulated and nine downregulated genes with no GO/KEGG enrichment. The gene *KHK*, encoding ketohexofructose/fructokinase, was expressed at higher levels in the anti-PD1 samples than the RA in this subset indicating a difference in the fructolysis pathway may have been elicited by either disease state. Differences in the metabolism of glucose could also be present between RA and the anti-PD1 arthritis samples as *RRAD* has been shown to inhibit glucose uptake and was expressed at higher levels in the RA samples (Bustamante et al., 2017; Garcia-Carbonell et al., 2016). The increase in *RRAD* expression was also found in the comparison of the CD90^{-}GP38^{+} subset between RA and treated RA patients with higher levels in the RA samples again (Table 64). In this comparison 11 genes were found to be upregulated in the RA samples and 31 upregulated in the treated RA samples. An important finding of this comparison was the increased expression of *TNFRSF10B*, encoding TRAILR2/Death Receptor 5 (DR5), in the treated RA samples. A final DE gene of interest found at higher levels the treated RA samples was *NDUT6*. *NDUT6* is transcribed from the antisense strand of the *FGF2* gene and therefore may actually be involved in the downregulation of FGF2 protein expression which has been demonstrated as detrimental in arthritides (Manabe et al., 1999; Nakano et al., 2004; Yamashita et al., 2002).

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000107551	RASSF4	6.38104	0.000558
ENSG00000279465		5.648922	0.04831
ENSG00000177409	SAMD9L	5.462681	0.04831
ENSG00000140391	TSPAN3	5.335383	0.04831
ENSG00000166592	RRAD	5.177511	0.03631

ENSG00000138030	KHK	-4.24968	0.04831
ENSG00000227660	UST-AS1	-4.31555	0.03631
ENSG00000204186	ZDBF2	-4.32309	0.04831
ENSG00000076053	RBM7	-4.47211	0.03631
ENSG00000247728		-4.8814	0.04831
ENSG00000173113	TRMT112	-4.98429	0.031455
ENSG00000254693		-5.21715	0.031455
ENSG00000235347		-5.24634	0.04831
ENSG00000273747		-6.07678	0.009519

Table 63: Significantly differentially expressed genes in the CD90^{-}GP38^{+} subset in the RA patients compared to the anti-PD1 patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000019582	CD74	5.809231	0.019917
ENSG00000186283	TOR3A	5.704267	0.029681
ENSG00000279465		5.608922	0.024836
ENSG00000131746	TNS4	5.534926	0.024836
ENSG00000109323	MANBA	5.515379	0.019917
ENSG00000115282	TTC31	5.335068	0.036245
ENSG00000256928		5.179047	0.036245
ENSG00000166592	RRAD	5.112516	0.018782
ENSG00000104325	DECR1	5.050201	0.011237
ENSG00000171067	C11orf24	4.616714	0.035995
ENSG00000137992	DBT	4.276096	0.034103
ENSG00000176490	DIRAS1	-3.01715	0.024633
ENSG00000237207	RBM17P3	-3.50673	0.011237
ENSG00000212747	RTL8B	-3.58364	0.017821
ENSG00000203778	FAM229B	-3.74214	0.024836
ENSG00000274547		-3.84599	0.000624
ENSG00000036672	USP2	-3.8655	0.032039
ENSG00000262558		-3.91177	0.024836
ENSG00000166780	C16orf45	-4.01378	0.030371
ENSG00000114698	PLSCR4	-4.12779	0.047509
ENSG00000104067	TJP1	-4.13242	0.034103
ENSG00000226430	USP17L7	-4.15655	0.024836
ENSG00000273611	ZNHIT3	-4.16221	0.024836
ENSG00000274415		-4.20084	0.015406
ENSG00000245060	LINC00847	-4.25868	0.017821
ENSG00000250424		-4.26814	0.029196
ENSG00000278574	ZNHIT3	-4.30119	0.019917
ENSG00000123243	ITIH5	-4.33992	0.024836
ENSG00000106341	PPP1R17	-4.40834	0.024836
ENSG00000277401	TJP1	-4.44113	0.041702
ENSG00000173511	VEGFB	-4.4705	0.024836
ENSG00000237070		-4.54554	0.039511
ENSG00000103024	NME3	-4.6088	0.011237

ENSG00000236287	ZBED5	-4.70481	0.028914
ENSG00000143570	SLC39A1	-4.73114	0.011237
ENSG00000120889	TNFRSF10B	-4.87225	0.009013
ENSG00000248664		-5.15086	0.044061
ENSG00000151743	AMN1	-5.20813	0.004862
ENSG00000097033	SH3GLB1	-5.32562	0.005828
ENSG00000091317	CMTM6	-5.35686	0.004335
ENSG00000254693		-5.65329	0.00254
ENSG00000170917	NUDT6	-6.09343	0.000361

Table 64: Significantly differentially expressed genes in the CD90$^-$GP38$^+$ subset in the RA patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

The final comparison of fibroblast subsets was of the CD90$^-$GP38$^+$ subset in anti-PD1 arthritis patients compared to treated RA patients. In this comparison two genes were upregulated, both of which were uncharacterised, in the anti-PD1 samples and five genes were upregulated in the treated RA samples (Table 65). Of the five genes higher in the treated RA samples four were characterised, however, with the exception of the aforementioned *NDUT6* (*FGF2* antisense) the interpretation of these genes in the context of this work was unclear.

Ensembl ID	HGNC symbol	Log2 fold change	Adjusted P value
ENSG00000273747		6.077605	0.004314
ENSG00000258012		5.292796	0.027492
ENSG00000176490	DIRAS1	-3.02651	0.029015
ENSG00000267986		-3.80091	0.029015
ENSG00000189320	FAM180A	-3.93847	0.007624
ENSG00000170917	NUDT6	-5.25461	0.024516
ENSG00000135299	ANKRD6	-5.42767	0.020564

Table 65: Significantly differentially expressed genes in the CD90$^-$GP38$^+$ subset in the anti-PD1 patients compared to the treated RA patients sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

5.4 Discussion

In the work in this chapter samples from untreated RA patients, RA patients with inadequate responses to methotrexate, and cancer patients treated with anti-PD1 agents who had subsequently developed arthritis were collected and cryopreserved before digesting the samples, cell sorting fibroblast and macrophage populations, and isolating the RNA from the populations. It is important to remember that all of these samples were taken from patients with actively inflamed joints, and a significant degree of global disease activity; there is therefore no “uninflamed” control, and it is possible that subtle differences between disease groups, for instance, could be obscured by genes involved in the inflammatory response. To some extent this is unavoidable as biopsy of minimally inflamed joints is a developing technique. The proportions of stromal populations, in particular the $CD34^-$ subset, correlated with ESR and synovial hypertrophy. RNA-sequencing of fibroblast populations found few clear differences in function between subsets or disease group, however it did discover differences of possible interest between disease groups. Sequencing of macrophage populations identified a $CD206^+CD163^{+/-}$ subset and a $CD206^-CD163^-$ mixed lymphocyte population.

Investigating the genes expressed at higher levels in the $CD206^+CD163^+$ subset than the DN subset confirmed the suspected macrophage identity of the subset with complement genes such as *C1QB*, *C2*, and *C1QA*, scavenger receptor genes such as *COLEC12* and *MARCO*, and other genes such as *SIGLEC1*, *CD14*, *MERTK*, and *TGFB1* all significantly higher in this subset (Baeten et al., 2000; Baeten et al., 2004; Canton, Neculai and Grinstein, 2013; De Ceulaer,

Papazoglou and Whaley, 1980; Hartnell et al., 2001; Hartung and Hadding, 1983). The finding of significant expression of *MERTK* is of interest in the context of work that suggested expression of *MERTK* can be used to discriminate macrophages from dendritic cells and identify tissue-resident macrophage populations (Gautier et al., 2012; Xue et al., 2014a).

Investigation of the genes upregulated in the DN subset compared to the double positive indicated the cells are heterogeneous and displayed little evidence of a macrophage phenotype. The genes alluded towards the presence of two main cell populations within the $CD45^{+}HLA\text{-}DR^{+}CD206^{-}CD163^{-}$ group with the expression of *NGK7*, *CD8A*, *CD3G*, *CD3E*, and *CD3D* all hinting at a cytotoxic T lymphocyte identity (Alarcon et al., 1988; Hidalgo et al., 2008; Turman et al., 1993). The other cell type present in the population seemed to be B lymphocytes identified by the expression of the immunoglobulin genes *IGHG3* and *IGHM* and the gene *JCHAIN* which encodes a protein involved in IgM secretion (Johansen, Braathen and Brandtzaeg, 2000). Two interesting points can be made from the expression of these genes, (1) the expression of *IGHG3* suggested that IgG3 is being produced by these cells and production of this isotype is known to be stimulated by IFNγ suggesting this cytokine is involved in RA, and (2) the expression of *IGHM* and *JCHAIN* suggested that IgM is being secreted which is of importance given the dominance of this isotype in the pool of rheumatoid factor antibodies present in RA patients (Gioud-Paquet et al., 1987; Snapper et al., 1992).

The expression of a gene well known in the cancer field, *CAV1* encoding the

caveolin-1 protein, was higher in the $CD90^{-}GP38^{+}$ fibroblast subset than in the $CD34^{+}$ cells as was the expression of *CTSK*, encoding cathepsin K. Caveolin-1 is used as a biomarker in the study of many cancers as downregulation of the protein in tumour stroma is associated with a poor prognosis (Simpkins et al., 2012; Sloan et al., 2009; Sotgia et al., 2011; Witkiewicz et al., 2009). Cathepsin K could clearly be involved in the pathogenesis of RA given the degradative nature of the enzyme and is elevated in the serum of RA patients, is correlated with radiological progression, and genetic knockout of the gene in a murine arthritis model improves inflammation, bone erosion, and cell infiltration (Hao et al., 2015; Skoumal et al., 2004). When comparing the $CD34^{+}$ subset between RA and anti-PD1 samples it was apparent that the regulation of extracellular matrix and synovial fluid constituents was different between the disease states either as a result of the anti-PD1 treatment or due to differences in the pathogenesis of the ailments. *COL24A1*, *COL18A1-AS1*, and *HAS2* were all differentially expressed in this comparison with the first gene higher in the RA samples and the second and third higher in the anti-PD1 samples.

Investigating subset differences it appeared that the $CD90^{-}GP38^{+}$ subset may have a more degradative role than the other two subsets, particularly the $CD34^{+}$ subset. Recent work identifying stromal subsets in RA and OA patients provides somewhat similar results to the findings here. Mizoguchi et al. (2017) found that $CD34^{-}CD90^{-}$ synovial fibroblasts had higher MMP1 production in response to TNFα stimulation than the other subsets, a finding similar with to the work above except the $CD34^{-}$ population in this thesis was defined as $CD90^{+}$. The study also found that the $CD34^{+}$ stromal cells expressed higher levels of IL-6 and CXCL12

when stimulated with TNFα and that the $CD34^+CD90^+$ subset, which would match the population in this thesis, had higher levels of RANKL expression than the other subsets.

An important issue with the RNA-sequencing dataset was highlighted by the detection of differential expression of the sex-linked genes *ZFY* and *XIST* between the $CD34^+$ subset in the RA and anti-PD1 samples. The ratio of female to male samples in the anti-PD1 samples was skewed towards female with only one of the five anti-PD1 patients being male. This influenced the findings of comparisons with these samples as the differences being detected could actually have been due to sex rather than due to the disease states. Although several computational methods exist to correct for unwanted variation in large datasets the viability of using a method on data with such a severe bias and relatively low replicates is questionable and may affect the validity of the results in other ways (Johnson, Li and Rabinovic, 2007; Leek, 2014; Leek and Storey, 2007; 2008). Therefore it is important to keep in mind that some of the differentially expressed genes detected in the analysis are probably due to sex differences.

Although the findings in this chapter show some interesting differences in gene expression both between fibroblast or macrophage subsets and within subsets between disease groups improvements could be made to greatly increase the robustness and possibly the number of significant findings in the RNA-sequencing dataset. To improve the utility of the dataset more samples would need to be sequenced both to increase the number of samples in each disease group but also aid in the balancing or even computational removal of sex

differences to reduce the impact of this on variation in the data. An larger number of samples were not sequenced to begin with for two reasons; (1) time constraints and the need to develop a digestion protocol prior to sample collection limited the number of samples that could be collected, and (2) RNA-sequencing is costly and although Novartis completed the sequencing resources were still transferred to facilitate this.

As sequencing control samples were included in the sequencing then, so long as the sequencing controls were included in any subsequent sequencing batches, batch correction would be possible to allow merging of the experiments. The inclusion of more samples would improve the power of statistical analyses, a necessary improvement as the number of samples in the final analysis was low. A larger sample number in each group would allow the use of algorithms such as comBat or surrogate-variable analysis to correct for unwanted sources of variation (Johnson, Li and Rabinovic, 2007). Another improvement to the dataset would be to add a ‘normal’, uninflamed control to the dataset so comparisons can be made to a non-diseased sample group.

For many of the comparisons small numbers of DE genes were detected with many genes being non-protein coding or uncharacterised. This casts doubt upon the robustness of these findings, particularly given the low number of samples in each group. If the differential expression of these genes was successfully validated then the biological significance would still be unclear as it is not possible to assign gross functional differences on the basis of such a small number of genes. In addition changes in the expression of genes does not always correlate

with changes in protein expression due to numerous post-transcriptional and post-translational regulatory mechanisms.

For the identification of cellular subsets, in particular the macrophage subsets, it could be argued that a targeted approach using pre-defined markers is a naïve method to identify the full spectrum of subsets. Given enough funds and time, bar-coded single-cell sequencing using platforms such as 10x could have been a better approach to address this due to the unbiased nature of the technique and would not only allow identification of macrophage subsets but also investigation of the genes expressed in each subset as attempted in this work. One issue consistently encountered during the establishment of a suitable panel of cell surface markers for sorting was the limited number of channels available. This was confounded by the small quantity of synovial biopsy tissue collected which prevented the use of the material in two separate sorts with more expansive stromal or macrophage panels. Using single cell sequencing one could sort broad populations such as $CD45^{-}CD235a^{-}CD326^{-}$ stromal cells and $CD45^{+}$ hematopoietic cells and use unsupervised cluster analysis to identify the sub-populations.

Finally, whilst the investigation of synovial cellular population data using the flow cytometry uncovered interesting correlations and differences it was all tested using proportions of cells rather than cell counts. This can make the interpretation of the data complex as changes in the proportion of one cell type will obviously affect the proportions of others whereas in reality it could be changes in the number of cells of one or more populations eliciting this effect. To account for this

effect further data should be collected whilst also noting the weight of tissue prior to digestion and also using counting beads to obtain accurate cell counts from flow cytometry data. This would allow the calculation of normalised cell counts which in turn would allow clearer interpretation of the results.

In summary synovial fibroblast and macrophage populations were identified by immunofluorescent staining of tissue sections and synovial cell suspensions. A protocol was established to allow uniform cryopreservation and digestion of synovial tissue across sites which facilitates comparisons between results. The protocol allowed the isolation of both fibroblast and macrophage populations. Fibroblast populations showed dynamic changes and appeared to expand with inflammation and immune cell infiltration into the synovium. Macrophages and a mixed lymphocyte population changed in proportion in different disease stages with RA having a higher proportion of the lymphocyte population than anti-PD1 induced arthritis patients and the opposite relationship displayed for $CD206^{+}CD163^{+}$ macrophages. The isolation of these populations from synovial biopsy material provided enough material for bulk RNA-sequencing. The findings of the RNA-sequencing dataset revealed that the CD206+CD163+ macrophage population had a scavenging phenotype, the $CD206^{-}CD163^{-}$ population in RA appeared to be enriched for cytotoxic T cells and IgM secreting B cells, and synovial fibroblasts displayed differences in phenotype both between subsets and between disease stages however the magnitude and relevance of the differences in both cases was not large.

The work in this chapter only partially addressed the third aim in section 1.9 that

different fibroblast and macrophage subsets exist within the synovium with distinct functions. The presence of different fibroblast subsets was demonstrated by staining of synovial tissue sections and cell suspensions and the subsequent RNA-sequencing data, however no clear functions could be assigned to the subsets using this dataset. Although 3 putative macrophage populations were identified in the synovial cell suspensions only the $CD206^{+}CD163^{-}$ and $CD206^{+}CD163^{+}$ populations appeared to be macrophages and very few differences were detected between these populations meaning they may actually represent a very similar, if not identical, population which fails to prove that different macrophage populations exist in the synovium.

The finding of *MERTK* expression in the $CD206^{+}CD163^{+}$ macrophage population, along with the expression of macrophage markers *CD14* and *SIGLEC1*, adds further evidence to the postulated tissue-resident macrophage identity of these cells as suggested by the flow cytometry data in chapter 4. In addition, the expression of genes associated with scavenger functions (*MARCO, COLEC12*), and tissue remodelling (*TGFB1*) cast the $CD206^{+}CD163^{+}$ population as an anti-inflammatory macrophage as hinted by the expression of CD206 and CD163. It is of particular interest that the $CD206^{+}CD163^{+}$ population constitutes a significantly higher proportion of the $CD45^{+}$HLA-DR^{+} population in anti-PD1 induced arthritis than in RA. Unfortunately it is not yet clear if anti-PD1 arthritis is an acute ailment or if it can develop into a chronic arthritis. If the former case is found true it is enticing to speculate that the higher proportion of tissue-resident macrophages may facilitate the resolution of inflammation whereas the number of tissue-resident macrophages is reduced in RA and as such results in

inadequate resolution.

The finding that the stromal $CD90^{+}GP38^{+}CD34^{-}$ population significantly correlates with inflammation and synovial hyperplasia in the RA samples is particular incriminating. Using the immunofluorescence data to assign locations to these subsets leads to the $CD90^{+}GP38^{+}CD34^{+/-}$ subsets appearing to be sublining subsets whereas the $CD90^{-}GP38^{+}$ subset represents lining layer fibroblasts. The expression of the gene encoding the adhesion molecule *CDH23* at higher levels in the $CD90^{-}GP38^{+}$ subset provides further evidence that this population represents the lining layer, a region of the synovium which is tightly associated and dependent upon cadherin expression for this function (Kiener et al., 2009; Lee et al., 2007). It could be postulated that the sublining $CD34^{-}$ cells drive inflammation in the RA synovium given the correlations, however the gene expression data did not detect clear pro-inflammatory cytokines as being enriched in this subset. Both the sublining $CD34^{-}$ population and the lining layer population express *MMP1* meaning both populations are capable of degrading collagen, a key pathological finding in RA. The $CD34^{+}$ population shows signs of regulating cell infiltration into the synovium as these fibroblasts express higher levels of *ACKR3* which is known to bind to CXCL12 and CXCL11 but it is not clear if this induces signalling or not (Burns et al., 2006). This may sequester available CXCL12 and CXCL11 inhibiting cell infiltration into the joint or alternatively generate a chemokine gradient to facilitate increased infiltration into the joint. Additionally the $CD34^{+}$ subset expresses *CTSK*, an alternative matrix remodelling enzyme to MMP1 but one in which overexpression in murine models has been shown to result in joint damage and synovitis again demonstrating the

destructive capacity of the fibroblast subpopulations (Morko et al., 2005). Whilst the functions of fibroblast subsets appear somewhat overlapping there is evidence that the method by which these functions are achieved are differs in each subset. This would be relevant to therapeutics targeting fibroblast populations as targeting *MMP1* expressing subsets for example may not successfully prevent joint destruction due to expression of *CTSK* by the $CD34^{+}$ subset.

6 CONCLUDING REMARKS

Three hypotheses were postulated and tested during this body of work. The first hypothesis, that fibroblasts isolated from different stages of RA progression differentially modulate monocyte or macrophage phenotype *in vitro*, was unable to be proven and as such the null hypothesis was accepted. Although the cultured fibroblast lines were unable to differentially modulate the phenotype of monocytes and macrophages according to the stage of RA from which the lines were isolated a coordinated modulation of cytokine production was elicited by co-culture of monocytes and fibroblasts regardless of disease stage (Figure 75). The modulation appeared to be a rapid event occurring within 24 hours of the initiation of co-culture, was cell contact in-dependent, and resulted in the regulation of cytokines and chemokines. Fibroblasts also appear to be able to influence to differentiation of monocytes into macrophages. This section of work confirmed published findings regarding fibroblast and monocyte interactions and contributes to the field by discovering that fibroblast disease stage does not affect the interactions.

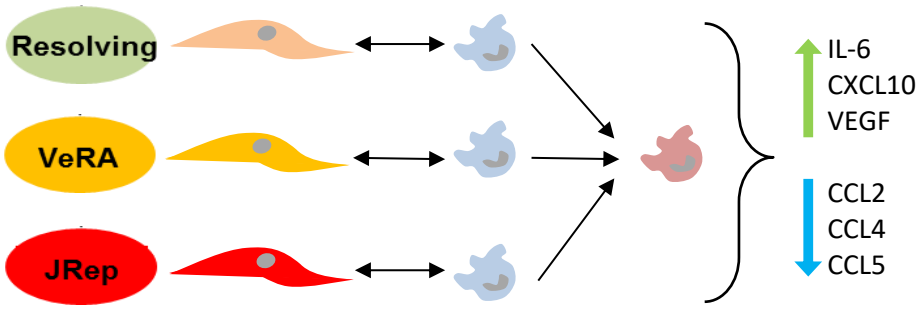


Figure 75: Co-culture of synovial fibroblasts and monocytes/macrophages does not result in differential modulation dependent upon the stage of RA from which the fibroblasts were isolated. However, co-culture does result in the modulation of the release of IL-6, CXCL10, VEGF, CCL2, CCL4, and CCL5 amongst others from the co-cultured cells.

The second hypothesis that different fibroblast and macrophage subsets exist within the synovium with distinct functions was partially proven although insufficient evidence was gathered to complete the testing of fibroblast and macrophage functions or to demonstrate the existence or lack of various synovial macrophage subsets. Using immunofluorescence staining both fibroblast and macrophage populations were identified in sections of synovial tissue with data generated by this work being included in and leading to publications (Choi et al., 2017; Croft et al., 2016; Misharin et al., 2014). In order to further analyse these populations a digestion protocol for the isolation of cells from synovial biopsy material was established in collaboration with members of the NIH AMP consortium, and this collaborative work is now being developed into two manuscripts for publication, one in review following resubmission after comments with partial data in a pre-publication repository (Mizoguchi et al., 2017). The protocol was then used to identify the stromal and macrophage populations in the synovium and to sort cells and isolate RNA for RNA-sequencing. The flow cytometry data generated during the cell sorting process identified significant correlations of RA fibroblast populations with inflammatory markers and differences in the proportion of leukocyte cells between RA and other arthritic diseases. A significant portion of these data form part of the current submission under review (Mizoguchi et al., 2017). The RNA-sequencing process was challenging given the low number of cells isolated for each population but yielded data of adequate quality for analysis. Subset analysis found differentially expressed genes between fibroblast subsets but provided no clear picture of functional differences. In the macrophage populations two populations were very similar with regards to gene expression but one sorted cell population appeared

to be a mixed leukocyte population. Differences in subsets between disease groups were present but did not provide a clear idea of functional differences. This section of work significantly contributed to the discovery and examination of fibroblast subsets in the synovium both through the development of a unified digestion protocol and through initial studies of subset differences in RA and other diseases (Figure 76). Contributions were also made towards the study of synovial macrophage populations with the RNA-sequencing of a definitive macrophage population.

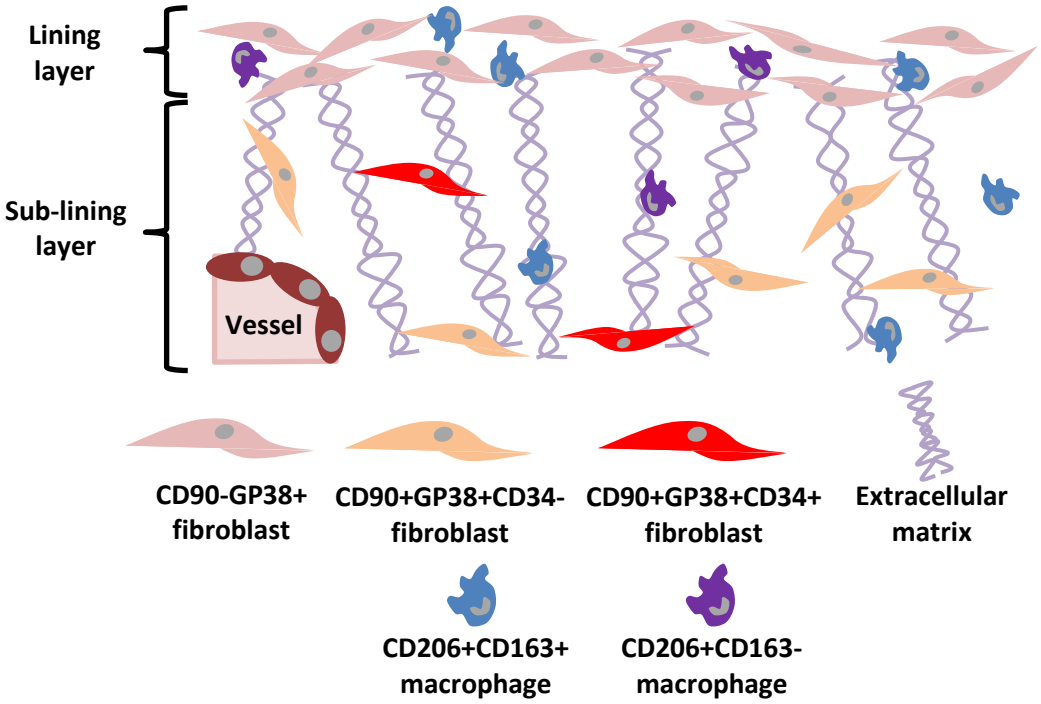


Figure 76: Synovial fibroblasts can be separated into at least three subpopulations and synovial macrophages appear to exist as two subpopulations. Synovial cell subpopulations can be identified by flow cytometry and show significant differences in gene expression as measured by RNA-sequencing.

Considerable evidence was found in support of the third hypothesis that fibroblasts isolated from different stages of RA progression maintain differential gene expression during *in vitro* co-culture and respond to stimuli in different ways.

Analysis of new microarray data, building on data generated by a previous student using fibroblast lines from different clinical outcome groups including different stages of RA progression with different stimuli and culture conditions found a large number of differentially expressed genes between disease stages in each treatment group. Differences in stromal subset markers were identified that validate the findings of flow cytometry and tissue staining data and highlighted a potential marker of very early RA that could be used alongside other novel biomarkers to stratify early disease.

This body of work has considerably expanded the knowledge regarding synovial fibroblasts and macrophages in RA and has opened new avenues to search for therapeutics. It was known that fibroblasts from late stage RA possess an 'imprinted aggressor' phenotype characterised by increased pro-inflammatory cytokine production and increased invasion and degredation of cartilage, however the point during RA progression at which the imprinting occurs is not known (Bottini and Firestein, 2013; Whitaker et al., 2013). The microarray data herein found that whilst fibroblasts cultured from the synovium of patients with resolving arthritis and very early RA both show differential gene expression when compared with fibroblasts from uninflamed joints there were no differentially expressed genes between the two groups, but differentially expressed genes were detected between the resolving arthritis and established RA or joint replacement RA groups. This indicates that the imprinted phenotype published in the literature may not occur during the intial stages of RA when symptoms first appear. This in turn may explain the therapeutic window seen whereby patients treated within the first three months of symptom onset have an improved disease

course with less joint destruction; it could be reasoned that the fibroblasts are more amenable to treatment during this early period before the imprinted phenotype is truly established. It would be pertinent to compare the differentially expressed genes between very early RA samples and the established RA and joint replacement RA samples and also between the resolving arthritis samples and the two later RA groups to assess if there are any genes common to the two lists that may explain the imprinted phenotype.

The data obtained on fibroblast-monocyte/macrophage co-cultures found no statistically significant difference between co-cultures with fibroblasts from different stages of RA with regards to secreted factors or gene expression. Regardless co-cultures with either monocytes or differentiating macrophages influenced protein and gene expression in both cases demonstrating the capability of fibroblasts to regulate the immune response. Of particular interest was the ability of fibroblasts to downregulate *TNF* and *IL10* expression in differentiating macrophages insinuating that synovial fibroblasts in the absence of exogenous factors elicit an anti-inflammatory phenotype from monocyte derived macrophages and can skew monocyte differentiation. This is somewhat suprising given the literature surrounding the pro-inflammatory phenotype of RA synovial fibroblasts. It may be that imprinted fibroblasts do not just drive aggressive disease mechanisms but still retain some features of maintaining homeostasis in the joint and the balance between the two has shifted in RA fibroblasts and is overwhelmed in inflamed joints. It was not investigated within this body of work but the interactions between fibroblasts and monocytes or macrophages in the presence of exogenous proinflammatory factors is of

significant interest, given the elements of regulation of inflammation displayed in the co-cultures it may be that in the presence of TNFα, for example, this vestige of regulation is lost and instead fibroblasts elicit an overtly inflammatory phenotype from monocytes and macrophages. Although no significant differences were found slight trends towards reduced inflammatory responses were seen in co-cultures with resolving fibroblasts. This should be further investigated as if proved true it would increase the resolution with which we can identify the point at which synovial fibroblasts become imprinted in RA which in turn could inform treatment regimes.

The work in chapters 4 and 5 demonstrated the existence of separate fibroblast subsets within the synovium beyond the distinction between lining and sublining layer. Of particular interest was the finding that the $CD34^-$ sublining population was significantly positively correlated with both systemic inflammation but also synovial hyperplasia, indicating a large portion of the expansion of the synovium during RA may be due to this population. This is in contrast to previous findings indicating that the lining layer is expanded and responsible for joint destruction in RA (Kiener et al., 2006; Kiener et al., 2009; Lee et al., 2007). The evidence demonstrates a discord between joint damage and inflammation with one state not necessarily accompanying the other. This can be explained by the hypothesis that the different fibroblast subpopulations drive different aspects of RA, for example the $CD34^-$ subset may drive inflammation whereas the $CD90^-GP38^+$ subset is responsible for destruction of the joint architecture. It would be interesting to return to the literature identifying different subtypes of RA, such as fibroid, lymphoid, or myeloid, with fibroblast subsets in mind and quantify the size

and activity of each subset in each subtype (Dennis et al., 2014). This could lead to a stratified medicine approach in RA with the subtype identified dictating which fibroblast subpopulation will be targeted.

In addition to the fibroblast subsets identified putative tissue-resident macrophages were identified in the synovium on the basis of CD206, CD163, and MerTK expression. It is not yet known if tissue-resident macrophages exist in the synovium however MerTK has been identified as a marker of tissue-resident macrophages and was expressed by the $CD206^{+}CD163^{+}$ macrophages identified herein (Gautier et al., 2012). Identifying tissue-resident macrophages in humans is difficult and as such the work is normally completed by fate-mapping in animal models, however one approach to confirm the residence and self-replenishing nature of the $CD206^{+}CD163^{+}$ population within the synovium could be to recruit patients who have undergone bone marrow transplants and through repeat biopsies and genetic typing measure the contribution of donor monocytes to the synovial macrophage pool. The genes expressed by the $CD206^{+}CD163^{+}$ population were indicative of an anti-inflammatory phenotype meaning this population could be responsible for maintaining homeostasis within the joint. Given the higher proportion of this population in the anti-PD1 induced arthritis patients than RA this may be one mechanism by which inflammation becomes chronic in RA, through a reduction in the number of tissue-resident macrophages in the synovium. In terms of therapeutics, targeting only monocytes and supporting tissue-resident macrophages may promote remission in RA patients as opposed to therapies targeting both cell types together.

In summary, this thesis has expanded knowledge of cellular subsets within the synovium in RA and on the phenotype of synovial fibroblasts during RA progression. Several avenues of investigation show promise as possible sources of new targets for therapeutics in RA and also may aid in the development of stratified treatments depending upon the subtype of RA with which patients present.

7 FUTURE WORK

Although the work in this thesis has addressed most of the hypotheses postulated further work should be completed both to improve the integrity of the data and findings so far and to validate and expand upon the initial findings. Future work for the co-cultures includes:

- Assess the impact of TNFα and IL-10, due to the clear pro- and anti-inflammatory effects of these cytokines, upon fibroblast:monocyte co-cultures.
- Investigate fibroblast:monocyte co-cultures after 24 hours to test for disease specific differences in the temporality of cytokine/chemokine release.
- Identify the sources of cytokines released during co-culture either through the using of secretion blocking agents such as Brefeldin A or through mRNA analysis.
- Co-culture freshly sorted synovial fibroblast subsets or macrophages and assess IL-6 and MMP1 production.

Future work for the synovial digestion and RNA-sequencing work:

- The protocol optimisation work requires an increase in sample size particularly in the fresh versus frozen digestion comparisons.
- Rather than using population proportions in the flow cytometry data for the protocol optimisation and the sample digestions for RNA-sequencing counting beads will be used to allow measurement of absolute cell count and normalising to tissue weight to be more accurate.
- The sample size in all groups will be increased by sequencing more samples and correcting for batch effects using the sequencing controls.

This will increase the power of the statistical tests to detect differentially expressed genes and increase confidence in the current results.

- Adding the BEACON cohort disease outcome groups to the RNA-sequencing dataset will allow comparison of fibroblast and macrophage subsets across disease stages, a comparison not possible in the microarray dataset.
- The results of the RNA-sequencing analysis requires validation by qRT-PCR or by functional assays to determine if clear functional readouts can be assigned to the genes. To this end basal and stimulated MMP1 production from fibroblast subsets will be measured and the metabolic state of fibroblast subsets from outcome groups will be measured using a Seahorse XF analyser.
- Fibroblasts and macrophage subsets must be sorted and *in vitro* cultures established. This is to allow validation of the RNA-sequencing findings regarding gene expression and functional phenotype and to address the question of longevity of subsets i.e. do the subsets maintain gene and marker expression over several passages in culture or do they progress towards a single phenotype. The longevity of fibroblast subsets will be assessed using RNA-sequencing and flow cytometry analysis of the subsets at each passage.

Future work for the microarray analysis:

- As with the RNA-sequencing experiments the results of the microarray must be validated through qRT-PCR of gene expression of through functional assays. The expression of FOS family members in Jrep fibroblasts, histone H1 family members in the normal samples, and subset

markers CD34 and FAP in all outcome groups will be measured.

- Present synovial fibroblast cultures generated from the BEACON cohort will be tested for the presence of subset markers using flow cytometry to address the questions raised by the microarray data regarding the homo-/heterogeneity of cultured synovial fibroblast lines.

Two first author publications will be generated using the work in this thesis:

- The microarray data will be used to address the hypothesis that fibroblasts from different clinical outcome groups show transcriptional and functional differences. Briefly the DE genes detected between outcome groups that are associated with the matrisome will be validated using a combination of qRT-PCR, ELISA, and matrix degradation assays. The presence or absence of subsets of fibroblasts in the cultured fibroblast lines will be confirmed by flow cytometric analysis as will the persistence of subsets in *in vitro* culture using freshly sorted synovial fibroblasts. The pathogenic behaviour of fibroblast lines will be confirmed using the SCID mouse cartilage implantation model as in (Muller-Ladner et al., 1996). Investigation of the mechanism behind the maintained differential expression of genes will be investigated by assaying chromatin accessibility ATAC-seq or through assaying gene methylation status using bisulphite sequencing.
- The current RNA-sequencing data and the associated flow cytometry data will be used to characterise the changes in synovial fibroblast subsets occurring during arthritis subsequent to anti-PD1 treatment and to investigate the effect of treatment upon cells in fibroblast subsets RA synovium. Briefly, the number of samples sequenced will

be increased and counting beads used during cell sorts to make the data more robust. The key differences in gene expression between subsets will be confirmed in RA samples and validated using qRT-PCR, ELISA, matrix degradation assays, chemotaxis assays using conditioned media, and co-cultures with leukocytes. Differences in subset gene expression and function across disease groups will then be investigated and again confirmed using the above assays. Histological differences between the three disease groups will be investigated using hematoxylin and eosin and also immunofluorescent staining of synovial tissue sections.

The overarching questions to address going forwards in this field are:

1. How do synovial fibroblast subsets contribute to the pathology of RA?
2. How do synovial fibroblast subsets differ in function and number during the progression of RA?
3. Do fibroblasts from resolving arthritis demonstrate different interactions with monocytes and macrophages than those from RA?
4. Does the $CD206^{+}CD163^{+}$ macrophage population represent a tissue-resident population within the synovium and does this subset dampen inflammation?

8 APPENDIX

8.1 Authored publications arising during PhD studies

Choi, I.Y., Karpus, O.N., Turner, J.D., Hardie, D., Marshall, J.L., De Hair, M.J.H., Maijer, K.I., Tak, P.P., Raza, K., Hamann, J., Buckley, C.D., Gerlag, D.M. and Filer, A. (2017) 'Stromal cell markers are differentially expressed in the synovial tissue of patients with early arthritis'. PLoS ONE, 12: (8): e0182751.

Mizoguchi, F., Slowikowski, K., Marshall, J.L., Wei, K., Rao, D.A., Chang, S.K., Nguyen, H.N., Noss, E.H., Turner, J.D., Earp, B.E., Blazar, P.E., Wright, J., Simmons, B.P., Donlin, L.T., Kalliolias, G.D., Goodman, S.M., Bykerk, V.P., Ivashkiv, L.B., Lederer, J.A., Hacohen, N., Nigrovic, P.A., Filer, A., Buckley, C.D., Raychaudhuri, S. and Brenner, M.B. (2017) 'Single Cell Transcriptomics And Flow Cytometry Reveal Disease-Associated Fibroblast Subsets In Rheumatoid Arthritis'. bioRxiv.
**This is a pre-print deposited in a repository. At the time of submission a more extensive manuscript is under review.*

Ross, E.A., Naylor, A.J., Neil, J.D., Crowley, T., Ridley, M.L., Crowe, J., Smallie, T., Tang, T.J., Turner, J.D., Norling, L.V., Dominguez, S., Perlman, H., Verrills, N.M., Kollias, G., Vitek, M.P., Filer, A., Buckley, C.D., Dean, J.L. and Clark, A.R. (2017) 'Treatment of inflammatory arthritis via targeting of tristetraprolin, a master regulator of pro-inflammatory gene expression'. Annals of the Rheumatic Diseases, 76: (3): 612.
**I contributed to this work by staining synovial tissue markers for stromal and macrophage markers and tristetraproline and imaging the tissues*

Croft, A.P., Naylor, A.J., Marshall, J.L., Hardie, D.L., Zimmermann, B., Turner, J., Desanti, G., Adams, H., Yemm, A.I. and Müller-Ladner, U. (2016) 'Rheumatoid synovial fibroblasts differentiate into distinct subsets in the presence of cytokines and cartilage'. Arthritis Research & Therapy, 18: (1): 270.

Turner, J.D. and Filer, A. (2015) 'The role of the synovial fibroblast in rheumatoid arthritis pathogenesis.' Current Opinion in Rheumatology, 27: (2): 175-182.

Misharin, Alexander v., Cuda, Carla m., Saber, R., Turner, Jason d., Gierut, Angelica k., Haines, G.K., Iii, Berdnikovs, S., Filer, A., Clark, Andrew r., Buckley, Christopher d., Mutlu, Gökhan m., Budinger, G.R.S. and Perlman, H. (2014) 'Nonclassical Ly6C− Monocytes Drive the Development of Inflammatory Arthritis in Mice'. Cell Reports, 9: (2): 591-604.

Abstracts

Montero-Melendez, T., Buckley, C., Filer, A., Turner, J. and Perretti, M. (2016) 'AB0077 Targeting Synovial Fibroblasts with Melanocortin Drugs'. Annals of the Rheumatic Diseases, 75: (Suppl 2): 923-923.

Turner, JD., Clark, A., Buckley, CD., Filer, A. (2016). 'Synovial fibroblast and monocyte interaction leads to differential cytokine regulation in direct co-culture.' Keystone Stromal Cells in Immunity Conference 2016

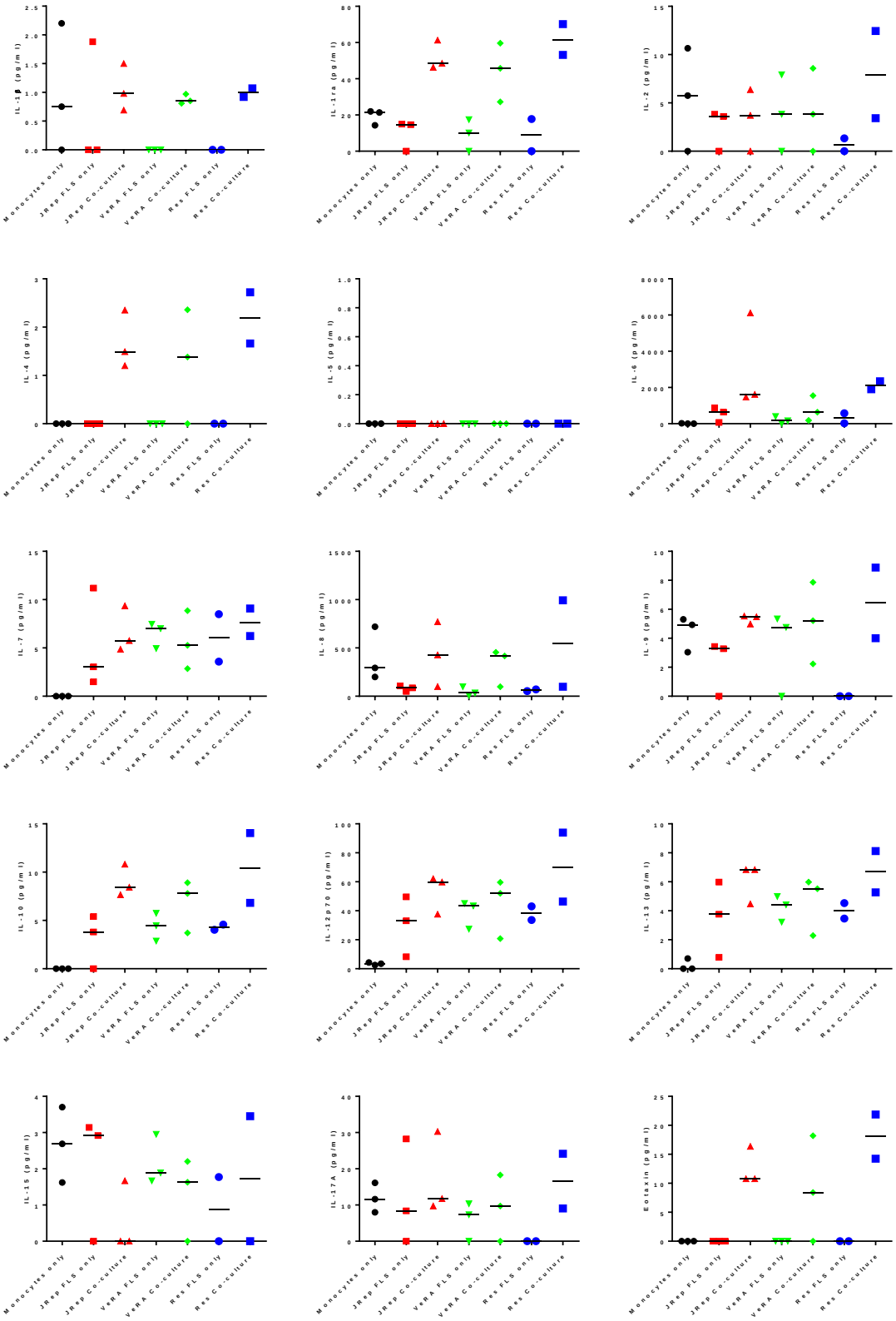
Desanti, G., Naylor, A., Navarro-Nunez, L., Turner, J., Saghir, A., Lowe, K., Filer, A., Watson, S. and Buckley, C. (2015) 'THU0035 A Key Role for Platelet-Derived Clec-2 in the Regulation of Synovial Inflammation'. Annals of the Rheumatic Diseases, 74: (Suppl 2): 205-205.

Ross, E., Smallie, T., Naylor, A., Desanti, G., Crowe, J., O'neil, J., Turner, J., Yemm, A., Norling, L. and Perlman, H. (2015) 'A8. 1 Tristetraprolin is a novel therapeutic target for rheumatoid arthritis'. Annals of the Rheumatic Diseases, 74: (Suppl 1): A81-A81.

Heron, K., Turner, J.D., Hardie, D., Adams, H., Raza, K., Buckley, C.D., Tennant, D. and Filer, A. (2014) 'O20. Hypoxic Regulation of Epigenetic Modifications During Disease Progression in Rheumatoid Arthritis'. Rheumatology, 53: (suppl 1): i36-i36.

Filer, A., Hardie, DL., Choi, YK., Karpus, O., Turner, JD., Hamann, J., Raza, K., Buckley, CD., and Gerlag, DM. (2013) 'Analysis of immunohistological images of stromal cell populations in ultrasound guided biopsies of synovium to help predict patient outcomes in rheumatoid arthritis.' CYTO Conference 2013

8.2 *In vitro* monocyte:fibroblast luminex data



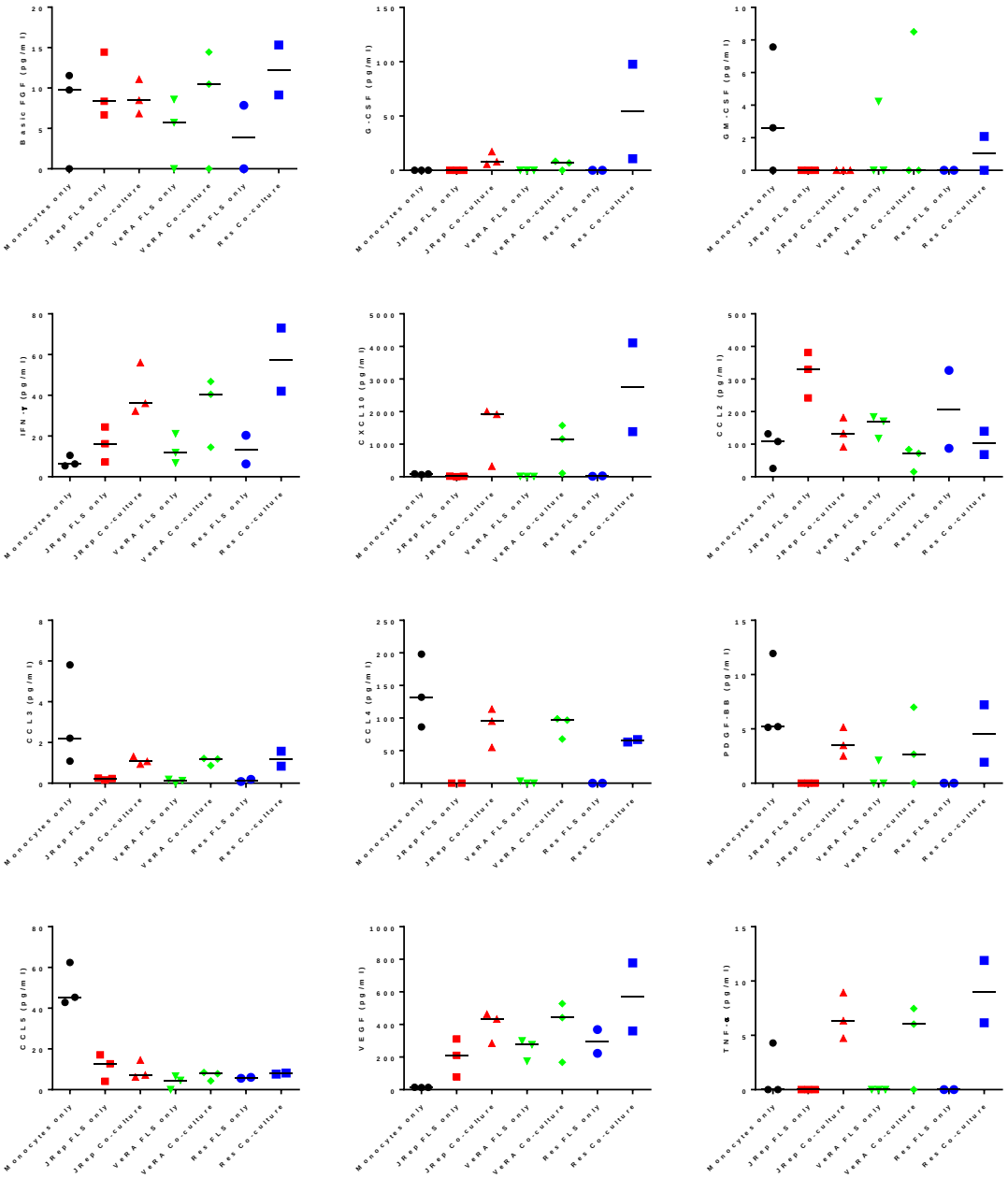


Figure 77: Results of all factors analysed in the multiplex analysis of supernatants harvested from unstimulated monocyte:fibroblast co-cultures. Bars show the median.

8.3 Synovial tissue digestion protocol

1. Turn on the heating plate for the digestion. Warm serum-free RPMI-1640 containing 0.97mM sodium orthopyruvate, 1.94mM glutamine, 97U/ml penicillin, 97µg/ml streptomycin, and [0.97x] MEM non-essential amino acids to 37^{o}C in preparation for use in the digestion buffer.

2. Drain media/saline from tissue using 100µm cell filter.

3. Place tissue in Petri dish and cut into pieces 1mm^3 using scalpels. Weigh tissue

4. Divide tissue between 5ml FACS tubes at a max of 0.02g/tube. Add magnetic flea to tube.

5. Make up digestion buffer.
    i. 10ml RPMI-1640 (serum free with additives as in step 1)
    ii. 200ul Liberase TM (50µg/ml final concentration)
    iii. 40ul DNase1 (40µg/ml final concentration)

6. Add 1ml digestion buffer to each tube and digest at 37°C with a magnetic stirrer for 30 minutes. Stirrer revolutions set at 380rpm

7. After 15 minutes of digestion agitate the tissue via repeated aspiration using a 14g needle. Continue digestion afterwards.

8. Thoroughly resuspend/agitate digested suspensions and pass through 70µm cell strainer. Rinse filter with 10ml RPMI 10% FBS (with NEAA, SOP, GPS).

9. Spin down at 1400rpm for 6 minutes and continue with staining protocol

<u>Freezing tissue</u>

1. Pre-cool Mr Frosty to 4^{o}C
2. Collect tissues by biopsy needle. (Size of the tissue is about 16G diameter x 1-10 mm length)
3. Mince tissue with scalpels until pieces are approximately 1mm diameter
4. Weigh tissue and place into 1ml cryovial with 1ml of Cryostor 10. Agitate vial to separate tissue pieces out.
5. Put the tube on ice for 10 min.
6. Put the tube into pre-chilled Mr Frosty and place into -80^{o}C Freezer

<u>Thawing tissue</u>

1. Place cryovial into 37^{o}C water bath to rapidly thaw vial
2. As soon as sample has thawed decant tissue into a 70µm filter and rinse with 10ml serum-free RPMI-1640
3. Transfer tissue into vessel for digestion and continue with usual digestion protocol

8.4 R script for the analysis of RNA-sequencing data from *ex vivo* synovial populations

```
#Required packages
library("BiocInstaller")
library(Biobase)
library(limma)
library(edgeR)
library(Glimma)
library(gplots)
library(org.Hs.eg.db)
library(RColorBrewer)
library("DESeq2")
library("genefilter")
library("ggplot2")
library("ggfortify")

# Reading in the data
seqdata <- read.delim("NG00009.0-Ensembl-human-76-bowtie2-pe-gene-2.cnt", stringsAsFactors = FALSE)
sampleinfo <- read.delim("SampleInfoCleaned2.txt", comment.char = "", stringsAsFactors = TRUE, fill=TRUE)
sampleinfo$Group <- paste(sampleinfo$Diagnosis, sampleinfo$Population, sep = ".")
head(seqdata)
dim(seqdata)
countdata <- seqdata[,-(1)]
head(countdata)
rownames(countdata) <- seqdata[,1]
colnames(countdata) <- sampleinfo$Sample.Name
table(colnames(countdata)==sampleinfo$Sample.Name)


#removal of samples identified as outliers through several iterations of the normalisation and data preparatio
n
sampleinfo <- sampleinfo[c(-2,-5,-6,-12,-14,-15,-16,-17,-18,-19,-20,-21,-25,-31,-72,-73,-74,-75,-76,-77,-79,-8
6,-99,-100,-108),]
countdata <- countdata[,c(-2,-5,-6,-12,-14,-15,-16,-17,-18,-19,-20,-21,-25,-31,-72,-73,-74,-75,-76,-77,-79,-86
,-99,-100,-108)]
rownames(sampleinfo) <- sampleinfo$Sample.Name

#Remove genes not in any samples
countdata2 <- countdata[rowSums(countdata)>0,]

#Collection additional annotation information and adding the current annotation information
library(biomaRt)
ensembl = useMart(biomart = "ENSEMBL_MART_ENSEMBL",dataset="hsapiens_gene_ensembl", host="www.ensembl.org")
Attributes <- listAttributes(ensembl)
Filters <- listFilters(ensembl)
BMimport =getBM(attributes= c("ensembl_gene_id", "hgnc_symbol", "entrezgene", "external_gene_name",
                              "start_position", "end_position"), values= rownames(countdata),
                filter="ensembl_gene_id", mart= ensembl)
BMimport$gene_length <- (BMimport$end_position - BMimport$start_position) +1
library("plyr")
df <- data.frame(row.names = rownames(countdata2), stringsAsFactors=FALSE)
df$ensembl_gene_id <- rownames(df)
df2 <- join(df, BMimport, by="ensembl_gene_id", type="left", match="first")


#Assessing log2 CPM to find where the boundry with noise lies
logcounts <- cpm(countdata2,log=TRUE)
plot(logcounts[,1],logcounts[,2])

# Create expression set, function provided by Sabina Pfister
mr.analysis.get.eset <- function(x.data, x.annots, x.genes)
{
  require(Biobase)
  pData <- new("AnnotatedDataFrame",x.annots)
  fData <- new("AnnotatedDataFrame", x.genes)
  eset <- new("ExpressionSet", exprs = as.matrix(x.data), phenoData = pData, featureData= fData)
  return(eset)
}

esetdata <- mr.analysis.get.eset(countdata2, sampleinfo, df2)


# Filter read if intensity is at least 1 in 75% of the samples in at least one group
# Yann Abraham code from Sabina Pfister
mr.analysis.filter <- function(eset, groups, p=0.75, val=1, na.rm=TRUE)
{
  require(genefilter)
  require(edgeR)

  cat(nrow(eset),'total identifiers\n')

  ffun <- filterfun(pOverA(p=p,val,na.rm=na.rm))
```

```
  fltr <- lapply( levels(groups), function(name)
  {
    sub <- cpm(eset[,groups==name],log=FALSE)
    ft <- genefilter(sub, ffun)
    return( ft )
  })
  fltr <- do.call('cbind',fltr)
  fltr <- rowSums(fltr)

  eset <- eset[fltr!=0,]
  cat(nrow(eset),'identifiers passed filtering\n')
  return(eset)
}

esetdataFilt <- mr.analysis.filter(esetdata,
                                   groups = factor(sampleinfo$Group),
                                   val= 1)

#Voom transformation and quantile normalisation
y <- DGEList(exprs(esetdataFilt), genes = fData(esetdataFilt))
barplot(y$samples$lib.size, ylab = "Library size", main = "Sample library sizes",names = colnames(y), las=2, c
ex.names = 0.75)
y <- calcNormFactors(y)
y$samples
group <- factor(sampleinfo$Group)
design <- model.matrix(~ 0 + group)
colnames(design) <- levels(group)
v <- voom(y,design,plot = TRUE, normalize.method = "quantile")

par(mar=c(7,5,3,3))
boxplot(v$E, xlab="", ylab="Log2 counts per million",las=2,main="Voom transformed logCPM", cex.axis = 0.75)
boxplot(cpm(y, log = TRUE), xlab="", ylab="Log2 counts per million",las=2,main="non-transformed logCPM",cex.ax
is = 0.75)

#THIS CODE RUN AFTER INITIAL DE GENE ANALYSIS USING PREPARED DATA
#PURPOSE IS MAKE DATA AMENABLE TO GSEA AND GO ENRICHMENT
#USE THIS STEP TO PREPARE DATA FOR SECOND DEG ANALYSIS AND SUBSEQUENT GSEA AND GO ANALYSIS
#Remove non-primary assembly genes and deprecated genes (using logic of higher ensembl ID are non-assembly for
#duplicated genes and NA in external gene name are deprecated)
dim(v)
v <- v[!duplicated(v$genes$external_gene_name),]
dim(v)
v <- v[!is.na(v$genes$external_gene_name),]
#END OF GSEA DATA PREP


#Multidimensional scaling plots
plotMDS(v)
par(mfrow=c(3,2))
#Coloured according to variables
#Diagnosis
levels(sampleinfo$Diagnosis)
col.dia <- c("purple","orange", "red", "blue")[sampleinfo$Diagnosis]
data.frame(sampleinfo$Diagnosis,col.dia)
plotMDS(v,col=col.dia, pch=16)
legend("bottomleft",fill=c("purple","orange", "red", "blue"),legend=levels(sampleinfo$Diagnosis))
title("Diagnosis")
#Cell type
levels(sampleinfo$CellType)
col.cell <- c("purple","orange", "red")[sampleinfo$CellType]
data.frame(sampleinfo$CellType,col.cell)
plotMDS(v,col=col.cell, pch=16)
legend("bottomleft",fill=c("purple","orange", "red"),legend=levels(sampleinfo$CellType))
title("Cell type")
#Population
levels(sampleinfo$Population)
col.pop <- c("purple","orange", "red", "blue", "green", "black", "brown", "grey")[sampleinfo$Population]
data.frame(sampleinfo$Population,col.pop)
plotMDS(v,col=col.pop, pch=16)
legend("bottomleft",fill=c("purple","orange","red", "blue", "green", "black", "brown", "grey"),legend=levels(s
ampleinfo$Population))
title("Cell population")
#Sex
levels(sampleinfo$Sex)
col.sex <- c("purple","orange", "red")[sampleinfo$Sex]
data.frame(sampleinfo$Sex,col.sex)
plotMDS(v,col=col.sex, pch=16)
legend("bottomleft",fill=c("purple","orange", "red"),legend=levels(sampleinfo$Sex))
title("Sex")
#Site
levels(sampleinfo$Site)
col.site <- c("purple","orange", "red", "blue")[sampleinfo$Site]
data.frame(sampleinfo$Site,col.site)
plotMDS(v,col=col.site, pch=16)
legend("bottomleft",fill=c("purple","orange", "red", "blue"),legend=levels(sampleinfo$Site))
```

```
title("Site")
#Digest batch
levels(sampleinfo$DigestBatch)
col.digbat <- c("purple","orange", "red", "blue")[sampleinfo$DigestBatch]
data.frame(sampleinfo$DigestBatch,col.digbat)
plotMDS(v,col=col.digbat, pch=16)
legend("bottomleft",fill=c("purple","orange", "red", "blue"),legend=levels(sampleinfo$DigestBatch))
title("Digest batch")


#Checking batch vs diagnosis and diagnosis vs cell type
par(mfrow=c(1,2))
char.batch <- c(1,2,3,4)[sampleinfo$DigestBatch]
plotMDS(v,col=col.dia, pch=char.batch,cex=2)
legend("topleft",fill=c("purple","orange", "red", "blue"),legend=levels(sampleinfo$Diagnosis))
legend("bottomright",legend=levels(sampleinfo$DigestBatch),pch=c(1,2,3,4))

char.cell <- c(1,2,3)[sampleinfo$CellType]
plotMDS(y,col=col.dia, pch=char.cell, cex=2)
legend("topleft",fill=c("purple","orange", "red", "blue"),legend=levels(sampleinfo$Diagnosis))
legend("bottomright",legend=levels(sampleinfo$CellType),pch=c(1,2,3))

#Interactive MDS plot with Glimma
labels <- paste(sampleinfo$SampleName, sampleinfo$Diagnosis, sampleinfo$CellType, sampleinfo$Population, sampl
einfo$Sex, sampleinfo$Site)
Group <- paste(sampleinfo$Diagnosis, sampleinfo$CellType, sampleinfo$Population, sampleinfo$Sex, sampleinfo$Si
te,sep=".")
Group <- factor(Group)
glMDSPlot(v, labels=sampleinfo$Sample.Name, groups=factor(sampleinfo$Diagnosis), folder="mdsVoomFirstdemul")

#PCA
pca <- prcomp(t(v$E), scale=TRUE, center = TRUE)
autoplot(pca, data = sampleinfo, colour = 'Population', size = 2,
         xlab="PC1: 17.2% of variance", ylab = "PC2: 4.3% of variance",
         loadings = TRUE)

#PCAs per subset
vpca <- v
#For macrophages
vpca.mphage <- vpca[, grep("Macrophage", sampleinfo$CellType)]
col.dia.mphage <- col.dia[grep("Macrophage", sampleinfo$CellType)]

mp.pca <- prcomp(t(vpca.mphage$E), scale=TRUE, center = TRUE)
autoplot(mp.pca, x= 2, y= 3, data = sampleinfo[grep("Macrophage", sampleinfo$CellType),], colour = 'Population
', size = 3,
         frame = TRUE, frame.type = 'norm', xlab= "PC2: 7.23% of variance", ylab ="PC3: 3.97% of variance")
#3D PCA
library(ggbiplot)
library(pca3d)
pca3d(pca, components = 4:6, col = col.dia.mphage, axes.color = "grey", bg = "white", radius = 1, labels.col =
"black",
      show.axes = TRUE, group=factor(sampleinfo$Diagnosis[grep("Macrophage", sampleinfo$CellType)]),
      show.axe.titles = TRUE,  show.plane = TRUE,
      show.shapes = TRUE, show.centroids = FALSE, show.ellipses = FALSE, legend= "bottomleft")

#Fibroblasts
vpca.fibro <- vpca[, grep("Fibroblast", sampleinfo$CellType)]
col.dia.fibro <- col.dia[grep("Fibroblast", sampleinfo$CellType)]
col.pop.fibro <- col.pop[grep("Fibroblast", sampleinfo$CellType)]

fb.pca <- prcomp(t(vpca.fibro$E), scale=TRUE, center = TRUE)

autoplot(fb.pca, x= 2, y = 3, data = sampleinfo[grep("Fibroblast", sampleinfo$CellType),], colour = 'Populatio
n', size = 3,
         frame = TRUE, frame.type = 'norm', xlab = "PC2: 4.24% of variance", ylab="PC3: 3.76% of variance")
#3D PCA
library(pca3d)
pca3d(fb.pca, components = 1:3, col = col.dia.fibro, axes.color = "grey", bg = "white", radius = 1, labels.col
= "black",
      show.axes = TRUE, group=factor(sampleinfo$Diagnosis[grep("Fibroblast", sampleinfo$CellType)]),
      legend = "bottomleft")
pca2d(fb.pca, components = 2:3, col = col.pop.fibro, axes.color = "grey", bg = "white", radius = 1, labels.col
= "black",
      show.axes = TRUE, group=factor(sampleinfo$Population[grep("Fibroblast", sampleinfo$CellType)]),
      legend = "bottomleft", biplot = TRUE)


#tSNE
library("Rtsne")
Res.tSNE <- Rtsne(t(v$E), dims = 5, initial_dims = 15, perplexity = 20,
      theta = 0.5, check_duplicates = TRUE, pca = TRUE, max_iter = 1000,
      verbose = TRUE)
plot(Res.tSNE$Y[,1], Res.tSNE$Y[,2], xlab="tSNE dim 1", ylab="tSNE dim 2", col=col.dia, pch=16)
#CD34 colour gradient
ggplot(Res.tSNE$Y, aes(Res.tSNE$Y[,1], Res.tSNE$Y[,2])) +
```

```
  geom_point(aes(colour = v$E["ENSG00000174059",])) +
  scale_colour_gradient2(low = muted("black"), mid = "grey", high ="purple") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "black"))


#Hierarchical clustering (Pearson's correlation distance of top 400 most variable genes across samples)
#Macrophages
var_genes <- apply(v[,grep("Macrophage", sampleinfo$CellType)], 1, var)
head(var_genes)
select_var <- names(sort(var_genes, decreasing=TRUE))[1:400]
head(select_var)
highly_variable_vmphage <- v$E[select_var, grep("Macrophage", sampleinfo$CellType)]
dim(highly_variable_vmphage)
head(highly_variable_vmphage)
#Pearson correlation distance, labels changed if required
hc <- hclust(as.dist(1-cor(highly_variable_vmphage)))
plot(hc, labels = group[grep("Macrophage", sampleinfo$CellType)])

#Fibroblasts
var_genes <- apply(v[,grep("Fibroblast", sampleinfo$CellType)], 1, var)
head(var_genes)
select_var <- names(sort(var_genes, decreasing=TRUE))[1:400]
head(select_var)
highly_variable_vfibro <- v$E[select_var, grep("Fibroblast", sampleinfo$CellType)]
dim(highly_variable_vfibro)
head(highly_variable_vfibro)
#Pearson correlation distance, labels changed if required
hc <- hclust(as.dist(1-cor(highly_variable_vfibro)))
plot(hc, labels = group[grep("Fibroblast", sampleinfo$CellType)])

#Sanity checks on sorted markers
par(mar=c(10,5,3,3))
sampleinfo$Population <- factor(sampleinfo$Population, c("DN", "206p163n", "206p163p", "34n", "34p", "90n38p",
"Control"))
boxplot(v$E["ENSG00000081237",]~factor(sampleinfo$Population),vertical=TRUE,las=2,
          ylab="log2CPM",main="PTPRC", col = c("red", "red", "red", "blue", "blue", "blue", "green4"))
boxplot(v$E["ENSG00000261371",]~factor(sampleinfo$Population),vertical=TRUE,las=2,
          ylab="log2CPM",main="PECAM1", col = c("red", "red", "red", "blue", "blue", "blue", "green4"))
boxplot(v$E["ENSG00000154096",]~factor(sampleinfo$Population),vertical=TRUE,las=2,
          ylab="log2CPM",main="THY1", col = c("red", "red", "red", "blue", "blue", "blue", "green4"))
boxplot(v$E["ENSG00000162493",]~factor(sampleinfo$Population),vertical=TRUE,las=2,
          ylab="log2CPM",main="PDPN",  col = c("red", "red", "red", "blue", "blue", "blue", "green4"))
boxplot(v$E["ENSG00000174059",]~factor(sampleinfo$Population),vertical=TRUE,las=2,
          ylab="log2CPM",main="CD34", col = c("red", "red", "red", "blue", "blue", "blue", "green4"))
boxplot(v$E["ENSG00000204287",]~factor(sampleinfo$Population),vertical=TRUE,las=2,
          ylab="log2CPM",main="HLA-DRA",  col = c("red", "red", "red", "blue", "blue", "blue", "green4"))
boxplot(v$E["ENSG00000177575",]~factor(sampleinfo$Population),vertical=TRUE,las=2,
          ylab="log2CPM",main="CD163",  col = c("red", "red", "red", "blue", "blue", "blue", "green4"))
boxplot(v$E["ENSG00000260314",]~factor(sampleinfo$Population),vertical=TRUE,las=2,
          ylab="log2CPM",main="MRC1",  col = c("red", "red", "red", "blue", "blue", "blue", "green4"))
boxplot(v$E["ENSG00000126353",]~factor(sampleinfo$Population),vertical=TRUE,las=2,
          ylab="log2CPM",main="CCR7",  col = c("red", "red", "red", "blue", "blue", "blue", "green4"))
boxplot(v$E["ENSG00000188389",]~group,vertical=TRUE,las=2,cex.axis=0.8,
          ylab="log2CPM",main="PDCD1" )

#Fit the linear model
fit <- lmFit(v)
names(fit)

cont.matrix <- makeContrasts(
  #Comparions of FLS to each other and mphages to each other within disease
  RA.34pv34n=RA.34p - RA.34n,
  RA.34pv90n38p=RA.34p - RA.90n38p,
  RA.34nv90n38p=RA.34n - RA.90n38p,
  RA.206p163pv206p163n=RA.206p163p - RA.206p163n,
  RA.206p163pvDN=RA.206p163p - RA.DN,
  RA.206p163nvDN=RA.206p163n - RA.DN,

  PD1.34pv34n=PD1.34p - PD1.34n,
  PD1.34pv90n38p=PD1.34p - PD1.90n38p,
  PD1.34nv90n38p=PD1.34n - PD1.90n38p,
  PD1.206p163pv206p163n=PD1.206p163p - PD1.206p163n,
  PD1.206p163pvDN=PD1.206p163p - PD1.DN,
  PD1.206p163nvDN=PD1.206p163n - PD1.DN,

  TreatedRA.34pv34n=TreatedRA.34p - TreatedRA.34n,
  TreatedRA.34pv90n38p=TreatedRA.34p - TreatedRA.90n38p,
  TreatedRA.34nv90n38p=TreatedRA.34n - TreatedRA.90n38p,
  TreatedRA.206p163pv206p163n=TreatedRA.206p163p - TreatedRA.206p163n,
  TreatedRA.206p163pvDN=TreatedRA.206p163p - TreatedRA.DN,
  TreatedRA.206p163nvDN=TreatedRA.206p163n - TreatedRA.DN,

  #Comparison of populations across disease
  RA.34pvPD1.34p=RA.34p - PD1.34p,
```

```
  RA.34pvTreatedRA.34p=RA.34p - TreatedRA.34p,
  PD1.34pvTreatedRA.34p=PD1.34p - TreatedRA.34p,

  RA.34nvPD1.34n=RA.34n - PD1.34n,
  RA.34nvTreatedRA.34n=RA.34n - TreatedRA.34n,
  PD1.34nvTreatedRA.34n=PD1.34n - TreatedRA.34n,

  RA.90n38pvPD1.90n38p=RA.90n38p - PD1.90n38p,
  RA.90n38pvTreatedRA.90n38p=RA.90n38p - TreatedRA.90n38p,
  PD1.90n38pvTreatedRA.90n38p=PD1.90n38p - TreatedRA.90n38p,

  RA.206p163pvPD1.206p163p=RA.206p163p - PD1.206p163p,
  RA.206p163pvTreatedRA.206p163p=RA.206p163p - TreatedRA.206p163p,
  PD1.206p163pvTreatedRA.206p163p=PD1.206p163p - TreatedRA.206p163p,

  RA.206p163nvPD1.206p163n=RA.206p163n - PD1.206p163n,
  RA.206p163nvTreatedRA.206p163n=RA.206p163n - TreatedRA.206p163n,
  PD1.206p163nvTreatedRA.206p163n=PD1.206p163n - TreatedRA.206p163n,

  RA.DNvPD1.DN=RA.DN - PD1.DN,
  RA.DNvTreatedRA.DN=RA.DN - TreatedRA.DN,
  PD1.DNvTreatedRA.DN=PD1.DN - TreatedRA.DN,

  #INTERACTIONS TO ASSESS DIFFERENCES IN SUBSETS BETWEEN DISEASES I.E. 34NV90N38P IN RA V TREATEDRA
  #34nv90n38p
  RA.34nv90n38p.VS.TreatedRA.34nv90n38p=(RA.34n-RA.90n38p)-(TreatedRA.34n-TreatedRA.90n38p),
  RA.34nv90n38p.VS.PD1.34nv90n38p=(RA.34n-RA.90n38p)-(PD1.34n-PD1.90n38p),
  TreatedRA.34nv90n38p.VS.PD1.34n90n38p=(TreatedRA.34n-TreatedRA.90n38p)-(PD1.34n-PD1.90n38p),
  #34pv90n38p
  RA.34pv90n38p.VS.TreatedRA.34pv90n38p=(RA.34p-RA.90n38p)-(TreatedRA.34p-TreatedRA.90n38p),
  RA.34pv90n38p.VS.PD1.34pv90n38p=(RA.34p-RA.90n38p)-(PD1.34p-PD1.90n38p),
  TreatedRA.34pv90n38p.VS.PD1.34p90n38p=(TreatedRA.34p-TreatedRA.90n38p)-(PD1.34p-PD1.90n38p),
  #34nv34p
  RA.34nv34p.VS.TreatedRA.34nv34p=(RA.34n-RA.34p)-(TreatedRA.34n-TreatedRA.34p),
  RA.34nv34p.VS.PD1.34nv34p=(RA.34n-RA.34p)-(PD1.34n-PD1.34p),
  TreatedRA.34nv34p.VS.PD1.34n34p=(TreatedRA.34n-TreatedRA.34p)-(PD1.34n-PD1.34p),
  #206p163pvDN
  RA.206p163pv206p163n.VS.TreatedRA.206p163pv206p163n=(RA.206p163p-RA.206p163n)-(TreatedRA.206p163p-TreatedRA.
206p163n),
  RA.206p163pv206p163n.VS.PD1.206p163pv206p163n=(RA.206p163p-RA.206p163n)-(PD1.206p163p-PD1.206p163n),
  TreatedRA.206p163pv206p163n.VS.PD1.206p163p206p163n=(TreatedRA.206p163p-TreatedRA.206p163n)-(PD1.206p163p-PD
1.206p163n),
  #206p163pvDN
  RA.206p163pvDN.VS.TreatedRA.206p163pvDN=(RA.206p163p-RA.DN)-(TreatedRA.206p163p-TreatedRA.DN),
  RA.206p163pvDN.VS.PD1.206p163pvDN=(RA.206p163p-RA.DN)-(PD1.206p163p-PD1.DN),
  TreatedRA.206p163pvDN.VS.PD1.206p163pDN=(TreatedRA.206p163p-TreatedRA.DN)-(PD1.206p163p-PD1.DN),
  #206p163nvDN
  RA.206p163nvDN.VS.TreatedRA.206p163nvDN=(RA.206p163n-RA.DN)-(TreatedRA.206p163n-TreatedRA.DN),
  RA.206p163nvDN.VS.PD1.206p163nvDN=(RA.206p163n-RA.DN)-(PD1.206p163n-PD1.DN),
  TreatedRA.206p163nvDN.VS.PD1.206p163nDN=(TreatedRA.206p163n-TreatedRA.DN)-(PD1.206p163n-PD1.DN),

  #Do the 206p163p and 206p163n differ from DN in the same way?
  RA.206p163pvDN.VS.RA.206p163nvDN=(RA.206p163p-RA.DN)-(RA.206p163n-RA.DN),
  TreatedRA.206p163pvDN.VS.TreatedRA.206p163nvDN=(TreatedRA.206p163p-TreatedRA.DN)-(TreatedRA.206p163n-Treated
RA.DN),
  PD1.206p163pvDN.VS.PD1.206p163nvDN=(PD1.206p163p-PD1.DN)-(PD1.206p163n-PD1.DN),

  #FLS v mphage test
  RA.34n.v206p163p = RA.34n-RA.206p163p,

  levels=design)

fit.cont <- contrasts.fit(fit, cont.matrix)
fit.cont <- eBayes(fit.cont)
dim(fit.cont)
summa.fit <- decideTests(fit.cont)
summary(summa.fit)
topTable(fit.cont,coef=1,sort.by="p")

#Venn diagrams of DE genes
#Within disease
vennDiagram(summa.fit[,1:3], cex = 0.9)
vennDiagram(summa.fit[,4:6], cex = 0.9)
vennDiagram(summa.fit[,7:9], cex = 0.9)
vennDiagram(summa.fit[,10:12], cex = 0.9)
vennDiagram(summa.fit[,13:15], cex = 0.9)
vennDiagram(summa.fit[,16:18], cex = 0.7)
#Between disease
vennDiagram(summa.fit[,19:21], cex = 0.7)
vennDiagram(summa.fit[,22:24], cex = 0.7)
vennDiagram(summa.fit[,25:27], cex = 0.7)
vennDiagram(summa.fit[,28:30], cex = 0.7)
vennDiagram(summa.fit[,31:33], cex = 0.7)
vennDiagram(summa.fit[,34:36], cex = 0.7)
```

```
#Writing tables of DE genes for all comparisons
write.table(x=topTable(fit.cont,coef=1,sort.by="p", n="Inf"), file = "RA.34pv34n.txt", row.names = FALSE, sep=
"\t")
write.table(x=topTable(fit.cont,coef=2,sort.by="p", n="Inf"), file = "RA.34pv90n38p.txt", row.names=FALSE, sep
="\t")
write.table(x=topTable(fit.cont,coef=3,sort.by="p", n="Inf"), file = "RA.34nv90n38p.txt", row.names=FALSE, sep
="\t")
write.table(x=topTable(fit.cont,coef=4,sort.by="p", n="Inf"), file = "RA.206p163pv206p163n.txt", row.names=FAL
SE, sep="\t")
write.table(x=topTable(fit.cont,coef=5,sort.by="p", n="Inf"), file = "RA.206p163pvDN.txt", row.names=FALSE, se
p="\t")
write.table(x=topTable(fit.cont,coef=6,sort.by="p", n="Inf"), file = "RA.206p163nvDN.txt", row.names=FALSE, se
p="\t")
write.table(x=topTable(fit.cont,coef=7,sort.by="p", n="Inf"), file = "PD1.34pv34n.txt", row.names=FALSE, sep="
\t")
write.table(x=topTable(fit.cont,coef=8,sort.by="p", n="Inf"), file = "PD1.34pv90n38p.txt", row.names=FALSE, se
p="\t")
write.table(x=topTable(fit.cont,coef=9,sort.by="p", n="Inf"), file = "PD1.34nv90n38p.txt", row.names=FALSE, se
p="\t")
write.table(x=topTable(fit.cont,coef=10,sort.by="p", n="Inf"), file = "PD1.206p163pv206p163n.txt", row.names=F
ALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=11,sort.by="p", n="Inf"), file = "PD1.206p163pvDN.txt", row.names=FALSE,
sep="\t")
write.table(x=topTable(fit.cont,coef=12,sort.by="p", n="Inf"), file = "PD1.206p163nvDN.txt", row.names=FALSE,
sep="\t")
write.table(x=topTable(fit.cont,coef=13,sort.by="p", n="Inf"), file = "TreatedRA.34pv34n.txt", row.names=FALSE
, sep="\t")
write.table(x=topTable(fit.cont,coef=14,sort.by="p", n="Inf"), file = "TreatedRA.34pv90n38p.txt", row.names=FA
LSE, sep="\t")
write.table(x=topTable(fit.cont,coef=15,sort.by="p", n="Inf"), file = "TreatedRA.34nv90n38p.txt", row.names=FA
LSE, sep="\t")
write.table(x=topTable(fit.cont,coef=16,sort.by="p", n="Inf"), file = "TreatedRA.206p163pv206p163n.txt", row.n
ames=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=17,sort.by="p", n="Inf"), file = "TreatedRA.206p163pvDN.txt", row.names=F
ALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=18,sort.by="p", n="Inf"), file = "TreatedRA.206p163nvDN.txt", row.names=F
ALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=19,sort.by="p", n="Inf"), file = "RA.34pvPD1.34p.txt", row.names=FALSE, s
ep="\t")
write.table(x=topTable(fit.cont,coef=20,sort.by="p", n="Inf"), file = "RA.34pvTreatedRA.34p.txt", row.names=FA
LSE, sep="\t")
write.table(x=topTable(fit.cont,coef=21,sort.by="p", n="Inf"), file = "PD1.34pvTreatedRA.34p.txt", row.names=F
ALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=22,sort.by="p", n="Inf"), file = "RA.34nvPD1.34n.txt", row.names=FALSE, s
ep="\t")
write.table(x=topTable(fit.cont,coef=23,sort.by="p", n="Inf"), file = "RA.34nvTreatedRA.34n.txt", row.names=FA
LSE, sep="\t")
write.table(x=topTable(fit.cont,coef=24,sort.by="p", n="Inf"), file = "PD1.34nvTreatedRA.34n.txt", row.names=F
ALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=25,sort.by="p", n="Inf"), file = "RA.90n38pvPD1.90n38p.txt", row.names=FA
LSE, sep="\t")
write.table(x=topTable(fit.cont,coef=26,sort.by="p", n="Inf"), file = "RA.90n38pvTreatedRA.90n38p.txt", row.na
mes=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=27,sort.by="p", n="Inf"), file = "PD1.90n38pvTreatedRA.90n38p.txt", row.n
ames=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=28,sort.by="p", n="Inf"), file = "RA.206p163pvPD1.206p163p.txt", row.name
s=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=29,sort.by="p", n="Inf"), file = "RA.206p163pvTreatedRA.206p163p.txt", ro
w.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=30,sort.by="p", n="Inf"), file = "PD1.206p163pvTreatedRA.206p163p.txt", r
ow.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=31,sort.by="p", n="Inf"), file = "RA.206p163nvPD1.206p163n.txt", row.name
s=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=32,sort.by="p", n="Inf"), file = "RA.206p163nvTreatedRA.206p163n.txt", ro
w.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=33,sort.by="p", n="Inf"), file = "PD1.206p163nvTreatedRA.206p163n.txt", r
ow.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=34,sort.by="p", n="Inf"), file = "RA.DNvPD1.DN.txt", row.names=FALSE, sep
="\t")
write.table(x=topTable(fit.cont,coef=35,sort.by="p", n="Inf"), file = "RA.DNvTreatedRA.DN.txt", row.names=FALS
E, sep="\t")
write.table(x=topTable(fit.cont,coef=36,sort.by="p", n="Inf"), file = "PD1.DNvTreatedRA.DN.txt", row.names=FAL
SE, sep="\t")
write.table(x=topTable(fit.cont,coef=37,sort.by="p", n="Inf"), file = "RA.34nv90n38p.VS.TreatedRA.34nv90n38p.t
xt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=38,sort.by="p", n="Inf"), file = "RA.34nv90n38p.VS.PD1.34nv90n38p.txt", r
ow.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=39,sort.by="p", n="Inf"), file = "TreatedRA.34nv90n38p.VS.PD1.34n90n38p.t
xt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=40,sort.by="p", n="Inf"), file = "RA.34pv90n38p.VS.TreatedRA.34pv90n38p.t
xt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=41,sort.by="p", n="Inf"), file = "RA.34pv90n38p.VS.PD1.34pv90n38p.txt", r
ow.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=42,sort.by="p", n="Inf"), file = "TreatedRA.34pv90n38p.VS.PD1.34p90n38p.t
```

```
xt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=43,sort.by="p", n="Inf"), file = "RA.34nv34p.VS.TreatedRA.34nv34p.txt", r
ow.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=44,sort.by="p", n="Inf"), file = "RA.34nv34p.VS.PD1.34nv34p.txt", row.nam
es=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=45,sort.by="p", n="Inf"), file = "TreatedRA.34nv34p.VS.PD1.34n34p.txt", r
ow.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=46,sort.by="p", n="Inf"), file = "RA.206p163pv206p163n.VS.TreatedRA.206p1
63pv206p163n.txt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=47,sort.by="p", n="Inf"), file = "RA.206p163pv206p163n.VS.PD1.206p163pv20
6p163n.txt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=48,sort.by="p", n="Inf"), file = "TreatedRA.206p163pv206p163n.VS.PD1.206p
163p206p163n.txt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=49,sort.by="p", n="Inf"), file = "RA.206p163pvDN.VS.TreatedRA.206p163pvDN
.txt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=50,sort.by="p", n="Inf"), file = "RA.206p163pvDN.VS.PD1.206p163pvDN.txt",
row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=51,sort.by="p", n="Inf"), file = "TreatedRA.206p163pvDN.VS.PD1.206p163pDN
.txt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=52,sort.by="p", n="Inf"), file = "RA.206p163nvDN.VS.TreatedRA.206p163nvDN
.txt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=53,sort.by="p", n="Inf"), file = "RA.206p163nvDN.VS.PD1.206p163nvDN.txt",
row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=54,sort.by="p", n="Inf"), file = "TreatedRA.206p163nvDN.VS.PD1.206p163nDN
.txt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=55,sort.by="p", n="Inf"), file = "RA.206p163pvDN.VS.RA.206p163nvDN.txt",
row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=56,sort.by="p", n="Inf"), file = "TreatedRA.206p163pvDN.VS.TreatedRA.206p
163nvDN.txt", row.names=FALSE, sep="\t")
write.table(x=topTable(fit.cont,coef=57,sort.by="p", n="Inf"), file = "PD1.206p163pvDN.VS.PD1.206p163nvDN.txt"
, row.names=FALSE, sep="\t")

#Checking DE using plots
par(mfrow=c(1,2))
plotMD(fit.cont,coef=1,status=summa.fit[,"RA.34pv34n"])
volcanoplot(fit.cont,coef=1,highlight=100,names=fit.cont$genes$hgnc_symbol)

glXYPlot(x=fit.cont$coefficients[,1], y=fit.cont$lods[,1],
         xlab="logFC", ylab="B", main="RA.34pv34n",
         counts=y$counts, groups=group, status=summa.fit[,1],
         anno=fit.cont$genes, side.main="ensembl_gene_id", folder="volcanoRA.34pv34n")


glMDPlot(fit.cont, coef=1, counts=v$E, groups=group,
         status=summa.fit, id.column="ensembl_gene_id", main="RA.34pv34n",
         folder="mdRA.34pv34n")


#GSEA
#camera
#Using msigdb gene sets
load("human_H_v5p2.rdata")#Hallmark
load("human_c2_v5p2.rdata")#Curated gene sets (Kegg etc)
load("human_c5_v5p2.rdata")#GO
load("human_c6_v5p2.rdata")#Oncogenic
load("human_c7_v5p2.rdata")#Immunologic
names(Hs.H)[1:5]
length(Hs.c2)


H.ind <- ids2indices(Hs.H, v$genes$entrezgene)
gst.camera.H <- camera(v,index=H.ind,design=design,contrast = cont.matrix[,1],inter.gene.cor=0.05)
gst.camera.H[1:5,] #Top 5 results
table(gst.camera.H$FDR < 0.25)
write.csv(gst.camera.H,file="gstcameraH_RA.34pv34n.csv")

c2.ind <- ids2indices(Hs.c2, v$genes$entrezgene)
gst.camera.c2 <- camera(v,index=c2.ind,design=design,contrast = cont.matrix[,1],inter.gene.cor=0.05)
gst.camera.c2[1:5,] #Top 5 results
table(gst.camera.c2$FDR < 0.25)
write.csv(gst.camera.c2,file="gstcamerac2_RA.34pv34n.csv")

c5.ind <- ids2indices(Hs.c5, v$genes$entrezgene)
gst.camera.c5 <- camera(v,index=c5.ind,design=design,contrast = cont.matrix[,1],inter.gene.cor=0.05)
gst.camera.c5[1:5,] #Top 5 results
table(gst.camera.c5$FDR < 0.25)
write.csv(gst.camera.c5,file="gstcamerac2_RA.34pv34n.csv")

c6.ind <- ids2indices(Hs.c6, v$genes$entrezgene)
gst.camera.c6 <- camera(v,index=c6.ind,design=design,contrast = cont.matrix[,1],inter.gene.cor=0.05)
gst.camera.c6[1:5,] #Top 5 results
table(gst.camera.c6$FDR < 0.25)
write.csv(gst.camera.c6,file="gstcamerac6_RA.34pv34n.csv")

c7.ind <- ids2indices(Hs.c7, v$genes$entrezgene)
```

```
gst.camera.c7 <- camera(v,index=c7.ind,design=design,contrast = cont.matrix[,1],inter.gene.cor=0.05)
gst.camera.c7[1:5,] #Top 5 results
table(gst.camera.c7$FDR < 0.25)
write.csv(gst.camera.c7,file="gstcamerac6_RA.34pv34n.csv")

#Plotting enrichment
head(fit.cont$coefficients)
head(fit.cont$t)
par(mfrow=c(1,1))
# barcode plot with logFCs, edit gene set names as required
barcodeplot(fit.cont$coeff[,1], index=c2.ind[["Gene set of interest"]], main="LogFC: Gene set name") #Edit nam
es as required
# barcode plot using t-statistics, edit gene set names as required
barcodeplot(fit.cont$t[,1], index=c2.ind[["Gene set of interest"]], main="T-statistic: Gene set name") #Edit n
ames as required

#GO and Kegg enrichment with gene length bias correction, repeat for each comparison tested
library(goseq)
#Macrophage
RA.206p163pvDN <- topTable(fit.cont,coef=5,sort.by="p", n="Inf")
genes.up=as.integer(RA.206p163pvDN$adj.P.Val[RA.206p163pvDN$logFC>0]<.05)
names(genes.up)=row.names(RA.206p163pvDN[RA.206p163pvDN$logFC>0,])
table(genes.up)
pwf=nullp(genes.up,"hg19","ensGene")
head(pwf)
GO.BP=goseq(pwf,"hg19","ensGene",test.cats=c("GO:BP"))
head(GO.BP)
GO.BP$adj.over_represented_pvalue <- p.adjust(GO.BP$over_represented_pvalue, method="BH")
GO.BP.sigover <- GO.BP[GO.BP$adj.over_represented_pvalue<0.05,]
enriched.GO=GO.BP[p.adjust(GO.BP$over_represented_pvalue, method="BH")<.05,]
head(enriched.GO)
library(GO.db)
for(go in enriched.GO[1:56, 1]){
  print(GOTERM[[go]])
  cat("--------------------------------------\n")
  }

KEGG=goseq(pwf,'hg19','ensGene',test.cats="KEGG")
head(KEGG)
KEGG$adj.over_represented_pvalue <- p.adjust(KEGG$over_represented_pvalue, method="BH")
KEGG.sigover <- KEGG[KEGG$adj.over_represented_pvalue<0.05,]
enriched.KEGG=KEGG$category[p.adjust(KEGG$over_represented_pvalue, method="BH")<.05]
head(enriched.KEGG)

genes.down=as.integer(RA.206p163pvDN$adj.P.Val[RA.206p163pvDN$logFC<0]<.05)
names(genes.down)=row.names(RA.206p163pvDN[RA.206p163pvDN$logFC<0,])
table(genes.down)
pwf=nullp(genes.down,"hg19","ensGene")
head(pwf)
GO.BP=goseq(pwf,"hg19","ensGene",test.cats=c("GO:BP"))
head(GO.BP)
GO.BP$adj.over_represented_pvalue <- p.adjust(GO.BP$over_represented_pvalue, method="BH")
GO.BP.sigover <- GO.BP[GO.BP$adj.over_represented_pvalue<0.05,]
enriched.GO=GO.BP$category[p.adjust(GO.BP$over_represented_pvalue, method="BH")<.05]
head(enriched.GO)
library(GO.db)
for(go in enriched.GO[1:62]){
  print(GOTERM[[go]])
  cat("--------------------------------------\n")
}

#FLS
RA.34pv34n <- topTable(fit.cont,coef=1,sort.by="p", n="Inf")
RA.34pv90n38p <- topTable(fit.cont,coef=2,sort.by="p", n="Inf")
RA.34nv90n38p <- topTable(fit.cont,coef=3,sort.by="p", n="Inf")
genes.up=as.integer(RA.34nv90n38p$adj.P.Val[RA.34nv90n38p$logFC<0]<.05)
names(genes.up)=row.names(RA.34nv90n38p[RA.34nv90n38p$logFC<0,])
table(genes.up)
pwf=nullp(genes.up,"hg19","ensGene")
head(pwf)
GO.BP=goseq(pwf,"hg19","ensGene",test.cats=c("GO:BP"))
head(GO.BP)
GO.BP$adj.over_represented_pvalue <- p.adjust(GO.BP$over_represented_pvalue, method="BH")
GO.BP.sigover <- GO.BP[GO.BP$adj.over_represented_pvalue<0.05,]
```

8.5 R script for the analysis of microarray data from fibroblast cultures

```
library("BiocInstaller")
library(Biobase)
library("limma")
library("genefilter")
library("RColorBrewer")
library("ggplot2")
library("ggfortify")
library("human.db0")
library("AMADID039494.db")
library("arrayQualityMetrics")
library("Glimma")
library("gplots")

#Line below for loading data if saved during analysis and returned to
load("MicroarrayAnalysis.RData")

#Read in raw data from feature extraction files into red/green format
targets <- readTargets(file="targets.txt")
RG_allsample <- read.maimages(targets, path="All OGT Agilent fibroblast array raw data files",
                              source="agilent")
targets$Group <- paste(targets$Diagnosis, targets$Serum, targets$Stimulated, sep=".")
targets$StimSerum <- paste(targets$Stimulated, targets$Serum, sep=".")

#Swap colours for correct ratio calculation
RG_allsampleRGSwap <- RG_allsample
RG_allsampleRGSwap$G <- RG_allsample$R
RG_allsampleRGSwap$Gb <- RG_allsample$Rb
RG_allsampleRGSwap$Rb <- RG_allsample$Gb
RG_allsampleRGSwap$R <- RG_allsample$G

genelist <- RG_allsampleRGSwap$genes


#Creating annotation database
a <- genelist$ProbeName
b <- genelist$SystematicName
AgilentAMADID039494ID <- cbind(a, b)
makeDBPackage("HUMANCHIP_DB",
              affy=FALSE,
              prefix="AMADID039494",
              fileName="C:/Users/JDT366/Desktop/Agilent analysis WD/Annotationtest.txt",
              baseMapType="gbNRef",
              outputDir = "C:/Users/JDT366/Desktop/Agilent analysis WD/Annotation Test",
              version="1.0.0",
              manufacturer = "Agilent",
              chipName = "agilent_sureprint_g3_ge_8x60k_v2",
              manufacturerUrl = "http://www.agilent.co.uk")
install.packages("D:/Jason's Folder/Microarray/Agilent analysis WD/Annotation Test/AMADID039494.
db ", repos=NULL, type="source")
library("AMADID039494.db")

#Diagnostic plots
plotDensities(RG_allsampleRGSwap, main = "Red and Green channel densities for each array")
imageplot(RG_allsampleRGSwap$R[,1], layout = RG_allsampleRGSwap$printer)

#Normalising arrays, background correction, and log2 ratio calculation
MA_allsampleRGSwap <- normalizeWithinArrays(RG_allsampleRGSwap, method="loess", bc.method="norme
xp", offset=50)
plotMA(MA_allsampleRGSwap[,2])
plotDensities(MA_allsampleRGSwap)
MA_RGSwapNorm <- normalizeBetweenArrays(MA_allsampleRGSwap, method="Gquantile")
plotDensities(MA_RGSwapNorm)
imageplot(MA_RGSwapNorm$M[,17], layout=MA_RGSwapNorm$printer)

# Odd MA plots, 9, 17, 44, 67, 69, 73, 80, 145

adf <- new("AnnotatedDataFrame", data=targets)
adfgene <- new("AnnotatedDataFrame", data=genelist)
experimentData <- new("MIAME", name="Maria Juarez, Jason Turner", lab="Andrew Filer", contact="a
.filer@bham.ac.uk", title="Gene expression in RA longitudinal synovial fibroblasts in response t
o TNF and/or Serum")
```

```
esetMA_RGSwapNorm <- new("ExpressionSet", exprs=MA_RGSwapNorm$M, phenoData= adf, experimentData=
experimentData, featureData=adfgene, annotation="AMADID039494")
arrayQualityMetrics(esetMA_RGSwapNorm, outdir = "QC", intgroup = "StimSerum")

#Arrayqualitymetrics outliers 30, 51, 52, 53, 67, 83, 129, 134, 135, 139, 140

#Excluding manual MA plot checks and all ArrayQualityMetrics outliers
dim(esetMA_RGSwapNorm)
table(esetMA_RGSwapNorm$Group)
esetMA_RGSwapNormClean <- esetMA_RGSwapNorm[,c(-9,-17,-30,-44,-51,-52,-53,-67,-69,-73,-80,-83,-1
29,-134, -139,-140,-145)]
dim(esetMA_RGSwapNormClean)
table(esetMA_RGSwapNormClean$Group)

dim(MA_RGSwapNorm)
MA_RGSwapNormClean <- MA_RGSwapNorm[,c(-9,-17,-30,-44,-51,-52,-53,-67,-69,-73,-80,-83,-129,-134,
-139,-140,-145)]

#Automated imageplotting for checking artefacts
for(i in 1:132){
  imageplot(MA_RGSwapNormCleanCsex$M[,i], layout=MA_RGSwapNormCleanCsex$printer, main = i)
  Sys.sleep(1.5)
}
#7,38, 55

#Checking sample sex using XIST expression and removing samples which seemingly have incorrect s
ex info
XISTexprs <- exprs(esetMA_RGSwapNormClean)[subsetting<-grep("XIST", esetMA_RGSwapNormClean@featu
reData$GeneName),]
XISTexprsmean <- apply(XISTexprs, 2, mean)
hist(XISTexprsmean)
sex <- ifelse(XISTexprsmean>-1.2, "F", "M")
CorrectSex <- as.matrix(esetMA_RGSwapNormClean$Gender == sex)
table(CorrectSex)
CorrectSex
Csex <- grep("TRUE", CorrectSex)
MA_RGSwapNormCleanCsex <- MA_RGSwapNormClean[,Csex]
dim(MA_RGSwapNormCleanCsex)


#Averaging replicate probes
AvExprs<- avereps(MA_RGSwapNormCleanCsex$M, ID=MA_RGSwapNormCleanCsex$genes$ProbeUID)
AvGenelist <- MA_RGSwapNormCleanCsex$genes[match(unique(MA_RGSwapNormCleanCsex$genes$ProbeName),
                                                 MA_RGSwapNormCleanCsex$genes$ProbeName),]
rownames(AvGenelist) <- AvGenelist$ProbeName
rownames(AvExprs) <- rownames(AvGenelist)
#Removing previous esets
rm(esetMA_RGSwapNorm)
rm(esetMA_RGSwapNormClean)

#Generating new expression set
adf <- new("AnnotatedDataFrame", data=MA_RGSwapNormCleanCsex$targets)
adfgene <- new("AnnotatedDataFrame", data=AvGenelist)
experimentData <- new("MIAME", name="Maria Juarez, Jason Turner", lab="Andrew Filer", contact="a
.filer@bham.ac.uk", title="Gene expression in RA longitudinal synovial fibroblasts in response t
o TNF and/or Serum")
eset_RGSwap.Norm.Clean.Csex.AveRep <- new("ExpressionSet", exprs=AvExprs, phenoData= adf, experi
mentData=experimentData, featureData=adfgene, annotation="AMADID039494")
dim(eset_RGSwap.Norm.Clean.Csex.AveRep)

#Non-specific filtering
eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt <- nsFilter(eset_RGSwap.Norm.Clean.Csex.AveRep, requir
e.entrez = FALSE, remove.dupEntrez = FALSE, var.func = sd, var.cutoff = 0.4)
dim(eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset)
eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset@phenoData$StimSerum <- paste(eset_RGSwap.Norm.Cle
an.Csex.AveRep.NSFilt$eset@phenoData$Stimulated, eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset@
phenoData$Serum, sep=".")
arrayQualityMetrics(eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset, outdir = "QCEset", intgroup
= "StimSerum")

#---- End of data preparation


#PCA, basic scripts but then modified to investigate other PCs
```

```
pca <- prcomp(t(exprs(eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset)))
autoplot(pca, x= 1, y= 2, data = pData(eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset),
         colour = 'StimSerum', size = 3,
         frame = TRUE, frame.type = 'norm', xlab="PC1: 29.08% of variance", ylab="PC2: 11.01% of
variance")

#Subset data in treatment groups
esetHigh <- eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset[,grep("High", eset_RGSwap.Norm.Clean.
Csex.AveRep.NSFilt$eset@phenoData$Serum)]
table(esetHigh@phenoData$Serum)
esetHighTNF <- esetHigh[,grep("TNF", esetHigh@phenoData$Stimulated)]
table(esetHighTNF@phenoData$Stimulated)
esetHighUnstim <- esetHigh[,grep("Unstim", esetHigh@phenoData$Stimulated)]
table(esetHighUnstim@phenoData$Stimulated)
esetLow <- eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset[,grep("Low", eset_RGSwap.Norm.Clean.Cs
ex.AveRep.NSFilt$eset@phenoData$Serum)]
table(esetLow@phenoData$Serum)

#PCA for individual treatment groups
HiTNF.pca <- prcomp(t(exprs(esetHighTNF)))
autoplot(HiTNF.pca, x= 1, y= 2, data = pData(esetHighTNF),
         colour = 'Diagnosis', size = 3,
         frame = TRUE, frame.type = 'norm')

HiUnstim.pca <- prcomp(t(exprs(esetHighUnstim)))
autoplot(HiUnstim.pca, x= 1, y= 2, data = pData(esetHighUnstim),
         colour = 'Diagnosis', size = 3,
         frame = TRUE, frame.type = 'norm')

Low.pca <- prcomp(t(exprs(esetLow)))
autoplot(Low.pca, x= 1, y= 2, data = pData(esetLow),
         colour = 'Diagnosis', size = 3,
         frame = TRUE, frame.type = 'norm')


# 3D PCA plots for visualisation
library(pca3d)
pca3d(HiTNF.pca, components = 3:6, axes.color = "grey", bg = "white", radius = 1, labels.col = "
black",
      show.axes = TRUE, group=factor(esetHighTNF@phenoData$Diagnosis),
      show.axe.titles = TRUE,  show.plane = TRUE,
      show.shapes = TRUE, show.centroids = TRUE, show.ellipses = TRUE, legend= "bottomleft")



#Fitting a model to sex and sit then using residuals for second model fit to disease
design <- model.matrix(~ 0 + eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset@phenoData$Site
                       + eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset@phenoData$Gender)
lmFitPatient <- lmFit(eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset, design)
resid <- residuals.MArrayLM(lmFitPatient, y= exprs(eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$ese
t))

design <- model.matrix(~ 0 + eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset@phenoData$Group)
colnames(design) <- levels(eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt$eset@phenoData$Group)

eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt.Resid <-  new("ExpressionSet", exprs=resid,
                                                        phenoData= new("AnnotatedDataFrame",
                                                                       pData(eset_RGSwap.Norm.Cl
ean.Csex.AveRep.NSFilt$eset)),
                                                        experimentData=experimentData,
                                                        featureData=new("AnnotatedDataFrame",
                                                                        fData(eset_RGSwap.Norm.C
lean.Csex.AveRep.NSFilt$eset)),
                                                        annotation="AMADID039494")
dim(eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt.Resid)

fitPatientremoved <- lmFit(eset_RGSwap.Norm.Clean.Csex.AveRep.NSFilt.Resid, design)

#Contrast matrix
cont.matrix <- makeContrasts(
  #Within treatment comparisons (What is DE between outcomes treated the same)
  TNFJrepvEstRA=Jrep.High.TNF-EstRA.High.TNF,
  TNFJrepvEarly=Jrep.High.TNF-Early.High.TNF,
  TNFJrepvRes=Jrep.High.TNF-Res.High.TNF,
```

```
  TNFJrepvNorm=Jrep.High.TNF-Norm.High.TNF,
  TNFEstRAvEarly=EstRA.High.TNF-Early.High.TNF,
  TNFEstRAvRes=EstRA.High.TNF-Res.High.TNF,
  TNFEstRAvNorm=EstRA.High.TNF-Norm.High.TNF,
  TNFEarlyvRes=Early.High.TNF-Res.High.TNF,
  TNFEarlyvNorm=Early.High.TNF-Norm.High.TNF,
  TNFResvNorm=Res.High.TNF-Norm.High.TNF,

  HighJrepvEstRA=Jrep.High.Unstim-EstRA.High.Unstim,
  HighJrepvEarly=Jrep.High.Unstim-Early.High.Unstim,
  HighJrepvRes=Jrep.High.Unstim-Res.High.Unstim,
  HighJrepvNorm=Jrep.High.Unstim-Norm.High.Unstim,
  HighEstRAvEarly=EstRA.High.Unstim-Early.High.Unstim,
  HighEstRAvRes=EstRA.High.Unstim-Res.High.Unstim,
  HighEstRAvNorm=EstRA.High.Unstim-Norm.High.Unstim,
  HighEarlyvRes=Early.High.Unstim-Res.High.Unstim,
  HighEarlyvNorm=Early.High.Unstim-Norm.High.Unstim,
  HighResvNorm=Res.High.Unstim-Norm.High.Unstim,

  LowJrepvEstRA=Jrep.Low.Unstim-EstRA.Low.Unstim,
  LowJrepvEarly=Jrep.Low.Unstim-Early.Low.Unstim,
  LowJrepvRes=Jrep.Low.Unstim-Res.Low.Unstim,
  LowJrepvNorm=Jrep.Low.Unstim-Norm.Low.Unstim,
  LowEstRAvEarly=EstRA.Low.Unstim-Early.Low.Unstim,
  LowEstRAvRes=EstRA.Low.Unstim-Res.Low.Unstim,
  LowEstRAvNorm=EstRA.Low.Unstim-Norm.Low.Unstim,
  LowEarlyvRes=Early.Low.Unstim-Res.Low.Unstim,
  LowEarlyvNorm=Early.Low.Unstim-Norm.Low.Unstim,
  LowResvNorm=Res.Low.Unstim-Norm.Low.Unstim,

  #Between treatment contrasts (How does treatment affect each outcome group)
  JrepTNFvHigh=Jrep.High.TNF-Jrep.High.Unstim,
  JrepHighvLow=Jrep.High.Unstim-Jrep.Low.Unstim,

  EstRATNFvHigh=EstRA.High.TNF-EstRA.High.Unstim,
  EstRAHighvLow=EstRA.High.Unstim-EstRA.Low.Unstim,

  EarlyTNFvHigh=Early.High.TNF-Early.High.Unstim,
  EarlyHighvLow=Early.High.Unstim-Early.Low.Unstim,

  ResTNFvHigh=Res.High.TNF-Res.High.Unstim,
  ResHighvLow=Res.High.Unstim-Res.Low.Unstim,

  NormTNFvHigh=Norm.High.TNF-Norm.High.Unstim,
  NormHighvLow=Norm.High.Unstim-Norm.Low.Unstim,

  #Differences in TNF response from high serum samples between outcomes
  #(What is different in the response to TNF between outcome groups in high serum)
  JrepHighStimUnsvsEstRAHighStimUns=(Jrep.High.TNF-Jrep.High.Unstim)-(EstRA.High.TNF-EstRA.High.
Unstim),
  JrepHighStimUnsvsEarlyHighStimUns=(Jrep.High.TNF-Jrep.High.Unstim)-(Early.High.TNF-Early.High.
Unstim),
  JrepHighStimUnsvsResHighStimUns=(Jrep.High.TNF-Jrep.High.Unstim)-(Res.High.TNF-Res.High.Unstim
),
  JrepHighStimUnsvsNormHighStimUns=(Jrep.High.TNF-Jrep.High.Unstim)-(Norm.High.TNF-Norm.High.Uns
tim),

  EstRAHighStimUnsvsEarlyHighStimUns=(EstRA.High.TNF-EstRA.High.Unstim)-(Early.High.TNF-Early.Hi
gh.Unstim),
  EstRAHighStimUnsvsResHighStimUns=(EstRA.High.TNF-EstRA.High.Unstim)-(Res.High.TNF-Res.High.Uns
tim),
  EstRAHighStimUnsvsNormHighStimUns=(EstRA.High.TNF-EstRA.High.Unstim)-(Norm.High.TNF-Norm.High.
Unstim),

  EarlyHighStimUnsvsResHighStimUns=(Early.High.TNF-Early.High.Unstim)-(Res.High.TNF-Res.High.Uns
tim),
  EarlyHighStimUnsvsNormHighStimUns=(Early.High.TNF-Early.High.Unstim)-(Norm.High.TNF-Norm.High.
Unstim),

  ResHighStimUnsvsNormHighStimUns=(Res.High.TNF-Res.High.Unstim)-(Norm.High.TNF-Norm.High.Unstim
),

  #Differences in serum response between outcomes
  #(What is different in the response to serum between outcome groups)
```

```
  JrepUnsHighLowvsEstRAUnsHighLow=(Jrep.High.Unstim-Jrep.Low.Unstim)-(EstRA.High.Unstim-EstRA.Lo
w.Unstim),
  JrepUnsHighLowvsEarlyUnsHighLow=(Jrep.High.Unstim-Jrep.Low.Unstim)-(Early.High.Unstim-Early.Lo
w.Unstim),
  JrepUnsHighLowvsResUnsHighLow=(Jrep.High.Unstim-Jrep.Low.Unstim)-(Res.High.Unstim-Res.Low.Unst
im),
  JrepUnsHighLowvsNormUnsHighLow=(Jrep.High.Unstim-Jrep.Low.Unstim)-(Norm.High.Unstim-Norm.Low.U
nstim),

  EstRAUnsHighLowvsEarlyUnsHighLow=(EstRA.High.Unstim-EstRA.Low.Unstim)-(Early.High.Unstim-Early
.Low.Unstim),
  EstRAUnsHighLowvsResUnsHighLow=(EstRA.High.Unstim-EstRA.Low.Unstim)-(Res.High.Unstim-Res.Low.U
nstim),
  EstRAUnsHighLowvsNormUnsHighLow=(EstRA.High.Unstim-EstRA.Low.Unstim)-(Norm.High.Unstim-Norm.Lo
w.Unstim),

  EarlyUnsHighLowvsResUnsHighLow=(Early.High.Unstim-Early.Low.Unstim)-(Res.High.Unstim-Res.Low.U
nstim),
  EarlyUnsHighLowvsNormUnsHighLow=(Early.High.Unstim-Early.Low.Unstim)-(Norm.High.Unstim-Norm.Lo
w.Unstim),

  ResUnsHighLowvsNormUnsHighLow=(Res.High.Unstim-Res.Low.Unstim)-(Norm.High.Unstim-Norm.Low.Unst
im),

  levels=design)

fit.contPR <- contrasts.fit(fitPatientremoved, cont.matrix)
fit.contPR <- eBayes(fit.contPR, trend = TRUE)
summa.fitPR <- decideTests(fit.contPR)
summary(summa.fitPR)

#Generating tables of DE genes for assesment, functions modified to generate table for each comp
arison
topTable(fit.contPR, coef = "LowJrepvEstRA", number = "inf", sort.by = "p")
write.table(topTable(fit.contPR, coef = "TNFJrepvEstRA", number = "inf",
                     sort.by = "p", p.value = 0.05), file= "TNFJrepvEstRA.txt",
            row.names=FALSE, sep="\t")

#After DE gene lists manually curated and specific genes selected the genes are imported as an i
ndex and used to
#select only those genes in the comparisons
IndexDups <- read.table(file="/DE genes after dual models one for donor variation/TNF/IndexwithD
ups.csv", sep = ",")
Index <- IndexDups[!duplicated(IndexDups[,1]),1]
Index <- read.table(file="DE genes after dual models one for donor variation/High/HighIndex.csv"
, sep = ",")
Index <- Index[,1]
Index <- read.table(file="DE genes after dual models one for donor variation/Low/LowIndex.csv",
sep = ",")
Index <- Index[,1]
write.table(topTable(fit.contPR, coef = "LowResvNorm", number = "inf",
                     sort.by = "p", p.value = 0.05)[which(topTable(fit.contPR, coef = "LowResvNo
rm", number = "inf",
                                                                  sort.by = "p", p.value = 0.05
)$GeneName %in% Index),],
            file= "LowResvNormIndexed.txt",
            row.names=FALSE, sep="\t")
```

8.6 Differentially expressed gene lists from the TNF stimulated samples in the microarray experiments

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
HAS3	NM_005329	1.054702895	0.036434211
IL6	ENST00000420258	1.036563257	0.01672477
COX1	ENST00000361624	1.002815221	0.020692944
TSLP	NM_033035	0.838692824	0.043702812
IL4R	NM_001008699	0.824232506	0.000207513
IL34	NM_152456	0.649308856	0.020359111
COQ10B	NM_025147	0.577607612	0.028807291
RASA2	NM_006506	0.469517126	0.026552842
SERPINA4	NM_006215	-0.27276859	0.039570443
COQ4	ENST00000372875	-0.291628217	0.013266841
TFAP2B	NM_003221	-0.310597678	0.022509149
SYK	NM_003177	-0.312360373	0.015280483
SERPINA11	NM_001080451	-0.315908297	0.024683341
COX6B2	NM_144613	-0.357205493	0.042127734
PRODH2	NM_021232	-0.357826783	0.047409007
IGSF1	NM_205833	-0.367907039	0.022412505
SCARF2	NM_153334	-0.368129934	0.029150229
LTB4R	NM_181657	-0.369842067	0.044842034
JAG2	NM_002226	-0.377159478	0.008659967
NKAPP1	NR_027131	-0.381701206	0.03398931
GZMM	NM_005317	-0.388316219	0.047812223
CARD14	NM_024110	-0.401338952	0.001675318
SERPINB5	NM_002639	-0.404592656	0.031972296
SERPINF2	NM_000934	-0.411482236	0.007020224
HYAL4	NM_012269	-0.413309382	0.023995728
ALDH1L1	NM_012190	-0.416419317	0.010569011
PTPRC	NM_002838	-0.435473418	0.009750039
CCR2	NM_001123041	-0.440047841	0.035002169
DUSP9	NM_001395	-0.442209666	0.021787886
CD8B	NM_172102	-0.452619432	0.018455135
SOX10	NM_006941	-0.464535869	0.004956761
CEBPD	NM_005195	-0.479820555	0.009755944
FGFR4	NM_213647	-0.490677476	0.002823283
OSCAR	NM_206818	-0.49149354	0.018022606
CD3G	NM_000073	-0.524078376	0.015664904
IGSF1	NM_001555	-0.525779399	0.001542396
ZAP70	NM_001079	-0.540821832	0.003781331
FGA	NM_021871	-0.585354185	0.000774919
LCK	NM_005356	-0.595633057	0.003621978
SMAD1	NM_005900	-0.662092385	0.048686729
RUNX1	NM_001001890	-0.702080226	0.026728522
SERPINA3	NM_001085	-0.745231337	0.024683341
ZFP36	NM_003407	-0.961766073	0.047528418
COL7A1	NM_000094	-1.275588743	0.009791834
CD86	NM_006889	-1.534046141	0.000774919
MMP9	NM_004994	-1.550404411	0.03395765
FOSB	NM_006732	-2.020307681	0.01672477
FOS	NM_005252	-2.051984537	0.005459512

Table 66: Selected significantly differentially expressed genes between JRep and EstRA fibroblasts in the stimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
HAS3	NM_005329	1.388361371	0.001314505
CLEC3B	NM_003278	0.882403683	0.014027336
IL6	ENST00000420258	0.843053943	0.045422769
FGFR1	NM_001174066	0.720790581	0.025589189
TP53	NM_000546	0.473628873	0.001842865
CDK8	NM_001260	0.461395358	0.030149868
ADAM1	NR_036636	0.401599092	0.040740541
TFAP2A	ENST00000478375	0.27549022	0.037181421
SOX13	ENST00000367203	0.275129638	0.033419206
IKBKE	NM_014002	-0.433752833	0.011167086
RUNX1	NM_001122607	-0.467088011	0.016944944
EGR4	NM_001965	-0.631455659	0.047512838
ICAM4	NM_022377	-0.648142081	0.044741214
BCL6	NM_001130845	-0.732160211	0.045440797
SMAD1	NM_005900	-0.764398121	0.009307684
HIST2H2BE	NM_003528	-0.82659043	0.014068168
RUNX1	NM_001122607	-1.065735825	0.001983654
ZFP36	NM_003407	-1.220850599	0.003432037
CLIC2	NM_001289	-1.225785884	0.034111787
EGR1	NM_001964	-1.409174392	0.018170981
FGF13	NM_004114	-1.827580743	0.017386367
FOS	NM_005252	-2.283156057	0.001127113

FOSB	NM_006732	-2.491572913	0.001170927

Table 67: Selected significantly differentially expressed genes between JRep and veRA fibroblasts in the stimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
HAS3	NM_005329	1.220568339	0.015890829
IKBKE	NM_014002	-0.440116965	0.018062278
KMO	NM_003679	-0.652033079	0.032306399
TNF	NM_000594	-0.786212456	0.024519039
TLR3	NM_003265	-0.902515387	0.049960793
ZFP36	NM_003407	-0.984824292	0.048576639
CLIC2	NM_001289	-1.279694689	0.043166782
FOSB	NM_006732	-1.827600248	0.0389089
FOS	NM_005252	-1.852227398	0.016089606

Table 68: Selected significantly differentially expressed genes between JRep and Res fibroblasts in the stimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
HAS2	NM_005328	1.6493715	0.031226592
CDH13	NM_001257	1.609310774	0.049496744
CDK6	NM_001259	1.442275755	0.00552671
CD200	NM_001004196	1.416415032	0.001807355
IRX1	NM_024337	1.391417635	0.01109573
TAGLN	NM_001001522	1.356057454	0.004544098
LIMS2	NM_001161404	1.313388544	0.03874136
IL6	ENST00000420258	1.247783845	0.000766548
SERPINE1	NM_000602	1.242720696	0.030802221
HAS3	NM_005329	1.235658372	0.004644338
CDH13	NM_001220491	1.164016697	0.043781079
HIST1H4L	NM_003546	1.14053117	0.03755767
HIST1H4C	NM_003542	1.102695626	0.031591204
TIMP3	NM_000362	1.058173824	0.002665742
HIST1H4B	NM_003544	1.047977395	0.04682513
HMGB3	NM_005342	0.861838574	0.003304848
TSLP	NM_033035	0.852520582	0.021481896
TRAIP	NM_005879	0.842965855	0.046220202
TAGLN	NM_001001522	0.83759227	0.013915622
LIMS2	NM_001161404	0.800592518	0.045026249
SERPINA9	NM_175739	0.734939	0.015377522
CD200	NM_001004196	0.715674783	0.020030768
COL6A1	NM_001848	0.68740846	0.014888864
CASP6	NM_001226	0.423294377	0.01574508
OXSR1	NM_005109	0.407115297	0.044814047
ILF3	NM_004516	0.406824919	0.045026249
CCL17	NM_002987	-0.274552983	0.049766715
CARD14	NM_024110	-0.288585323	0.019267016
SGPP2	ENST00000321276	-0.293577187	0.023349752
RUNX1	NM_001001890	-0.347789165	0.044939921
HPGDS	NM_014485	-0.383558214	0.001100795
IL18BP	NM_173042	-0.414382685	0.046810494
IL17C	NM_013278	-0.446974695	0.046105813
IL15RA	ENST00000379971	-0.453949707	0.027239622
HSD17B7	NM_016371	-0.456440071	0.002406027
STAT3	NM_213662	-0.488457338	0.026734798
IFNAR1	NM_000629	-0.51884563	0.012192049
KMO	NM_003679	-0.528427282	0.002261617
ZFP36L1	ENST00000408913	-0.53023163	0.009946613
IL10RB	NM_000628	-0.548664261	0.022957886
NFKB1	NM_003998	-0.570133675	0.008605139
ALDH6A1	NM_005589	-0.581385757	0.042162771
HSD17B7	NM_016371	-0.597279098	0.00110405
C1RL	NM_016546	-0.602289744	0.001374465
SGPP2	NM_152386	-0.60302848	0.002194067
RUNX1	NM_001001890	-0.624297051	0.033746459
IFNGR2	NM_005534	-0.634977189	0.014936099
HSD17B7	NM_016371	-0.636658897	0.001457456
IL13RA1	NM_001560	-0.652716092	0.024094389
ZFP36L1	NM_004926	-0.65944701	0.044561829
TGFA	NM_003236	-0.66034474	0.028178336
IFNAR2	NM_000874	-0.661848666	0.008913532
OSMR	NM_001168355	-0.665886636	0.002987967
OSMR	NM_003999	-0.689530989	0.007458511
DUSP8	NM_004420	-0.703434358	0.017017269
IL15RA	NM_172200	-0.714050165	0.030377979
JMJD7	NM_001114632	-0.715963215	0.001345267
CD274	NM_014143	-0.735134724	0.003347188
FOXP2	NM_148900	-0.740218665	0.009416389
CD274	NM_014143	-0.768517611	0.000425461
IFNAR2	NM_207585	-0.775137389	4.12E-05
IRAK3	NM_007199	-0.778015073	0.003347188
NFKBIE	NM_004556	-0.838944803	0.000706544

CTSO	NM_001334	-0.861289321	0.016867391
CCL3L3	NM_001001437	-0.86287025	0.043562106
SPHK1	NM_182965	-0.869544118	0.010158493
BCL6	NM_001130845	-0.908790696	0.006944638
C1RL	NM_016546	-0.918542705	0.034086614
TNF	NM_000594	-0.923172673	0.001972305
CD44	NM_001202557	-0.937638469	0.00968053
SMAD1	NM_005900	-0.938082696	0.00071194
LYVE1	NM_006691	-0.941841145	0.020505762
CXCL10	NM_001565	-0.953793859	0.004921272
IFI35	NM_005533	-0.964948123	0.00968053
MMP2	NM_004530	-0.968656737	0.001089634
ALDH2	NM_000690	-0.988034137	0.030698328
DUSP1	NM_004417	-0.999226302	0.030698328
IL15RA	NM_001243539	-1.000799325	0.011020048
TLR3	NM_003265	-1.026124493	0.008573067
CD38	NM_001775	-1.053950124	0.009484482
CSF3	NM_000759	-1.054566136	0.001682605
SOCS2	NM_003877	-1.093477979	0.044179519
C1QTNF5	NM_015645	-1.104644232	0.040394877
CTSL1	NM_001912	-1.148471258	0.008229376
HLA-F	NM_001098478	-1.174982123	0.039154415
PPARG	NM_138711	-1.195556426	0.019516039
CCL4	NM_002984	-1.200036135	0.009799091
PPARG	NM_138711	-1.205511203	0.040394877
ZFP36	NM_003407	-1.213005913	0.003264632
DUSP8	NM_004420	-1.227056592	0.000530288
LBP	NM_004139	-1.304005346	0.004500834
CSF3	NM_000759	-1.337794611	0.011119258
KMO	NM_003679	-1.350190315	2.08E-08
CCL4	NM_002984	-1.371833412	0.007458511
IL36B	NM_173178	-1.416273994	0.022957886
CCL3L3	NM_001001437	-1.497368381	0.003771971
PDGFRL	NM_006207	-1.564402795	0.04368698
CD34	NM_001773	-1.589703908	0.035779095
CX3CL1	NM_002996	-1.595375062	3.89E-05
CTSK	NM_000396	-1.602221735	0.001126142
NOD2	NM_022162	-1.658431646	7.18E-05
CCL20	NM_004591	-1.783319571	0.009264125
IDO1	NM_002164	-1.844872808	0.015106843
LEF1	NM_016269	-1.978667398	1.84E-05
MMP9	NM_004994	-2.062437596	0.000751947
FOS	NM_005252	-2.12843472	0.001100795
CXCL9	NM_002416	-2.18179791	0.00334655
CXCL11	NM_005409	-2.211024027	0.009815964
FGF13	NM_004114	-2.233289418	0.001666051
CXCL10	NM_001565	-2.239079297	0.000421839
CLIC2	NM_001289	-2.345714422	1.74E-06
CCL8	NM_005623	-2.605995483	1.96E-06
IL23A	NM_016584	-2.876867793	0.004167708

Table 69: Selected significantly differentially expressed genes between JRep and Norm fibroblasts in the stimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
CD86	NM_006889	1.691988908	1.28E-06
HIST1H1D	NM_005320	1.370161882	0.048339094
HIST1H2AG	NM_021064	0.909831741	0.038024279
XIST	NR_001564	0.781036032	0.003032042
XIST	NR_001564	0.77012738	0.002708592
CLEC3B	NM_003278	0.760223125	0.00864492
SERPINA3	NM_001085	0.736809069	0.002201534
XIST	NR_001564	0.728273701	0.006557277
FGA	NM_021871	0.711047069	2.03E-07
HIST1H2BF	NM_003522	0.69768458	0.037913163
CX3CR1	NM_001337	0.681183077	0.010977937
CD3G	NM_000073	0.633343327	5.10E-05
HIF3A	ENST00000457865	0.630607099	0.000137984
LCK	NM_005356	0.626589207	4.19E-05
XIST	NR_001564	0.612346772	0.002362684
HAPLN2	NM_021817	0.608447078	0.011728367
XIST	NR_001564	0.585688384	0.028960163
ADAM8	NM_001109	0.584218902	0.019632455
PTPRC	NM_002838	0.581900029	2.24E-06
XIST	NR_001564	0.572566779	0.00218547
ADAMTS7	NM_014272	0.571086422	0.00056656
FGA	NM_021871	0.570498787	0.029496042
CD69	ENST00000416624	0.552114648	3.96E-05
CD8B	NM_172102	0.546413516	6.80E-05
C4BPB	NM_000716	0.543892004	0.000337099
ZFP36L1	ENST00000408913	0.531648189	0.000500841
FGFR4	NM_213647	0.530368081	1.61E-05

TNFSF12	NM_003809	0.499036225	0.023950109
SERPINA6	NM_001756	0.493966888	0.003901558
IL2RG	NM_000206	0.472615111	6.74E-05
CCR2	NM_001123041	0.472546071	0.001457813
SERPINB5	NM_002639	0.461032925	0.000545111
SMAD5	NM_001001419	0.457758655	0.003535859
ZAP70	NM_001079	0.453024816	0.001362799
CD52	NM_001803	0.451231471	0.02980525
PTPRC	NM_002838	0.448590629	0.006506869
LCK	NM_005356	0.445543218	0.002427911
PTPRC	NM_002838	0.441392369	0.00531806
TFAP2A	ENST00000478375	0.438124734	7.13E-06
KLB	NM_175737	0.429161261	0.00034775
PDPR	NM_017990	0.420840653	0.00027076
SERPINA7	NM_000354	0.415856591	0.001059914
ADAM22	NM_021721	0.411186828	0.002002345
SERPINC1	NM_000488	0.405037595	0.008920011
HDAC10	NM_032019	0.403284537	0.00880644
CD5	NM_014207	0.40212043	0.010897038
PRODH2	NM_021232	0.397831018	0.001763908
CD69	NM_001781	0.395857927	0.011573273
CD247	NM_198053	0.389671489	0.003239389
IGSF1	NM_205833	0.388551629	0.020041898
CD33	NM_001772	0.383877957	0.019266639
CASP2	NM_032982	0.383059396	0.004015127
CD6	NM_006725	0.381438125	0.006879023
SYK	NM_003177	0.381433568	4.12E-05
ALDH1L1	NM_012190	0.380645722	0.00158906
JAG2	NM_002226	0.370181	0.000420857
IGSF1	NM_001555	0.365763079	0.004782584
SERPINA11	NM_001080451	0.364689486	0.000268538
IGSF1	NM_205833	0.36211567	0.001868555
COL11A2	NM_080680	0.360598939	0.009615331
COL2A1	NM_001844	0.359664813	0.002988953
COQ4	ENST00000372875	0.354794766	3.22E-05
COL11A2	NM_080680	0.354765832	0.007253614
MAPK4	NM_002747	0.354228566	0.016123701
CDK3	NM_001258	0.349222949	0.002864536
GZMM	NM_005317	0.341028948	0.018935952
SOX10	NM_006941	0.326774672	0.010697961
CARD14	NM_024110	0.325154201	0.000838469
CD8B	NM_004931	0.320845322	0.033084139
SOX13	ENST00000367203	0.313364698	0.001259891
SERPINF2	NM_000934	0.308958109	0.007937836
ADH6	NM_000672	0.285227063	0.007269505
SERPINA4	NM_006215	0.283669087	0.00277543
ICAM5	NM_003259	0.272813937	0.010977937
MYB	ENST00000528345	0.265743606	0.028484998
TRAF3IP3	NM_025228	0.252138632	0.028733741
ALOX5AP	NM_001629	0.2340983	0.024922621
IKBKB	ENST00000517917	0.232276725	0.046353465
HS3ST4	NM_006040	0.225007532	0.047861668
SOX21	NM_007084	0.22119704	0.036847802
ALOX5	NM_000698	0.220905386	0.018052175
SSX2	NM_175698	0.216502857	0.028484998
HSD17B7	ENST00000367915	0.20804293	0.039378397
FGF14	NM_175929	-0.257522526	0.029001674
HDAC2	NM_001527	-0.265998612	0.020405397
SDHB	NM_003000	-0.337718627	0.017233207
EGFR	NM_201282	-0.377178989	0.034524637
IRF2	NM_002199	-0.382769818	0.035031078
IKBKE	NM_014002	-0.415870733	0.002300704
COQ10B	NM_025147	-0.416384555	0.026970803
IRAK3	NM_007199	-0.512078496	0.02926486
IL4R	NM_001008699	-0.535459807	0.002098042
EGR4	NM_001965	-0.540265742	0.036183372
SOCS6	NM_004232	-0.576148706	0.003186845
IFNE	NM_176891	-0.60598381	0.041157673
RUNX1	NM_001122607	-0.638687262	0.038938588
COX1	ENST00000361624	-0.840009757	0.009253926
CDH11	ENST00000394156	-0.859638518	0.006426459

Table 70: Selected significantly differentially expressed genes between estRA and veRA fibroblasts in the stimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
CD86	NM_006889	1.389353652	0.000110696
SERPINA3	NM_001085	0.631049849	0.016910971
FGA	NM_021871	0.597545548	1.95E-05
FGFR4	NM_213647	0.526193187	4.56E-05
PTPRC	NM_002838	0.517090051	5.43E-05
HIF3A	ENST00000457865	0.514604096	0.003998147

LCK	NM_005356	0.489726667	0.00262486
ADAMTS7	NM_014272	0.478062428	0.008058912
CD3G	NM_000073	0.456491408	0.007315394
XIST	NR_001564	0.454882468	0.049413966
ZFP36L1	ENST00000408913	0.439942983	0.008279145
CD69	ENST00000416624	0.41418873	0.004147856
ZAP70	NM_001079	0.413148222	0.006511869
TFAP2A	ENST00000478375	0.402771447	5.91E-05
IGSF1	NM_001555	0.393891505	0.003564809
CD8B	NM_172102	0.381374764	0.012388163
CD6	NM_006725	0.360756503	0.017904322
CARD14	NM_024110	0.344221985	0.00071852
SERPINB5	NM_002639	0.340080032	0.024858009
GZMM	NM_005317	0.33880869	0.03003083
KLB	NM_175737	0.334828834	0.011502591
IL2RG	NM_000206	0.332472138	0.01141724
JAG2	NM_002226	0.326914767	0.003691402
COQ4	ENST00000372875	0.322244331	0.000274342
IGSF1	NM_205833	0.318299758	0.011937787
SERPINA11	NM_001080451	0.316737737	0.003055865
CDK3	NM_001258	0.303960575	0.017996744
ALDH1L1	NM_012190	0.2997419	0.027345873
ICAM5	NM_003259	0.292426984	0.009031449
SERPINA4	NM_006215	0.273377538	0.006720263
IKBKB	ENST00000517917	0.261226007	0.031079901
SYK	NM_003177	0.261037521	0.010663505
SOX13	ENST00000367203	0.259644339	0.015159833
IKBKE	NM_014002	-0.422234866	0.003158306
IL4R	NM_001008699	-0.59549829	0.000887368
COX1	ENST00000361624	-0.938249452	0.004908659
CDH11	ENST00000394156	-0.971410631	0.002778639

Table 71: Selected significantly differentially expressed genes between estRA and Res fibroblasts in the stimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
CD86	NM_006889	1.563459581	0.000123487
CDH13	NM_001257	1.537029867	0.037313916
HAS2	NM_005328	1.502620566	0.03357635
UBE2C	NM_181803	1.489955088	0.027924586
HMMR	NM_012484	1.467669381	0.024310377
HIST1H2AG	NM_021064	1.335573456	0.00466954
SERPINE1	NM_000602	1.16013258	0.028524381
HIST1H4C	NM_003542	1.125427058	0.016167988
HIST1H4L	NM_003546	1.108510183	0.026179302
TRAIP	NM_005879	1.085933188	0.003670476
SERPINA3	NM_001085	1.070070001	0.025490247
HIST1H4B	NM_003544	0.994390986	0.037047428
HMGB2	NM_002129	0.932994223	0.047323612
SERPINA3	NM_001085	0.927220414	0.000863092
HMGB3	NM_005342	0.911422971	0.001211103
CD200	NM_001004196	0.910262838	0.044512677
HIST1H2BF	NM_003522	0.859317407	0.018699219
HIST1H4L	NM_003546	0.751531177	0.038554682
COL6A1	NM_001848	0.709016063	0.007144767
SERPINA9	NM_175739	0.707830572	0.012668494
CD200	NM_001004196	0.68175105	0.017166993
HMGB3	NM_005342	0.671335463	0.022557566
XIST	NR_001564	0.658895783	0.03321227
CX3CR1	NM_001337	0.619154324	0.043721172
XIST	NR_001564	0.616633531	0.0411869
IL24	NM_001185156	0.614822074	0.012573438
MMP15	NM_002428	0.613820863	0.002400954
ADAM8	NM_001109	0.612417674	0.028729551
LCK	NM_005356	0.584613629	0.001166384
FGA	NM_021871	0.557814954	0.000371438
IGSF1	NM_001555	0.54567491	0.000187907
TNFSF12	NM_003809	0.544674139	0.026622437
CD3G	NM_000073	0.541644498	0.003224028
HIST1H2AC	NM_003512	0.52565108	0.019338646
HDAC10	NM_032019	0.509516579	0.003307606
ZAP70	NM_001079	0.493057055	0.002555697
XIST	NR_001564	0.484847915	0.026872866
CD8B	NM_172102	0.483123053	0.002774021
PTPRC	NM_002838	0.478132549	0.000950172
SOX10	NM_006941	0.454839192	0.001527243
FGFR4	NM_213647	0.447820798	0.001932167
SERPINA6	NM_001756	0.44604894	0.02532153
IL2RG	NM_000206	0.418665962	0.002666762
LCK	NM_005356	0.412883239	0.015733663
ALDH1L1	NM_012190	0.400957983	0.004119238
PRODH2	NM_021232	0.399089585	0.006910346

C4BPB	NM_000716	0.397645789	0.028242353
HYAL4	NM_012269	0.390610299	0.011115279
ILF3	NM_012218	0.390018509	0.028242353
SERPINF2	NM_000934	0.383232306	0.003575748
CCR2	NM_001123041	0.383188797	0.028856134
JAG2	NM_002226	0.378324734	0.002035801
COQ4	ENST00000372875	0.377325398	0.000139996
DUSP9	NM_001395	0.375589856	0.021701483
SOCS4	NM_199421	0.374660478	0.009046961
PDPR	NM_017990	0.370248054	0.00665152
IGSF1	NM_205833	0.365517775	0.006817367
CD8B	NM_004931	0.355195396	0.032318038
CD6	NM_006725	0.354842673	0.029099309
CDK3	NM_001258	0.348961865	0.010034215
SERPINB5	NM_002639	0.34170054	0.032111992
SERPINA11	NM_001080451	0.326307291	0.005795158
COL11A2	NM_080680	0.321306573	0.044851662
COX6B2	NM_144613	0.303749477	0.039307975
NKAPP1	NR_027131	0.298702221	0.049673575
HSD17B7	ENST00000367915	0.292236869	0.007183296
SYK	NM_003177	0.283521149	0.009882345
SERPINA4	NM_006215	0.268678191	0.01454975
TFAP2A	ENST00000478375	0.24187981	0.039975867
SMAD4	NM_005359	-0.297128508	0.04456685
CCL17	NM_002987	-0.359780165	0.003487231
RASA2	NM_006506	-0.376087841	0.037460651
IFNAR1	NM_000629	-0.39556539	0.045615513
HPGDS	NM_014485	-0.422237447	0.000206038
STAT5A	NM_003152	-0.43525795	0.022187366
EGFR	NM_201282	-0.438639945	0.025092737
JAK3	NM_000215	-0.443315153	0.003283529
IL34	NM_152456	-0.473032387	0.048959135
HSD17B7	NM_016371	-0.499442153	0.005742228
JMJD7	NM_001114632	-0.507511918	0.022187366
CD274	NM_014143	-0.512925039	0.037500662
SGPP2	NM_152386	-0.531868185	0.005667209
IL15RA	ENST00000379971	-0.542748294	0.003670476
CSF1	NM_000757	-0.544993197	0.022782128
DUSP8	NM_004420	-0.555988187	0.046325536
IL6ST	NM_001190981	-0.56338086	0.036277051
NFKBIE	NM_004556	-0.585423163	0.017604695
C1RL	NM_016546	-0.593295161	0.001288624
FGFR1	NM_001174066	-0.602380323	0.040371246
STAG2	NM_001042749	-0.603627691	0.037134847
HSD17B7	NM_016371	-0.616523976	0.001628698
CD274	NM_014143	-0.620277999	0.00390128
IFNAR2	NM_000874	-0.649999468	0.006516779
BCL6	NM_001130845	-0.650485044	0.045267101
TNF	NM_000594	-0.658195698	0.025809434
KMO	NM_003679	-0.668476509	5.53E-05
CSF3	NM_000759	-0.687700438	0.040647882
NFKB1	NM_003998	-0.689712459	0.000720464
CSF1	NM_172210	-0.704572097	0.013648841
IL6ST	NM_001190981	-0.720779249	0.007945021
FOXP2	NM_148900	-0.722520878	0.007112283
KGFLP1	NR_003674	-0.725823576	0.026622437
IL13RA1	NM_001560	-0.751814579	0.00455215
COQ10B	NM_025147	-0.760028117	0.000532896
COQ10B	NM_025147	-0.762325861	0.000160722
SERPING1	NM_000062	-0.763300222	0.041015229
IFNAR2	NM_207585	-0.77373329	4.87E-05
SPHK1	NM_182965	-0.804631376	0.01191349
MMP2	NM_004530	-0.817274774	0.005083732
IRAK3	NM_007199	-0.836042773	0.001091906
IL15RA	NM_172200	-0.871663288	0.003477945
CD44	NM_001202557	-0.873522492	0.010761402
CD55	NM_000574	-0.890591186	0.01162128
C1RL	NM_016546	-0.920338575	0.019751224
NOD2	NM_022162	-0.958713101	0.024983953
IL4R	NM_001008699	-0.966299923	1.53E-06
IL15RA	NM_172200	-1.026469726	0.040051804
DUSP8	NM_004420	-1.058484082	0.002349161
PPARG	NM_138711	-1.065016897	0.026532499
IL15RA	NM_001243539	-1.073430543	0.003669008
CD34	NM_001025109	-1.090273073	0.044202978
CX3CL1	NM_002996	-1.094933535	0.005795158
PPARG	NM_138711	-1.103628695	0.039339553
SERPING1	NM_000062	-1.106394233	0.004161134
KMO	NM_003679	-1.106728426	8.68E-06
COX1	ENST00000361624	-1.145758849	0.001603154
CDH11	ENST00000394156	-1.195238996	0.000863092
CTSK	NM_000396	-1.220489213	0.012093237

CCL3L3	NM_001001437	-1.276810215	0.010630695
LBP	NM_004139	-1.27718058	0.003539003
CCL4	NM_002984	-1.322453156	0.002380291
CD34	NM_001773	-1.452648314	0.036549931
PDGFRL	NM_006207	-1.479308971	0.03538884
CCL4	NM_002984	-1.500801871	0.001810964
CCL8	NM_005623	-1.568682821	0.005667209
LEF1	NM_016269	-1.572466135	0.000928608
IDO1	NM_002164	-1.573766315	0.028556762
CLIC2	NM_001289	-1.613207605	0.001364676
TNFSF15	NM_005118	-1.896569213	0.019547678
CCL20	NM_004591	-2.378568682	0.000181517

Table 72: Selected significantly differentially expressed genes between estRA and Norm fibroblasts in the stimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
HAS2	NM_005328	1.863641497	0.00204622
COL13A1	NM_080801	1.72684302	0.037177737
PTGS1	NM_000962	1.409304409	0.033029992
FOSB	NM_006732	1.343806018	0.036857004
UBE2C	NM_181803	1.335970453	0.026327227
HMMR	NM_012484	1.224741004	0.034055678
SERPINE1	NM_000602	1.056794368	0.024400147
TAGLN	NM_001001522	0.999164189	0.014197699
CD200	NM_001004196	0.956394177	0.015450002
TNFRSF10D	NM_003840	0.949449963	0.044283301
TMEM200A	NM_052913	0.944867292	0.026327227
HMGB3	NM_005342	0.880071025	0.000276546
FGF1	NM_000800	0.772345017	0.044679721
MAP4K4	NM_145686	0.763439786	0.009350926
TRAIP	NM_005879	0.714363828	0.034216507
HSD17B1	NM_000413	0.680544224	0.027388898
RUNX1	NM_001122607	0.667114716	0.026059734
TIMP3	NM_000362	0.664234108	0.030809454
HMGB3	NM_005342	0.634403347	0.014248049
SERPINA9	NM_175739	0.616161287	0.01433844
IL24	NM_001185156	0.539869397	0.013406659
ICAM3	NM_002162	0.446429413	0.007732348
NOSIP	NM_015953	0.408351845	0.002140324
MMP15	NM_002428	0.39081149	0.032790745
HIST4H4	NM_175054	0.37619479	0.034238824
SOCS4	NM_199421	0.356045743	0.004771034
TP53I13	NM_138349	0.343579953	0.03702182
ILF3	NM_012218	0.31470653	0.046986374
SERPINB10	NM_005024	0.277537304	0.038934749
HDAC2	NM_001527	0.26117545	0.03114751
CARD14	NM_024110	-0.212400571	0.038542368
SMAD4	NM_005359	-0.2963697	0.02308832
LTB	NM_002341	-0.313653098	0.022419105
JMJD4	NM_023007	-0.322082112	0.028357646
HPGDS	NM_014485	-0.323542722	0.001230834
JAK3	NM_000215	-0.326530066	0.015213878
CCL17	NM_002987	-0.35099288	0.000984883
LILRB4	NM_006847	-0.364234171	0.009443294
HSD17B7	NM_016371	-0.366669266	0.00446769
SGPP2	NM_152386	-0.367119936	0.033162086
STAG2	NM_001042751	-0.372625466	0.047188095
IL18BP	NM_173042	-0.378076569	0.022880559
ADAMTS7	NM_014272	-0.378895831	0.030722004
SERPING1	NM_000062	-0.387860616	0.005691281
C1R	NM_001733	-0.389349052	0.027265707
HLA-A	NM_001242758	-0.390188008	0.029792002
HLA-DOB	NM_002120	-0.390437831	0.011633555
TP53	NM_000546	-0.395200059	0.001895209
C1RL	NM_016546	-0.398276113	0.015377871
JHDM1D	NM_030647	-0.401245674	0.02632639
IL15RA	ENST00000379971	-0.40422517	0.015377871
SOX13	NM_005686	-0.405867686	0.031784568
ACADM	NM_000016	-0.410137428	0.020400179
CDK8	NM_001260	-0.416735566	0.012373031
ISLR	NM_005545	-0.428695732	0.007087651
IL4R	NM_001008699	-0.430840116	0.016195886
CD274	NM_014143	-0.431198986	0.048700588
HIF3A	ENST00000457865	-0.442057111	0.01100384
CSF1	NM_000757	-0.4445353	0.036588526
ADAMTSL4	NM_025008	-0.445928287	0.029007361
IFNAR2	NM_000874	-0.455881855	0.001030761
JMJD7	NM_001114632	-0.45635453	0.019892739
IFNGR2	NM_005534	-0.459795507	0.035908721
FCGR2A	NM_001136219	-0.464940197	0.042938344
KLB	NM_175737	-0.492938477	0.000832431

HSD17B7	NM_016371	-0.498084812	0.001464508
IL6ST	NM_001190981	-0.50219344	0.012311157
FGFR1	NM_023110	-0.510639562	0.003004326
CD79A	NM_001783	-0.515974885	0.000838652
IRF1	NM_002198	-0.518680395	0.044786146
HAPLN2	NM_021817	-0.524492457	0.031598294
KMO	NM_003679	-0.532783957	0.000220658
HLA-A	NM_002116	-0.533825288	0.020409831
HSD17B7	NM_016371	-0.536180663	0.001693074
IL10RB	NM_000628	-0.536729125	0.006109565
HLA-C	NM_001243042	-0.546105265	0.036639302
TNFAIP8	NM_014350	-0.551703298	0.029461256
HLA-F	NM_018950	-0.564036422	0.040773292
HLA-G	NM_002127	-0.572558138	0.015349289
CD274	NM_014143	-0.573820642	0.002226419
NFKB1	NM_003998	-0.582967088	0.00096499
IL15RA	NM_172200	-0.590941415	0.028357646
TGFA	NM_003236	-0.59400367	0.014784937
IL13RA1	NM_001560	-0.597456981	0.01115874
HLA-E	NM_005516	-0.61054245	0.009755457
IL6ST	NM_001190981	-0.61679541	0.010525634
TNXB	NM_032470	-0.622209606	0.031947739
IFNAR2	NM_000874	-0.628383005	0.002509507
KGFLP1	NR_003674	-0.628550613	0.030631377
IFNAR2	NM_207585	-0.636583968	0.00011241
TLR3	NM_003265	-0.662715456	0.046471537
IFI35	NM_005533	-0.669130085	0.034216507
TNXB	NM_019105	-0.672502287	0.04601849
STAT5A	NM_003152	-0.690858809	1.07E-05
CXCL10	NM_001565	-0.690933709	0.016992767
MMP2	NM_004530	-0.691262894	0.007456799
DUSP8	NM_004420	-0.724078434	0.002375228
CD47	NM_198793	-0.725866484	0.014405172
HIST1H4K	NM_003541	-0.760585567	0.040696845
CSF1	NM_172212	-0.777193589	0.014564613
PDPN	NM_198389	-0.800208524	0.017982015
NFKBIE	NM_004556	-0.804965437	0.000120562
IRF7	NM_004031	-0.812004224	0.022715878
CD55	NM_000574	-0.817286171	0.009022243
SERPING1	NM_000062	-0.825026652	0.011276122
IL15RA	NM_001243539	-0.83107606	0.011629609
C1RL	NM_016546	-0.849596212	0.014618917
CLEC3B	NM_003278	-0.851926692	0.002923129
APOL1	NM_145343	-0.881389901	0.006172614
TLR2	NM_003264	-0.894884242	0.018032703
CCL19	NM_006274	-0.908465024	0.034216507
HLA-B	NM_005514	-0.909949476	0.041417157
PPARG	NM_138711	-0.980284074	0.038060024
FGFR1	NM_001174066	-1.001676322	3.81E-05
SERPING1	NM_000062	-1.040922856	0.002011841
HIST1H4J	NM_021968	-1.044908184	0.036374972
KMO	NM_003679	-1.067345237	3.37E-07
PPARG	NM_138711	-1.106055931	0.008566536
LBP	NM_004139	-1.117578713	0.003857719
CLIC2	NM_001289	-1.119928538	0.012610989
CD38	NM_001775	-1.126813032	0.000618595
TNFSF13B	NM_006573	-1.148904249	0.041898051
DUSP8	NM_004420	-1.149465612	0.000123137
MMP9	NM_004994	-1.172035629	0.031620406
CCL4	NM_002984	-1.185064648	0.001898973
IL15RA	NM_172200	-1.222057249	0.004638356
CCL4	NM_002984	-1.24797393	0.003305644
IFIT2	NM_001547	-1.305962888	0.042315859
IDO1	NM_002164	-1.386307319	0.029811225
CX3CL1	NM_002996	-1.413203702	2.01E-05
CCL13	NM_005408	-1.456802722	0.045342654
CXCL10	NM_001565	-1.464218854	0.008618722
NOD2	NM_022162	-1.512045792	1.95E-05
CFB	NM_001710	-1.692376013	0.010839916
CCL20	NM_004591	-1.797212689	0.001304004
LEF1	NM_016269	-1.805530264	5.26E-06
IL23A	NM_016584	-1.95121438	0.024400147
TNFSF15	NM_005118	-1.989979448	0.004945469
CCL8	NM_005623	-2.013317151	2.08E-05

Table 73: Selected significantly differentially expressed genes between veRA and Norm fibroblasts in the stimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
HAS2	NM_005328	1.605294626	0.010409256
UBE2C	NM_181803	1.329194691	0.030058911
TAGLN	NM_001001522	1.049471578	0.011148016

FGFBP1	NM_005130	0.948710761	0.018982322
HIST1H4L	NM_003546	0.944511567	0.037092556
HIST1H4C	NM_003542	0.895178248	0.036546782
HMGB2	NM_002129	0.871309579	0.040442669
FGF1	NM_000800	0.868388854	0.025578643
TRAIP	NM_005879	0.779456651	0.02291392
HMGB3	NM_005342	0.769550947	0.00195089
TIMP3	NM_000362	0.653506391	0.037801077
TAGLN	NM_001001522	0.648770904	0.025891105
MAP4K4	NM_145686	0.613739929	0.045554189
HMGB3	NM_005342	0.546944889	0.041402942
HIST1H2AC	NM_003512	0.472057105	0.019800478
TP53I13	NM_138349	0.37672894	0.02436336
MMP15	NM_002428	0.372893355	0.048470785
TGFB1I1	NM_001042454	0.347825423	0.024291963
CARD14	NM_024110	-0.231468356	0.026509582
JAK3	NM_000215	-0.306742329	0.026801213
SERPING1	NM_000062	-0.32284551	0.026655279
HPGDS	NM_014485	-0.35819454	0.00032274
IL4R	NM_001008699	-0.370801633	0.047629563
FGFR1	NM_023110	-0.379921315	0.036076398
ISLR	NM_005545	-0.381155102	0.01991746
LILRB4	NM_006847	-0.397552063	0.004503745
JHDM1D	NM_030647	-0.401366921	0.028724831
CDK8	NM_001260	-0.402499248	0.018198402
IFNAR2	NM_000874	-0.414044206	0.003607002
JMJD7	NM_001114632	-0.415754388	0.039691302
KMO	NM_003679	-0.418442751	0.005233022
IL10RB	NM_000628	-0.425237954	0.037092556
CD79A	NM_001783	-0.427075748	0.007725523
CD274	NM_014143	-0.429089132	0.030190164
IL6ST	NM_002184	-0.432606192	0.03039811
HSD17B7	NM_016371	-0.433246887	0.007341888
C1RL	NM_016546	-0.448322734	0.006795621
OSMR	NM_001168355	-0.448883531	0.023110903
HLA-DOB	NM_002120	-0.452201724	0.003290349
OSMR	NM_003999	-0.457424277	0.04197604
HSD17B7	NM_016371	-0.458236226	0.009710607
IL6ST	NM_001190981	-0.485227285	0.04768393
STAT5A	NM_003152	-0.487121897	0.003315244
SGPP2	NM_152386	-0.488554967	0.003945037
IL6ST	NM_001190981	-0.506984705	0.013105508
ADAMTSL4	NM_025008	-0.511974847	0.012590337
TNFAIP8	NM_014350	-0.512318351	0.049841423
IFNAR2	NM_000874	-0.538931074	0.012608272
IFNAR2	NM_207585	-0.544091207	0.001239674
FOXP2	NM_148900	-0.556215769	0.023170565
ZFP36L1	NM_004926	-0.583732222	0.028681134
FGFR1	NM_001174066	-0.592941746	0.02436336
NFKB1	NM_003998	-0.601895751	0.000784427
KGFLP1	NR_003674	-0.612671056	0.039559794
IL6ST	NM_001190981	-0.629640607	0.010066901
IL13RA1	NM_001560	-0.645809245	0.00643236
DUSP8	NM_004420	-0.64999648	0.008293007
CXCL10	NM_001565	-0.658034135	0.02696493
MMP2	NM_004530	-0.666500731	0.011555876
NFKBIE	NM_004556	-0.669463144	0.001878585
SERPING1	NM_000062	-0.683084391	0.044154183
KMO	NM_003679	-0.698157236	0.001353739
CD44	NM_001202557	-0.702391103	0.02436336
CSF3	NM_000759	-0.703076504	0.018791106
SERPING1	NM_000062	-0.78941791	0.025578643
CTSF	NM_003793	-0.894206088	0.015383848
IL15RA	NM_172200	-0.928109147	0.040152718
C1RL	NM_016546	-0.94549329	0.006932611
PPARG	NM_138711	-0.962782133	0.04838014
CTSK	NM_000396	-0.997434414	0.02436336
NOD2	NM_022162	-1.004591972	0.00800878
CCL3L3	NM_001001437	-1.007222885	0.027059477
PPARG	NM_138711	-1.013718064	0.018484087
DUSP8	NM_004420	-1.031780566	0.000754751
LBP	NM_004139	-1.032863852	0.008988852
CX3CL1	NM_002996	-1.051437269	0.002673
CLIC2	NM_001289	-1.066019733	0.020485122
TLR2	NM_003264	-1.098272277	0.003323617
CCL4	NM_002984	-1.123349325	0.003708453
CXCL10	NM_001565	-1.157012988	0.046904497
CCL4	NM_002984	-1.226405852	0.004154838
MMP9	NM_004994	-1.384140623	0.011604694
CFB	NM_001710	-1.438140773	0.036622789
IDO1	NM_002164	-1.471665699	0.022970678
TNFSF15	NM_005118	-1.669297347	0.023049545
LEF1	NM_016269	-1.720464431	1.68E-05
CCL20	NM_004591	-1.904819383	0.000739246
CCL8	NM_005623	-2.278599649	8.38E-07

Table 74: Selected significantly differentially expressed genes between Res and Norm fibroblasts in the stimulated high serum

treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

8.7 Differentially expressed gene lists from the unstimulated high samples in the microarray experiments

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
MOK	NM_014226	1.212031885	0.031993944
MOK	NM_014226	0.985073433	0.006928529
WNT16	NM_057168	0.911042236	0.02143313
COX1	ENST00000361624	0.8426846	0.037000316
WISP1	NM_080838	0.672212179	0.02570778
IL4R	NM_001008699	0.485301642	0.028275438
LTK	NM_002344	0.42508451	0.021139755
IRF2BP1	NM_015649	0.419593023	0.019342288
CDK10	NM_052987	0.381575367	0.043636672
IL7	ENST00000379113	0.366623445	0.005031484
WDR5B	NM_019069	0.365800171	0.020916839
CASP8	NM_033358	0.256434879	0.039460661
COQ4	ENST00000372875	-0.243661756	0.027087867
HSD17B7	ENST00000367915	-0.276134198	0.01955496
SYK	NM_003177	-0.277499833	0.02010976
HSF1	NM_005526	-0.278693814	0.047857472
SERPINA4	NM_006215	-0.30019578	0.010603827
KLB	NM_175737	-0.316874141	0.044758102
BMS1P1	NR_026566	-0.321611718	0.048162082
TRAF3IP3	NM_025228	-0.323060461	0.017750981
BMP7	NM_001719	-0.336701939	0.040000247
H3F3B	NM_005324	-0.342043525	0.043980233
CR2	NM_001006658	-0.351364089	0.039930204
HYAL4	NM_012269	-0.356839297	0.035572628
FGFR3	NM_000142	-0.358480697	0.015974075
APOA1	NM_000039	-0.358632557	0.015903845
PTPRC	NM_002838	-0.367907797	0.019781181
CELSR1	NM_014246	-0.370813748	0.026384121
ITGAL	NM_002209	-0.37376675	0.027605632
RUNX3	NM_001031680	-0.373776493	0.043132022
SERPINB5	NM_002639	-0.384661797	0.024943244
BMP7	NM_001719	-0.411984721	0.000480334
SERPINA7	NM_000354	-0.412257219	0.009125375
ZAP70	NM_001079	-0.413709142	0.020916839
CD8B	NM_172102	-0.420678225	0.016852593
APOA1	NM_000039	-0.422214291	0.020657089
DUSP9	NM_001395	-0.425556829	0.01496303
IGSF1	NM_205833	-0.425793227	0.002585343
SERPINA11	NM_001080451	-0.426722593	0.000484185
LCK	NM_005356	-0.432791015	0.019342288
FGFR4	NM_213647	-0.463874487	0.002289808
SOX10	NM_006941	-0.470065562	0.001715966
SDHC	NM_003001	-0.482786437	0.000911182
ZFP36L1	ENST00000408913	-0.499919821	0.017554638
FGL1	NM_201553	-0.501189627	0.009974186
IGSF1	NM_001555	-0.525041012	0.000669135
SERPINA6	NM_001756	-0.54832397	0.008948615
CD3G	NM_000073	-0.57773592	0.002803615
MMP15	NM_002428	-0.584013325	0.007155178
FGA	NM_021871	-0.599219859	0.000313193
LCK	NM_005356	-0.615547957	0.000920668
EGR4	NM_001965	-0.621860784	0.045670072
HYAL1	NM_007312	-0.621986201	0.034601897
SOX2	NM_003106	-0.62783413	0.003399286
CD52	NM_001803	-0.637320722	0.008341503
XIST	NR_001564	-0.646141136	0.048218978
XIST	NR_001564	-0.723883296	0.030889521
XIST	NR_001564	-0.75393362	0.019516672
RUNX1	NM_001001890	-0.754022888	0.007868499
CELSR1	NM_014246	-1.055475544	0.009042443
FGB	NM_005141	-1.15772556	0.000827724
HIST2H2AA4	NM_001040874	-1.182713302	0.020594091
EGR1	NM_001964	-1.269785502	0.032739875
CD86	NM_006889	-1.514678902	0.000391181
FOS	NM_005252	-1.574686703	0.024943244
FOSB	NM_006732	-1.997930043	0.008839355
RASIP1	NM_017805	-2.241809536	0.010603827

Table 75: Selected significantly differentially expressed genes between JRep and estRA fibroblasts in the unstimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
INHBB	NM_002193	1.716276642	0.018436135
HAS3	NM_005329	1.422988857	0.000337483
ITGA8	NM_003638	1.269949514	0.038020853

MOK	NM_014226	1.126628239	0.041250473
WNT16	NM_057168	0.789952652	0.044532502
TNFRSF19	NM_018647	0.718280612	0.015976007
CLEC3B	NM_003278	0.703461198	0.047521126
BCOR	ENST00000501455	0.417180748	0.012991629
MAPK8	ENST00000374189	0.396928964	0.011364137
CD69	ENST00000416624	0.338302522	0.047521126
CIDEC	NM_022094	0.334484005	0.045801995
TP53	NM_000546	0.32382432	0.038020853
CD96	NM_198196	0.317535159	0.044098407
PTPRC	NM_002838	0.315997377	0.04370886
KLRK1	NM_007360	0.301313506	0.048538186
FGFR1	ENST00000496296	0.274771438	0.032238501
IGSF11	NM_152538	0.256345709	0.022551316
SS18L2	NM_016305	-0.28737696	0.014243734
ISLR2	NM_020851	-0.365256552	0.038477353
RB1	NM_000321	-0.433077435	0.038020853
IGLL5	NM_001178126	-0.447365215	0.019603004
ATF3	NM_001674	-0.909623962	0.048538186
ATF3	NM_001040619	-1.007989254	0.044801141
IER3	NM_003897	-1.268113304	0.033469761
IER3	NM_003897	-1.387274556	0.045364062
MME	NM_007289	-1.456482769	0.043554892
EGR1	NM_001964	-1.531252764	0.004543815
ATF3	NM_001040619	-1.595787866	0.025567647
FOS	NM_005252	-1.616867706	0.014786755
MME	NM_007289	-1.646768285	0.044801141
EGR2	NM_000399	-1.868377034	0.028098414
FOSB	NM_006732	-2.17854366	0.001754614

Table 76: Selected significantly differentially expressed genes between JRep and veRA fibroblasts in the unstimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
MOK	NM_014226	1.341542962	0.016654435
HAS3	NM_005329	1.233692671	0.003190146
MOK	NM_014226	0.909270069	0.01360078
NFATC4	NM_001136022	0.822747846	0.032651921
PADI2	NM_007365	0.587461014	0.002648371

Table 77: Selected significantly differentially expressed genes between JRep and Res fibroblasts in the unstimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
KRT34	NM_021013	2.561281789	0.010249823
KRT19	NM_002276	2.426719483	5.11E-05
KRT16	NM_005557	2.216353535	1.75E-06
KRT19P2	NR_036685	2.109597555	0.001114972
KRT16P2	NR_029392	2.089170757	0.000146855
COL8A1	NM_001850	2.058805438	0.018209918
KRT14	NM_000526	1.854161369	0.000995028
CDH13	NM_001257	1.798675797	0.013296404
MOK	NM_014226	1.730986201	0.00051115
SERPINE1	NM_000602	1.537731326	0.002725874
HAS2	NM_005328	1.446352169	0.046667314
KRT14	NM_000526	1.436596428	0.033976333
KRT14	NM_000526	1.397571154	0.007636063
TNC	NM_002160	1.3070678	0.046665367
CDH13	NM_001220491	1.300711577	0.011174189
LOXL2	NM_002318	1.263120376	0.022905693
MOK	NM_014226	1.254579338	0.000118619
HAS3	NM_005329	1.240530677	0.00246382
ITGA3	NM_002204	0.843342081	0.03913778
WISP1	NM_080838	0.819108665	0.002421392
NFATC4	NM_001136022	0.787615257	0.021244107
FGF5	NM_004464	0.772090975	0.036210115
COL22A1	NM_152888	0.696580929	0.048976219
IL27RA	NM_004843	0.67504455	0.039547336
LDHA	NM_005566	0.656215996	0.048676461
PADI2	NM_007365	0.611987758	0.001349021
TGFB1	NM_000660	0.607095008	0.035482378
FGF5	NM_033143	0.54449707	0.005785335
ARG2	NM_001172	0.533966036	0.026793723
CD3EAP	NM_012099	0.532715807	0.021423199
MOK	ENST00000520252	0.523323529	0.005562866
ITGA5	NM_002205	0.48814006	0.043635862
MLF2	NM_005439	0.432042084	0.007668477
STK24	NM_001032296	0.42315017	0.0022017
STK24	NM_001032296	0.416977939	0.004275039
CSK	NM_004383	0.409472398	0.046898056
ILF3	NM_004516	0.396858182	0.033289871
WDR53	NM_182627	-0.244446295	0.037270247
CCR3	NM_001837	-0.247132291	0.048522943
SS18L2	NM_016305	-0.323191666	0.004107214
HSD17B7	NM_016371	-0.337133292	0.024193337

COQ10A	NM_144576	-0.33880253	0.047030593
COL4A3BP	NM_001130105	-0.372614769	0.044675826
KLRG1	NM_005810	-0.420061607	0.010963434
VEGFB	NM_003377	-0.470658463	0.009932464
IFNAR1	NM_000629	-0.506308795	0.008845084
CTSA	NM_000308	-0.511282007	0.049396774
SOX13	NM_005686	-0.517805158	0.014265229
IRAK3	NM_007199	-0.548544738	0.039428232
HSD17B7	NM_016371	-0.558702561	0.004107214
HSD17B7	NM_016371	-0.564816404	0.001355841
DUSP8	NM_004420	-0.565025361	0.048976219
ZFP36L1	ENST00000408913	-0.582676302	0.0022017
HYAL1	NM_007312	-0.587148161	0.033289871
ZFP36L1	NM_004926	-0.60382969	0.048137308
MBP	NM_001025101	-0.651595952	0.033581214
ITM2B	NM_021999	-0.666292511	0.039340596
B2M	NM_004048	-0.683491573	0.041804792
HABP4	NM_014282	-0.711923103	0.02637757
MMP2	NM_004530	-0.743333383	0.010676081
ITM2B	NM_021999	-0.763153632	0.005639895
C1QTNF4	NM_031909	-0.801815856	0.010561496
IL17RD	NM_017563	-0.83844743	0.023313002
S100B	NM_006272	-0.968253225	0.049937442
CARD6	NM_032587	-0.998731823	0.020548565
LY96	NM_015364	-1.022549493	0.002055855
IL1R1	NM_000877	-1.032728285	0.046545608
ICAM1	NM_000201	-1.046967588	0.014770713
SOX8	NM_014587	-1.137090382	0.004505171
TLR5	NM_003268	-1.181659031	0.001372029
EGR1	NM_001964	-1.307340463	0.016820741
ITGB3	NM_000212	-1.345333403	0.013658454
TNFAIP2	NM_006291	-1.39163094	0.005785335
MME	NM_007289	-1.483103059	0.031466982
PDGFRL	NM_006207	-1.545078197	0.029744914
FOS	NM_005252	-1.582120022	0.014265229
FIGF	NM_004469	-1.59026749	0.009119436
CXCL14	NM_004887	-1.677358603	0.039325923
CD34	NM_001025109	-1.684696548	0.00118431
HS3ST1	NM_005114	-1.717329908	0.009650193
CTSK	NM_000396	-1.949735742	1.54E-05
CD34	NM_001773	-2.127962454	0.001382946
CTSH	NM_004390	-2.37244439	0.01906409

Table 78: Selected significantly differentially expressed genes between JRep and Norm fibroblasts in the unstimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
INHBB	NM_002193	1.487890678	0.017010949
CD86	NM_006889	1.399967653	4.64E-05
FGB	NM_005141	1.393160667	6.83E-07
XIST	NR_001564	0.786245615	0.001816541
XIST	NR_001564	0.77748283	0.002732686
XIST	NR_001564	0.770299567	0.003170536
XIST	NR_001564	0.748135407	0.002699061
CLEC3B	NM_003278	0.745513545	0.009243057
CELSR1	NM_014246	0.719555642	0.038962113
FGA	NM_021871	0.71561975	8.62E-08
CX3CR1	NM_001337	0.688899786	0.008849375
PTPRC	NM_002838	0.683905174	2.57E-08
CD3G	NM_000073	0.671539668	1.35E-05
SOX2	NM_003106	0.656740042	0.000124041
XIST	NR_001564	0.653625874	0.007015909
XIST	NR_001564	0.652306468	0.000889356
ITM2A	NM_004867	0.648737371	0.020853834
LCK	NM_005356	0.609034082	5.27E-05
C4BPB	NM_000716	0.593115844	7.07E-05
FYB	NM_001465	0.588154605	0.001566101
IL27RA	NM_004843	0.58547962	0.044760618
BMS1P1	NR_026566	0.5835468	1.65E-06
FGA	NM_021871	0.560545646	0.031003743
CD69	ENST00000416624	0.557874805	2.41E-05
SERPINA3	NM_001085	0.554754005	0.027784766
FA2H	NM_024306	0.554550586	0.000454189
PDGFA	NM_033023	0.547500634	0.010821429
BMP7	NM_001719	0.532902712	2.89E-08
CD8B	NM_172102	0.530294641	9.64E-05
XLOC_l2_014830	THC2583825	0.529957155	0.000422228
LCK	NM_005356	0.529287162	0.000203067
HAPLN2	NM_021817	0.524093117	0.034541478
SERPINB5	NM_002639	0.521226941	6.91E-05
FCGR2A	NM_001136219	0.519676422	0.023175855
ZAP70	NM_001079	0.515066922	0.000185218
DNMT3A	NM_175629	0.514240974	0.011570765
SERPINA7	NM_000354	0.511234195	3.82E-05
COL4A6	NM_033641	0.511214206	0.001555705
SERPINA6	NM_001756	0.509143299	0.002388778

PTPRC	NM_002838	0.508447063	0.001378175
CD247	NM_198053	0.507630009	6.71E-05
TNFSF12	NM_003809	0.504761489	0.019922519
CD69	NM_001781	0.49764473	0.000799468
SERPINA11	NM_001080451	0.496660909	5.48E-07
KLRK1	NM_007360	0.495966195	2.84E-05
ADAMTS7	NM_014272	0.490866705	0.003272827
MAPK4	NM_002747	0.487064861	0.000372136
CDKN1B	NM_004064	0.483293189	0.048982497
XIST	NR_001564	0.480774477	0.011321831
CD52	NM_001803	0.480739392	0.016853212
KLB	NM_175737	0.480370185	4.83E-05
STAG3	NM_012447	0.470596179	0.035555958
PTPRC	NM_002838	0.468056569	0.002447466
BCOR	ENST00000501455	0.464603192	0.000759932
CD33	NM_001772	0.461611704	0.002828963
FGL1	NM_201553	0.457518367	0.00329019
CCR9	NM_031200	0.456181617	0.002439727
IGSF1	NM_205833	0.454635989	5.55E-05
C3AR1	NM_004054	0.453154221	0.009859316
ADH4	NM_000670	0.440453414	0.013515688
CCR2	NM_001123041	0.43916982	0.002969279
LPL	NM_000237	0.435352199	0.004814471
LTB	NM_002341	0.428748989	0.000883674
CDK1	NM_001170406	0.425444799	0.000289346
DUSP13	NM_001007271	0.422373316	0.000411105
TFAP2A	ENST00000478375	0.421369394	1.31E-05
COL2A1	NM_001844	0.420327497	0.000323821
SERPINA5	NM_000624	0.418819585	0.038864277
IL10RA	NM_001558	0.418405975	0.022221312
CD5	NM_014207	0.418093652	0.006670375
SOX10	NM_006941	0.411445868	0.000681585
TRAF3IP3	ENST00000476050	0.411277535	0.003470093
SERPINC1	NM_000488	0.410952836	0.006862587
HIF3A	ENST00000457865	0.409529776	0.019350732
TRAF3IP3	NM_025228	0.402629649	0.011062
IL2RG	NM_000206	0.399643131	0.000715164
FGG	NM_000509	0.396850412	0.022027791
CD6	NM_006725	0.388716014	0.004983758
SYK	NM_003177	0.388574444	2.18E-05
CD46	NM_172350	0.388009659	0.015497876
IGSF1	NM_001555	0.386699644	0.002262719
CD28	NM_006139	0.385215027	0.001174158
CR2	NM_001006658	0.384497989	0.003558295
ITGAL	NM_002209	0.382221166	0.004044319
C4A	NM_001252204	0.381475529	0.01850763
BMS1P1	NR_026566	0.377859434	0.005031875
COL11A2	NM_080680	0.375163747	0.003549019
CDH3	NM_001793	0.372451994	0.022017955
C1R	NM_001733	0.366792502	0.042618033
ZFP36L1	ENST00000408913	0.360853985	0.026408269
APOB	NM_000384	0.354652114	0.002559743
PDPR	NM_017990	0.353616947	0.002340278
SOX3	NM_005634	0.351153373	0.012763739
RUNX3	NM_001031680	0.345306885	0.001848102
S100A1	NM_006271	0.345007172	0.006533316
IGSF11	NM_152538	0.340811035	0.000140291
CSF3R	NM_156039	0.339299795	0.024377944
ALDH1L1	NM_012190	0.334933896	0.005744704
FGA	NM_000508	0.331631604	0.044244797
RUNX3	NM_001031680	0.328854301	0.028586782
CDH1	NM_004360	0.322058218	0.04349493
C8A	NM_000562	0.319942664	0.025500708
COL11A2	NM_080680	0.317723291	0.024969792
MYB	ENST00000528345	0.314279739	0.006031665
HSD17B2	NM_002153	0.311560105	0.002323268
CD1E	NM_001042583	0.307505764	0.001174158
CD8B	NM_004931	0.307153303	0.041292581
FGFR3	NM_000142	0.302532134	0.012326744
S100Z	NM_130772	0.299431446	0.010620545
ICAM5	NM_003259	0.299148733	0.00390435
SERPINA4	NM_006215	0.299104297	0.001250447
BMP7	NM_001719	0.29842664	0.024893446
FGFR4	NM_213647	0.296876698	0.023451648
FBN3	NM_032447	0.29424809	0.008582602
SOX13	ENST00000367203	0.29285275	0.00239905
CD28	NM_006139	0.275956274	0.037291831
TRAF3IP3	NM_025228	0.273708279	0.013549551
TNFRSF19	ENST00000464735	0.270651264	0.035745641
ALOX5	NM_000698	0.260773496	0.003263474
CDK3	NM_001258	0.260501408	0.035216964
XIST	NR_001564	0.250684006	0.030179521
TNFRSF13C	NM_052945	0.245726179	0.040245869
CTSG	NM_001911	0.241970711	0.031175055
CD3E	NM_000733	0.237333058	0.040838797
COQ4	ENST00000372875	0.232092671	0.008105473

PDPR	NM_017990	0.231353441	0.036017693
COL23A1	NM_173465	0.226253206	0.042838597
HSD17B7	ENST00000367915	0.224734196	0.021033651
CASP8	NM_033358	-0.218036256	0.033176237
TRAF6	NM_145803	-0.275697768	0.026180046
DUSP11	NM_003584	-0.299752221	0.037364778
HSF1	NM_005526	-0.314683816	0.041553628
MAPKAP1	NM_001006617	-0.37653558	0.034492646
LTK	NM_002344	-0.381200477	0.010249542
ADAMTSL1	NM_001040272	-0.519168044	0.019555284
IRAK3	NM_007199	-0.553194337	0.014530198
COX1	ENST00000361624	-0.72731242	0.02753743

Table 79: Selected significantly differentially expressed genes between estRA and veRA fibroblasts in the unstimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
RASIP1	NM_017805	1.63676332	0.024675231
INHBB	NM_002193	1.354470241	0.03154465
KRT14	NM_000526	1.230862323	0.037520334
FGB	NM_005141	1.18335489	1.89E-05
CD86	NM_006889	1.141605364	0.000898732
XIST	NR_001564	0.805622156	0.00118974
FN1	NM_054034	0.784469446	0.00316798
CELSR1	NM_014246	0.781936624	0.019506106
FGA	NM_021871	0.765784068	1.85E-08
ITM2A	NM_004867	0.749535586	0.005252757
XIST	NR_001564	0.733949954	0.004914706
XIST	NR_001564	0.721719707	0.005706403
PTPRC	NM_002838	0.646477237	8.06E-08
LCK	NM_005356	0.627970825	2.62E-05
XIST	NR_001564	0.61383618	0.017022459
CD3G	NM_000073	0.613830605	5.64E-05
PDGFA	NM_033023	0.602181443	0.003891648
SOX2	NM_003106	0.58600697	0.000604195
XIST	NR_001564	0.578225701	0.018825563
CX3CR1	NM_001337	0.574301294	0.034397621
XIST	NR_001564	0.571939394	0.003947076
SERPINA6	NM_001756	0.571739347	0.00048264
C4BPB	NM_000716	0.535601046	0.000314538
IGSF1	NM_205833	0.52959373	2.39E-06
ADH4	NM_000670	0.517363702	0.002455152
ZAP70	NM_001079	0.508818681	0.000195445
FYB	NM_001465	0.508616455	0.007195476
SERPINA3	NM_001085	0.508066413	0.045429445
XLOC_I2_014830	THC2583825	0.506721699	0.000702904
CD52	NM_001803	0.502621977	0.010523776
TNFSF12	NM_003809	0.492037703	0.022219393
CD69	ENST00000416624	0.484762637	0.000218784
BMS1P1	NR_026566	0.482705015	5.89E-05
BMP7	NM_001719	0.472895938	5.70E-07
IGSF1	NM_001555	0.46878698	0.000141362
FCGR2A	NM_001136219	0.465716239	0.04449413
LCK	NM_005356	0.463349217	0.001132027
FGFR4	NM_213647	0.462659911	0.000127518
MAPK4	NM_002747	0.457469943	0.000795667
FA2H	NM_024306	0.456406316	0.004677174
S100A14	NM_020672	0.456203175	0.016125619
SERPINA11	NM_001080451	0.455521078	3.45E-06
CD8B	NM_172102	0.451734574	0.000872443
ADAMTS7	NM_014272	0.446566361	0.0079229
KLRK1	NM_007360	0.44602982	0.000148791
CD69	NM_001781	0.44464073	0.002891015
SERPIND1	NM_000185	0.442753979	0.03542021
FGL1	NM_201553	0.441882931	0.004432985
CR2	NM_001006658	0.437560259	0.000665857
COL2A1	NM_001844	0.431041181	0.000201329
LPL	NM_000237	0.428776745	0.005200617
SERPINB5	NM_002639	0.423787828	0.001288637
PTPRC	NM_002838	0.422934756	0.009292428
DNMT3A	NM_175629	0.42009638	0.047229901
CD5	NM_014207	0.417957782	0.006193604
C3AR1	NM_004054	0.415426391	0.018825563
CD247	NM_198053	0.411059482	0.001331955
COL4A6	NM_033641	0.406431707	0.015132598
SOX10	NM_006941	0.399467897	0.000904264
KLB	NM_175737	0.397448374	0.00076284
COL11A2	NM_080680	0.385876183	0.00232196
SERPINC1	NM_000488	0.383058302	0.012252796
SOX3	NM_005634	0.379859297	0.005707849
SERPINA7	NM_000354	0.377993935	0.002579068
PTPRC	NM_002838	0.377614764	0.018166892
IL2RG	NM_000206	0.375653632	0.001443823
TFAP2A	ENST00000478375	0.367821847	0.000123151
ITGAL	NM_002209	0.366309406	0.005742098
IGSF1	NM_205833	0.364696961	0.026626925

CDK1	NM_001170406	0.361228581	0.00228372
CCR9	NM_031200	0.358733421	0.022120312
HIF3A	ENST00000457865	0.357098811	0.045720892
BMP7	NM_001719	0.349870323	0.005707849
ZFP36L1	ENST00000408913	0.348702533	0.030910349
C4A	NM_001252204	0.346805404	0.033850497
CCR2	NM_001123041	0.34524836	0.025138173
S100A1	NM_006271	0.344668186	0.006111424
BCOR	ENST00000501455	0.343136897	0.017398604
TRAF3IP3	NM_025228	0.341218984	0.036309902
TRAF3IP3	NM_025228	0.338936149	0.001312646
SYK	NM_003177	0.338060062	0.000200579
COL11A2	NM_080680	0.335356002	0.015106922
TRAF3IP3	ENST00000476050	0.335323024	0.02101509
FGA	NM_000508	0.332787573	0.03915693
ICAM5	NM_003259	0.331074843	0.001034185
CD8B	NM_004931	0.328673471	0.023618662
DUSP9	NM_001395	0.326001223	0.023419725
CSF3R	NM_156039	0.322677537	0.031873117
CD28	NM_006139	0.315811031	0.00945551
ALDH1L1	NM_012190	0.312869907	0.010157806
RUNX3	NM_001031680	0.306010969	0.006398498
CDK3	NM_001258	0.302779	0.009869905
HSD17B2	NM_002153	0.30210379	0.002987411
APOB	NM_000384	0.301171995	0.012409249
SERPINA4	NM_006215	0.300911999	0.00101792
LTB	NM_002341	0.296550345	0.030935703
BMS1P1	NR_026566	0.295646399	0.035851997
PDPR	NM_017990	0.293820483	0.013797241
IL4	NM_000589	0.292525187	0.03542021
CD6	NM_006725	0.290015633	0.047897935
CELSR1	NM_014246	0.289972292	0.035185641
S100Z	NM_130772	0.282174472	0.01642312
CD3E	NM_000733	0.272426357	0.013443939
SERPINF2	NM_000934	0.269176481	0.0214147
DUSP13	NM_001007271	0.266818024	0.038529898
CD28	NM_006139	0.26031926	0.048705249
TNFRSF19	ENST00000464735	0.260260092	0.042062602
TNFRSF13C	NM_052945	0.257235816	0.026935309
FGFR3	NM_000142	0.256764268	0.03890405
FBN3	NM_032447	0.255438846	0.025503828
SOX13	ENST00000367203	0.253979704	0.009772756
APOA1	NM_000039	0.247464653	0.048757834
CD4	NM_000616	0.245106255	0.036309902
COL23A1	NM_173465	0.237349782	0.02855393
HS3ST4	NM_006040	0.230952134	0.034603836
IGSF11	NM_152538	0.224173136	0.017432898
ALOX5	NM_000698	0.21634879	0.017590967
IGSF3	NM_001542	0.199607414	0.03242972
CASP8	NM_033358	-0.216537723	0.031653926
COQ5	NM_032314	-0.254419817	0.032623367
COQ3	NM_017421	-0.269364163	0.049220135
DUSP11	NM_003584	-0.286353748	0.045411247
WDR5B	NM_019069	-0.323685978	0.010424637
CDK10	NM_052987	-0.360905187	0.01534259
LTK	NM_002344	-0.364567701	0.013957759
IL4R	NM_001008699	-0.385069621	0.033929214
IRAK3	NM_007199	-0.553635952	0.013493758
COX1	ENST00000361624	-0.756005691	0.018763722
CDH11	ENST00000394156	-0.876306051	0.004267432
IL1RL1	NM_016232	-1.01924266	0.027421888

Table 80: Selected significantly differentially expressed genes between estRA and Res fibroblasts in the unstimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
KRT14	NM_000526	2.18782001	0.000466137
KRT19	NM_002276	2.048846792	0.00070572
KRT19P2	NR_036685	1.824335547	0.003827506
COL8A1	NM_001850	1.781949078	0.031519895
CDH13	NM_001257	1.774000812	0.010473838
ITGB2	NM_000211	1.698896857	0.036273983
KRT16P2	NR_029392	1.65208547	0.002598368
HAS2	NM_005328	1.506856423	0.024315716
KRT14	NM_000526	1.381380857	0.012340073
CLEC11A	NM_002975	1.358567517	0.0012199
CD86	NM_006889	1.317134679	0.00078387
KRT16	NM_005557	1.267198321	0.005782468
KRT14	NM_000526	1.193109911	0.018527012
HIST1H3F	NM_021018	1.158694665	0.027388927
HIST1H3D	NM_003530	1.153052852	0.025895415
CTSZ	NM_001336	1.126481426	0.040047523
HIST1H3H	NM_003536	1.097987097	0.022066074
FGB	NM_005141	1.059320242	0.000871688
ITGA3	NM_002204	0.99582735	0.008098711
HIST1H2AG	NM_021064	0.93200215	0.042261826

CELSR1	NM_014246	0.923017787	0.010915795
TGFB1	NM_000660	0.910093471	0.00061894
HIST1H4C	NM_003542	0.908960938	0.044297156
VEGFC	NM_005429	0.879230635	0.036198755
IL27RA	NM_004843	0.7341059	0.015004893
CLEC18B	NM_001011880	0.730892279	0.003354964
HIST1H3G	NM_003534	0.726346066	0.024756061
HIST1H2BF	NM_003522	0.696416323	0.048574629
HIST2H2BF	NM_001161334	0.678433808	0.030894638
FGA	NM_021871	0.651168156	2.43E-05
ITM2A	NM_004867	0.649147672	0.031749896
SERPINA3	NM_001085	0.642135485	0.016156199
XIST	NR_001564	0.621802104	0.005141404
XIST	NR_001564	0.613145033	0.03753264
RUNX1	NM_001001890	0.607721775	0.017029413
PLOD1	NM_000302	0.6062584	0.028182629
SOX2	NM_003106	0.605706387	0.002002709
PDGFA	NM_002607	0.596336406	0.031034245
ARG2	NM_001172	0.596165199	0.008098711
CD3G	NM_000073	0.583874879	0.000894344
XIST	NR_001564	0.579060773	0.04459401
MIF	NM_002415	0.572487214	0.00826459
PDGFA	NM_033023	0.569450437	0.015164124
LCK	NM_005356	0.56816001	0.000951069
ZAP70	NM_001079	0.563427873	0.000354397
HIST1H3H	NM_003536	0.554588894	0.031207703
HIST2H3D	NM_001123375	0.55369177	0.037529427
TAGLN2	NM_003564	0.547504013	0.015654912
IGSF1	NM_001555	0.542179632	0.000138703
EGR4	NM_001965	0.535472046	0.048557952
BMP7	NM_001719	0.527419821	3.37E-06
PTPRC	NM_002838	0.515239575	0.00020675
S100P	NM_005980	0.515183629	0.036190892
LCK	NM_005356	0.498712605	0.002170409
SOX10	NM_006941	0.490448381	0.00038315
SERPINA6	NM_001756	0.486617541	0.009777015
LY86	NM_004271	0.485235719	0.027258562
IGSF1	NM_205833	0.479991243	0.00020675
JMJD6	NM_015167	0.47485814	0.023437741
DUSP9	NM_001395	0.473397642	0.002206318
ITGAL	NM_002209	0.468937424	0.00146411
CD3EAP	NM_012099	0.466690035	0.033619138
TRAF4	NM_004295	0.459762781	0.018264215
RUNX3	NM_001031680	0.447501231	0.004554604
C4BPB	NM_000716	0.440301809	0.010061543
CELSR1	NM_014246	0.437663994	0.002598368
FGL1	NM_201553	0.430490339	0.013338196
APOA1	NM_000039	0.423583888	0.008219187
CDK1	NM_001170406	0.413828204	0.00205211
CD8B	NM_172102	0.409136643	0.008487134
SERPINA11	NM_001080451	0.406789475	0.000317988
SERPINB6	NM_004568	0.406373202	0.028472752
PTPRC	NM_002838	0.405125806	0.024910217
APOA1	NM_000039	0.404348777	0.002095478
ICAM3	NM_002162	0.403350903	0.026843708
SERPINB5	NM_002639	0.399055598	0.007854629
MLF2	NM_005439	0.383560741	0.014322879
S100A1	NM_006271	0.382506038	0.029383089
VWF	NM_000552	0.382297297	0.042150663
RUNX1	NM_001122607	0.37776294	0.02706205
IGSF1	NM_205833	0.377701982	0.033577993
MAP4K2	NM_004579	0.377258808	0.03271615
COL11A2	NM_080680	0.375561677	0.012545969
TRAF3IP3	NM_025228	0.374370562	0.00184359
BMS1P1	NR_026566	0.373629382	0.007090595
FGFR3	NM_000142	0.372608572	0.004411069
IRF5	NM_001098627	0.37009359	0.038468333
SERPINF2	NM_000934	0.367716701	0.003492405
COL11A2	NM_080680	0.364429256	0.011090518
ADAMTSL5	NM_213604	0.356959447	0.007932017
DUSP15	NM_080611	0.354978882	0.00590183
S100Z	NM_130772	0.353679009	0.00557903
BMP7	NM_001719	0.348404979	0.013709404
FGFR4	NM_213647	0.347051332	0.012814052
KLB	NM_175737	0.344002616	0.011003689
CCR2	NM_001123041	0.341217981	0.042150663
CSF3R	NM_156039	0.33674423	0.037226604
SERPINA7	NM_000354	0.336453223	0.018098389
HDAC10	NM_032019	0.335082191	0.048275522
CR2	NM_001006658	0.334888589	0.023758771
BCOR	ENST00000501455	0.329143462	0.037722482
ITGB2	ENST00000545414	0.329007232	0.03479601
OGFR	NM_007346	0.32774514	0.018198571
IL2RG	NM_000206	0.326628143	0.01508551
SOX3	NM_005634	0.31996659	0.039036159
COQ4	ENST00000372875	0.318679881	0.000889904

CD8B	NM_004931	0.317026774	0.046076549
DUS1L	NM_022156	0.316204449	0.034874751
CD6	NM_006725	0.314670004	0.043452594
CD69	ENST00000416624	0.310779984	0.043051736
HYAL4	NM_012269	0.310661815	0.036245104
XIST	NR_001564	0.309275374	0.010999758
S100A1	NM_006271	0.303604829	0.031479719
ALDH1L1	NM_012190	0.295978435	0.028329683
CD3E	NM_000733	0.293858715	0.01475548
HYAL3	NM_003549	0.293675056	0.042108398
COL2A1	NM_001844	0.290274313	0.031432323
TLR9	NM_017442	0.287576316	0.049827066
TNFRSF19	ENST00000464735	0.277553034	0.042321311
SYK	NM_003177	0.276488176	0.008356923
STK24	NM_001032296	0.275214232	0.044297156
RUNX3	NM_001031680	0.267407224	0.033676305
COX8A	NM_004074	0.263362973	0.028225773
TNFRSF13C	NM_052945	0.259457609	0.039348443
CD4	NM_000616	0.257583347	0.040475175
SERPINA4	NM_006215	0.251699401	0.016470209
HSD17B2	NM_002153	0.231010656	0.048112599
IGSF3	NM_001542	0.230518674	0.020696963
HSD17B7	ENST00000367915	0.212382766	0.044439755
EDF1	NM_153200	0.207028684	0.049510627
WDR53	NM_182627	-0.216575164	0.049399844
IL7	ENST00000379113	-0.254239979	0.032807382
STK3	ENST00000424861	-0.261863767	0.032305966
HDAC3	NM_003883	-0.271219544	0.048391636
WDR5B	NM_019069	-0.295940397	0.035352388
TRAF6	NM_145803	-0.300743004	0.022614093
CCR3	NM_001837	-0.30152348	0.00835685
DUSP19	NM_080876	-0.324847496	0.011035235
ZFX	NM_003410	-0.331652265	0.032904079
COQ5	NM_032314	-0.337826995	0.007585476
MRFAP1L1	NM_203462	-0.338843035	0.010915795
TP53	NM_000546	-0.359580571	0.010127931
CASP8	NM_033358	-0.36803978	0.000586712
C1RL	NM_016546	-0.368341308	0.039822202
COQ5	NM_032314	-0.369498191	0.002422323
CASP8AP2	NM_012115	-0.370848314	0.035900648
IRF2BP1	NM_015649	-0.382458574	0.015390496
COQ10A	NM_144576	-0.388542935	0.012863641
MMP19	NM_002429	-0.389277398	0.044955217
IFNAR1	NM_000629	-0.393904724	0.035900648
SDHD	NM_003002	-0.394055901	0.009999465
TANK	NM_004180	-0.395261286	0.023084674
LTK	NM_002344	-0.451062408	0.005267769
IL6ST	NM_002184	-0.460787651	0.028182629
HEXB	NM_000521	-0.473653992	0.048216691
TANK	NM_133484	-0.475037376	0.031262966
EGFR	NM_005228	-0.489088208	0.026827549
CSF1	NM_000757	-0.503520148	0.027096026
VEGFB	NM_003377	-0.504630145	0.003855043
NFKB1	NM_003998	-0.512759294	0.008859069
IL4R	NM_001008699	-0.527921367	0.006079353
FGF16	NM_003868	-0.543628292	0.010642552
HSD17B7	NM_016371	-0.565959578	0.000994163
EGFR	NM_201282	-0.578005894	0.001772152
IL6ST	NM_001190981	-0.586791354	0.020542856
C1QTNF4	NM_031909	-0.637040108	0.035900648
CD302	NM_014880	-0.642257917	0.031971681
LY96	NM_015364	-0.653876879	0.047190042
HSD17B7	NM_016371	-0.685616652	0.000281158
FGFR1	NM_001174066	-0.689684237	0.012380576
WNT16	NM_057168	-0.711772481	0.044004218
DUSP8	NM_004420	-0.715862195	0.006188105
CSF1	NM_172212	-0.735604032	0.033080896
IL17RD	NM_017563	-0.762314586	0.028834618
SERPINF1	NM_002615	-0.770837353	0.043935366
SERPING1	NM_000062	-0.77481374	0.028225773
WARS2	NM_201263	-0.789939771	0.018198571
IL6ST	NM_001190981	-0.792841883	0.002227704
HSD11B1	NM_181755	-0.795113663	0.016661893
IRAK3	NM_007199	-0.815829948	0.000837308
MMP2	NM_004530	-0.884713834	0.00146411
C1S	NM_201442	-0.895301832	0.011435542
TLR5	NM_003268	-0.930755564	0.010062697
CEBPA	NM_004364	-0.935323864	0.018744202
TGFBR3	NM_003243	-0.963083349	0.035517317
IL1R1	NM_000877	-0.973405185	0.017542592
TNFAIP2	NM_006291	-0.989078563	0.045498872
CCL18	NM_002988	-0.989927163	0.042589861
S100B	NM_006272	-1.006830676	0.026484494
COX1	ENST00000361624	-1.201795133	0.000554592
SERPING1	NM_000062	-1.261808882	0.000635117
CD34	NM_001025109	-1.462568414	0.003894013

FIGF	NM_004469	-1.465115246	0.01235823
CTSK	NM_000396	-1.614271298	0.000482475
CD34	NM_001773	-1.767836622	0.006537856
CXCL14	NM_004887	-1.965472373	0.008544061
TNFSF15	NM_005118	-2.446820772	0.001422026
CTSH	NM_004390	-2.465455782	0.009752533

Table 81: Selected significantly differentially expressed genes between estRA and Norm fibroblasts in the unstimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
KRT19	NM_002276	2.189430413	3.99E-05
KRT19P2	NR_036685	1.940385599	0.000407656
MMP3	NM_002422	1.827997187	0.006693999
HAS2	NM_005328	1.789285196	0.001480909
HIST2H3A	NM_001005464	1.685362715	0.003919487
COL13A1	NM_080801	1.62669414	0.02876373
KRT34	NM_021013	1.604795472	0.048381524
KRT14	NM_000526	1.562681499	0.00314875
KRT14	NM_000526	1.489560735	0.001433639
CDH13	NM_001257	1.418059311	0.014711404
SERPINE1	NM_000602	1.377489298	0.001139117
HMMR	NM_012484	1.327489833	0.010156929
KRT16P2	NR_029392	1.292797676	0.005228299
LIF	NM_002309	1.255486032	0.019379977
ITGA10	NM_003637	1.23185833	0.0130741
KRT16	NM_005557	1.217348644	0.001761719
KRT14	NM_000526	1.180138715	0.005449955
IER3	NM_003897	1.169344158	0.023911901
HS3ST3B1	ENST00000360954	1.144851692	0.004562995
HIST1H3F	NM_021018	1.13914114	0.009308184
CDH2	NM_001792	1.132418694	0.03976666
FGF2	NM_002006	1.124920521	0.01438476
HIST1H4L	NM_003546	1.118802026	0.004735409
HIST1H3B	NM_003537	1.104808731	0.030528908
HIST1H3D	NM_003530	1.092487539	0.011309103
CDH13	NM_001220491	1.077322244	0.009083508
COL4A1	NM_001845	1.076079436	0.012811628
ADAMTS4	NM_005099	1.04372667	0.00474268
HIST1H3H	NM_003536	1.043429241	0.009279571
IER3	NM_003897	1.022540587	0.023623605
HIST1H4C	NM_003542	1.022057226	0.00634607
CLEC11A	NM_002975	1.014251424	0.004501923
CDK1	NM_001786	0.996648391	0.027989256
MAPK13	NM_002754	0.969919802	0.014472881
HIST1H4B	NM_003544	0.89692495	0.01725275
COL4A2	NM_001846	0.880828884	0.047541417
TGFB1	NM_000660	0.87625039	0.000200112
SOX9	NM_000346	0.827795139	0.022572237
FGF5	NM_004464	0.751958228	0.008472839
CTPS	NM_001905	0.717834199	0.010323912
ITGA3	NM_002204	0.694243394	0.028708223
MAPK13	NM_002754	0.688882877	0.015409767
PTPRB	NM_002837	0.674420497	0.00198808
JMJD6	NM_015167	0.618133726	0.00055147
COL22A1	ENST00000522546	0.605288043	0.022754527
HIF1A	NM_181054	0.600684758	0.031276781
IFNE	NM_176891	0.587794191	0.022291213
HIST1H3E	NM_003532	0.581963799	0.042848394
CDH4	NM_001794	0.573944523	0.018374224
COL22A1	NM_152888	0.568403419	0.037699184
HIST1H3G	NM_003534	0.561804355	0.039146149
HIST2H3D	NM_001123375	0.561195838	0.011183396
IGLL5	NM_001178126	0.546560306	0.000362893
HIST2H2BF	NM_001161334	0.543561557	0.039557566
HIST1H2BE	NM_003523	0.539579871	0.029636511
P4HA2	NM_004199	0.539128029	0.021032893
LIF	NM_002309	0.518898513	0.016046527
CTPS	NM_001905	0.502366133	0.019456206
WISP1	NM_080838	0.493922375	0.026587838
HIST1H3H	NM_003536	0.487328886	0.023626675
S100P	NM_005980	0.48646279	0.017760143
MAPK12	NM_002969	0.483728638	0.010602695
HIST1H2BB	NM_021062	0.479511234	0.035576248
CASP3	NM_004346	0.479139829	0.02692071
ICAM3	NM_002162	0.476738714	0.001836711
ADAMTS6	NM_197941	0.474134363	0.028379159
CD3EAP	NM_012099	0.461933843	0.011456968
MIF	NM_002415	0.461498398	0.011236929
TGIF1	NM_170695	0.456820101	0.049874114
ITGB1	NM_133376	0.448741617	0.03236398
MLF2	NM_005439	0.447208728	0.000771277
PLOD1	NM_000302	0.438760597	0.022243127
CSK	NM_004383	0.429708183	0.006664079
TAGLN2	NM_003564	0.427965002	0.02510745
HIST1H2BO	NM_003527	0.427105713	0.030748857

CD99P1	NR_033381	0.417902985	0.033758005
ITGA5	NM_002205	0.414729926	0.026788002
TRAF4	NM_004295	0.411524965	0.011607931
HIST1H2BJ	NM_021058	0.408091466	0.043364553
HIST1H2BH	NM_003524	0.401033706	0.04670105
STK24	NM_001032296	0.390500091	0.001080073
STK24	NM_001032296	0.385180202	0.000771277
IL17RE	NM_153483	0.377673162	0.003750549
HS3ST3B1	NM_006041	0.372470756	0.047281914
ALDH4A1	NM_170726	0.368786031	0.022334293
RB1	NM_000321	0.35794809	0.023252579
HIST4H4	NM_175054	0.350552464	0.028105365
COL17A1	NM_000494	0.342519841	0.042950103
TBRG4	NM_030900	0.335408326	0.020683122
VWF	NM_000552	0.324811933	0.038579498
MAP4K2	NM_004579	0.300208544	0.04320633
COX8A	NM_004074	0.299497699	0.002983065
ISLR2	NM_020851	0.298337439	0.025409091
CD8A	NM_001145873	0.287896487	0.018142244
HSF1	NM_005526	0.281727877	0.037780423
ILF3	NM_012218	0.280727183	0.048784378
DUSP15	NM_080611	0.260776143	0.016355341
COX10	NM_001303	0.255349494	0.023140176
RUNX1	NM_001001890	0.24643083	0.048956557
ITGB7	NM_000889	0.246309898	0.023015171
TRAF7	NM_032271	0.24178975	0.034893898
ADAMTSL5	NM_213604	0.237075476	0.037216073
SOX30	NM_178424	0.215369068	0.038177193
TFAP2A	ENST00000478375	-0.201671648	0.030458255
SOX13	ENST00000367203	-0.204211082	0.025725872
INHA	NM_002191	-0.213977373	0.034801
IGSF11	NM_152538	-0.221581122	0.010081613
FBN3	NM_032447	-0.223505456	0.030874478
SDHC	NM_003001	-0.227047969	0.046242082
HDAC3	NM_003883	-0.229775592	0.044642733
MRFAP1L1	NM_203462	-0.230104863	0.041443182
IL16	NM_001172128	-0.232712227	0.036106811
CD99L2	NM_031462	-0.257241394	0.033743699
COQ5	NM_032314	-0.263594723	0.010156929
IFNAR2	NM_000874	-0.266939162	0.041030723
HSD17B7	NM_016371	-0.269451016	0.022569051
CCR3	NM_001837	-0.273080533	0.004646255
CDH1	NM_004360	-0.276791487	0.04844452
COX15	NM_078470	-0.290055547	0.009134035
CD46	NM_172350	-0.291704254	0.046789339
CD247	NM_198053	-0.300078855	0.014238982
BCL9	NM_004326	-0.301649187	0.046175213
FA2H	NM_024306	-0.312191655	0.041283878
ZFP36L1	ENST00000408913	-0.31453019	0.045915852
FGG	NM_000509	-0.318457965	0.0418904
DUSP13	NM_001007271	-0.320542623	0.005199743
IFNAR1	NM_000629	-0.32624008	0.038224047
LTB	NM_002341	-0.328482011	0.00757596
CD81	NM_004356	-0.329909617	0.027685955
KLRK1	NM_007360	-0.334254467	0.003608484
COL6A6	NM_001102608	-0.334336072	0.039170871
JMJD4	NM_023007	-0.337737888	0.010081613
COL4A3BP	NM_001130105	-0.34016734	0.041622859
EGFR	NM_201282	-0.340246268	0.029079764
C1R	NM_001733	-0.342904927	0.030086179
KLRG1	NM_005810	-0.366614698	0.005922683
ADAMTSL4	NM_025008	-0.368918756	0.044925904
TANK	NM_004180	-0.371751454	0.010499659
ZFX	NM_003410	-0.373125955	0.004238638
SDHD	NM_003002	-0.373628169	0.003882858
MMP19	NM_002429	-0.37473214	0.020086567
FGFR1	NM_023110	-0.378136509	0.017089095
COQ10A	NM_144576	-0.393478708	0.00284653
TIMP2	NM_003255	-0.395938758	0.039221408
ADAMTS7	NM_014272	-0.396746295	0.011236929
HEXB	NM_000521	-0.406911751	0.041360455
CD33	NM_001772	-0.41005615	0.004871501
NFKBIE	NM_004556	-0.41408966	0.036795144
ISLR	NM_005545	-0.422733971	0.003688789
FYB	NM_001465	-0.429049026	0.014819547
HIF3A	ENST00000457865	-0.430299618	0.006374595
SERPING1	NM_000062	-0.436854819	0.000756435
IL6R	NM_000565	-0.439768951	0.022728597
XLOC_l2_014830	THC2583825	-0.440766155	0.002359484
SRC	NM_005417	-0.445894852	0.022431042
CD79A	NM_001783	-0.450236766	0.00186247
IRF1	NM_002198	-0.459039357	0.048298168
VEGFB	NM_003377	-0.461774698	0.001836711
HSD17B7	NM_016371	-0.468675404	0.001368289
IL3RA	NM_002183	-0.471487964	0.039026837
SOX13	NM_005686	-0.486060177	0.004238638

TANK	NM_133484	-0.490904884	0.00757596
CD248	NM_020404	-0.49412441	0.035624611
IL34	NM_152456	-0.497337416	0.008605863
FCGRT	NM_004107	-0.499161307	0.006415354
HLA-A	NM_002116	-0.515631512	0.02876373
NFKB1	NM_003998	-0.531629228	0.001391569
TNFAIP8	NM_014350	-0.537547014	0.017754306
HSD17B7	NM_016371	-0.544250835	0.000724018
CD300A	NM_007261	-0.549174049	0.012341706
IL6ST	NM_002184	-0.556535928	0.001630918
IL6ST	NM_001190981	-0.56140759	0.008027909
HLA-A	NM_002116	-0.564290364	0.00641976
B2M	NM_004048	-0.564892158	0.029866885
TP53	NM_000546	-0.571058551	2.54E-06
IL6ST	NM_001190981	-0.571888797	0.00178201
WNT16	NM_057168	-0.590682898	0.045347928
STAT5A	NM_003152	-0.60199831	0.000146006
ADAMTSL4	NM_025008	-0.602557028	0.03706661
HABP4	NM_014282	-0.606307275	0.015831113
HAPLN2	NM_021817	-0.609525617	0.005208465
C1S	NM_201442	-0.611940404	0.041516119
ITM2B	NM_021999	-0.612407232	0.014091016
HLA-J	NR_024240	-0.626966478	0.036117481
HLA-E	NM_005516	-0.629206671	0.003486408
CLEC3B	NM_003278	-0.637542625	0.015409767
HLA-F	NM_018950	-0.637717689	0.00947685
IL6ST	NM_002184	-0.641183624	0.011896465
ITM2B	NM_021999	-0.646536604	0.00396943
HLA-G	NM_002127	-0.654875547	0.002318004
TNFRSF19	NM_018647	-0.662161966	0.004522701
MBP	NM_001025101	-0.662859104	0.00544915
TNXB	NM_019105	-0.663687359	0.027618778
CXCL16	NM_022059	-0.664207573	0.018268841
C1RL	NM_016546	-0.686031449	0.02973981
IGSF10	NM_178822	-0.693404506	0.000689217
CD55	NM_000574	-0.699364711	0.013325148
PDGFRB	NM_002609	-0.71335253	0.021788497
IL6ST	NM_001190981	-0.718578681	0.001092391
C1QTNF4	NM_031909	-0.722830593	0.004484807
CDKN1B	NM_004064	-0.723230615	0.000756435
TNXB	NM_019105	-0.724207198	0.046434742
DUSP8	NM_004420	-0.732458371	0.00098843
MMP2	NM_004530	-0.744975505	0.001630918
CCL28	NM_148672	-0.749833786	0.010769207
C1S	NM_001734	-0.766423582	0.013562522
HLA-C	M26429	-0.76955304	0.016695038
IL17RD	NM_017563	-0.781509573	0.007252692
TNXB	NM_019105	-0.787706494	0.020086567
CD302	NM_014880	-0.787784734	0.001720166
HSD11B1	NM_181755	-0.799860649	0.004338195
CCL18	NM_002988	-0.806024488	0.048521774
FGFR1	NM_001174066	-0.835534941	0.000457219
CARD6	NM_032587	-0.851636177	0.012261966
CSF1	NM_172212	-0.853422715	0.00314875
IL15RA	NM_172200	-0.873654131	0.027387919
CTSF	NM_003793	-0.875634554	0.007365432
TGFBR3	NM_003243	-0.882754621	0.021039118
TLR5	NM_003268	-0.891353432	0.003567363
IL1R1	NM_000877	-0.892734652	0.025843516
ICAM1	NM_000201	-0.896757871	0.008925009
CXCL16	NM_001100812	-0.940704676	0.03079632
SERPING1	NM_000062	-0.949968985	0.001410684
SERPING1	NM_000062	-0.972915925	0.001892867
SOX8	NM_014587	-0.997267448	0.00227453
CEBPA	NM_004364	-1.058895917	0.001630918
PPAP2A	NM_176895	-1.07111984	0.012809485
IFI6	NM_022873	-1.140804558	0.027165255
CD34	NM_001025109	-1.180240433	0.005824153
CTSK	NM_000396	-1.203166852	0.00218423
CD14	NM_001174104	-1.268307047	0.011888954
PDGFRL	NM_006207	-1.290924312	0.020449063
FGF9	NM_002010	-1.293298131	0.007176846
C1R	NM_001733	-1.325051515	0.011115066
FIGF	NM_004469	-1.333780281	0.006737687
AHR	NM_001621	-1.341253134	0.001930314
AHR	NM_001621	-1.342170198	0.004736242
TNFAIP2	NM_006291	-1.433200695	0.000589664
INHBB	NM_002193	-1.500385674	0.007753531
CD34	NM_001773	-1.516570755	0.005689158
CXCL14	NM_004887	-1.648422153	0.008722197
CCRL1	NM_178445	-1.750719297	0.018294532
S100B	NM_006272	-1.777244154	5.01E-06
C3	NM_000064	-1.924817255	0.004225252
TNFSF15	NM_005118	-2.024213571	0.00187624
C3	NM_000064	-2.351077604	0.004972093
CTSH	NM_004390	-2.526888113	0.001738199

Table 82: Selected significantly differentially expressed genes between veRA and Norm fibroblasts in the unstimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
KRT19	NM_002276	2.047334531	9.18E-05
MMP3	NM_002422	2.018429113	0.002682924
KRT19P2	NR_036685	1.881864767	0.000594591
HAS2	NM_005328	1.596286355	0.003797106
HIST2H3A	NM_001005464	1.435765945	0.01201747
HS3ST3B1	ENST00000360954	1.239448628	0.001930996
SERPINE1	NM_000602	1.21906278	0.003461711
KRT16	NM_005557	1.210111808	0.001679199
HMMR	NM_012484	1.175656838	0.02122041
ITGA10	NM_003637	1.132300721	0.021118066
KRT14	NM_000526	1.115853709	0.014660786
KRT16P2	NR_029392	1.093749498	0.01664042
ADAMTS4	NM_005099	1.063785195	0.003594231
HIST1H3B	NM_003537	1.045823022	0.038294596
FGF2	NM_002006	1.03740845	0.022634663
CDH13	NM_001220491	1.010532539	0.012905526
TNC	NM_002160	1.000584288	0.045998211
HIST1H3F	NM_021018	0.993951307	0.021613185
CDK1	NM_001786	0.991851284	0.026340432
HIST1H4L	NM_003546	0.972926646	0.012423182
HIST1H3D	NM_003530	0.95987386	0.024551792
HIST1H3H	NM_003536	0.954674799	0.015972925
HIST1H4C	NM_003542	0.931454641	0.011204044
KRT14	NM_000526	0.926980337	0.027243403
HIST1H4B	NM_003544	0.825071508	0.026452876
TGFB1	NM_000660	0.821645797	0.000341475
MAPK13	NM_002754	0.781374991	0.048379629
FGF1	NM_000800	0.773624931	0.022157629
CTPS	NM_001905	0.757814833	0.006000501
CLEC11A	NM_002975	0.723258585	0.04002402
HIST1H2AI	NM_003509	0.713470431	0.041634742
VEGFC	NM_005429	0.701183619	0.043125094
ITGA3	NM_002204	0.662251123	0.034478237
P4HA2	NM_004199	0.646702387	0.004840892
LDHA	NM_005566	0.616037173	0.01447501
CTPS	NM_001905	0.587062773	0.005438252
HIF1A	NM_181054	0.578991676	0.035596405
FGF5	NM_004464	0.575226857	0.042759075
WISP1	NM_080838	0.564518835	0.009437422
HIST2H2BF	NM_001161334	0.563320884	0.029707394
MAPK13	NM_002754	0.557436504	0.049386941
IFNE	NM_176891	0.554626315	0.028756527
CD3EAP	NM_012099	0.542471002	0.002982706
PLOD1	NM_000302	0.54233831	0.017096416
JMJD6	NM_015167	0.542286385	0.001911257
HIST1H3G	NM_003534	0.538299535	0.044961252
HIST1H2BE	NM_003523	0.523776616	0.032503049
HIST1H2AE	NM_021052	0.507594767	0.03873044
MIF	NM_002415	0.502615719	0.005163134
CASP3	NM_004346	0.499051318	0.018906721
IFRD2	NM_006764	0.49354181	0.016302988
HIST2H3D	NM_001123375	0.490346494	0.025027181
CSK	NM_004383	0.462771477	0.003353746
HIST1H3H	NM_003536	0.459165769	0.030453607
HIST1H2BB	NM_021062	0.45710849	0.042255261
HIST1H2BH	NM_003524	0.438338033	0.026475862
PLOD1	NM_000302	0.435667034	0.020979642
TAGLN2	NM_003564	0.430912512	0.021734892
S100P	NM_005980	0.422233417	0.037521235
ALDH4A1	NM_170726	0.421401854	0.007791741
CDH15	NM_004933	0.418983324	0.032503049
HIST1H2BO	NM_003527	0.417071336	0.032515904
TRAF4	NM_004295	0.416957802	0.009361588
MAPKAP1	NM_001006618	0.416491784	0.028926409
MAP4K2	NM_004579	0.412506598	0.004491357
IGLL5	NM_001178126	0.409559577	0.00496458
HIST1H2BD	NM_021063	0.401755761	0.04695147
STK24	NM_001032296	0.391736493	0.000707628
STK24	NM_001032296	0.39101455	0.001046832
ICAM3	NM_002162	0.390152147	0.009066335
ITGA5	NM_002205	0.389728848	0.034837852
MLF2	NM_005439	0.377715908	0.003746528
COL17A1	NM_000494	0.377134984	0.022795194
HS3ST3B1	NM_006041	0.364933101	0.048920256
CSF1R	NM_005211	0.356938191	0.048551731
RUNX1	NM_001001890	0.350938765	0.004091494
ILF3	NM_012218	0.341037712	0.047238055
ISLR2	NM_020851	0.314674613	0.016374661
H2AFY	NM_138610	0.304963364	0.022817781
SERBP1	NM_001018067	0.288756173	0.036318774
COX8A	NM_004074	0.28431829	0.004054495

ILF3	NM_004516	0.284173455	0.049348411
DUS1L	NM_022156	0.268757555	0.02948322
DUS1L	NM_022156	0.260774181	0.038509909
COX5A	NM_004255	0.254838794	0.027243403
COX5A	NM_004255	0.249911962	0.03586374
COL5A1	ENST00000464187	0.242604379	0.034837852
DUSP15	NM_080611	0.234375582	0.029126088
COX10	NM_001303	0.226214478	0.041775149
MMP25	NM_022468	0.225577746	0.025974129
SDHA	NM_004168	-0.200116024	0.049072951
COQ5	NM_032314	-0.206182289	0.042962242
MRFAP1L1	NM_203462	-0.22452493	0.043238224
STK3	ENST00000424861	-0.230252571	0.022282778
IL16	NM_001172128	-0.242451979	0.026269672
IGBP1	NM_001551	-0.28424553	0.040104853
KLRK1	NM_007360	-0.284318092	0.011190323
S100A5	NM_002962	-0.290472249	0.026269672
CD99L2	NM_031462	-0.293932377	0.013418062
BCL9	NM_004326	-0.302344515	0.042211769
CD81	NM_004356	-0.307622681	0.037735037
CCR3	NM_001837	-0.325546109	0.000868697
COL4A3BP	NM_001130105	-0.326609765	0.020711014
SERPING1	NM_000062	-0.32879833	0.009024133
SDHD	NM_003002	-0.336610595	0.007868495
SOX13	NM_005686	-0.34695403	0.038493068
FYB	NM_001465	-0.349510876	0.046328841
ADAMTS7	NM_014272	-0.35244595	0.022744046
TP53	NM_000546	-0.354534845	0.002514867
TANK	NM_133484	-0.361739865	0.048192497
HIF3A	ENST00000457865	-0.377868653	0.015034889
HSD17B7	NM_016371	-0.387229014	0.001063161
EGFR	NM_005228	-0.389757616	0.032745005
CTSA	NM_000308	-0.391707108	0.048943414
IL6ST	NM_002184	-0.399974798	0.0209413
VEGFB	NM_003377	-0.40106952	0.005735737
IGBP1	NM_001551	-0.402519423	0.003609053
ZFP36L1	ENST00000408913	-0.407906963	0.008003802
COL4A3BP	NM_001130105	-0.414123641	0.010951731
NFKBIE	NM_004556	-0.414599873	0.033579293
FGF16	NM_003868	-0.414890297	0.018856271
XLOC_I2_014830	THC2583825	-0.417530699	0.00346791
ISLR	NM_005545	-0.421685474	0.00341857
ADAMTSL1	NM_001040272	-0.425575399	0.019444972
TANK	NM_004180	-0.439343404	0.002491602
COQ10A	NM_144576	-0.441387175	0.000891036
IL34	NM_152456	-0.453073529	0.015331786
IL6ST	NM_001190981	-0.456696361	0.029539258
CDKN1B	NM_004064	-0.461736924	0.019090656
IL6ST	NM_001190981	-0.462416883	0.009485344
CD79A	NM_001783	-0.464203173	0.001292339
TIMP2	NM_003255	-0.477737365	0.010710119
HSD17B7	NM_016371	-0.484950352	0.000919765
HAPLN2	NM_021817	-0.494981405	0.021532197
NFKB1	NM_003998	-0.507172297	0.002018948
IGSF10	NM_178822	-0.507461436	0.009946398
FCGRT	NM_004107	-0.508765009	0.00482355
MMP19	NM_002429	-0.529848754	0.001046832
ITM2B	NM_021999	-0.53033219	0.032247889
ADAMTSL4	NM_025008	-0.535020138	0.003198202
HSD17B7	NM_016371	-0.537060039	0.00083748
STAT5A	NM_003152	-0.582817137	0.000167772
HABP4	NM_014282	-0.597118367	0.016056846
LY96	NM_015364	-0.601129637	0.025333527
CD55	NM_000574	-0.603672573	0.031246006
IL6ST	NM_001190981	-0.615626577	0.004251092
MBP	NM_001025101	-0.638262753	0.006662972
TNXB	NM_019105	-0.640177949	0.031180623
PDGFRB	NM_002609	-0.649812111	0.034180186
C1RL	NM_016546	-0.653939924	0.036086863
CDKN1B	NM_004064	-0.66834713	0.001619918
CSF1	NM_172212	-0.669320494	0.018308811
ITM2B	NM_021999	-0.671698527	0.00254402
SERPINF1	NM_002615	-0.684649067	0.028926409
MMP2	NM_004530	-0.689665736	0.003195805
CD302	NM_014880	-0.7213299	0.003517454
TNXB	NM_019105	-0.733940787	0.031644171
C1QTNF4	NM_031909	-0.736930586	0.003408916
C1S	NM_001734	-0.737708418	0.016302988
ADAMTSL4	NM_025008	-0.747882775	0.008099988
DUSP8	NM_004420	-0.749874147	0.000820339
SERPING1	NM_000062	-0.752951532	0.009321448
FGF9	NM_002010	-0.78992954	0.003777009
TNXB	NM_019105	-0.793983953	0.017047976
TNXB	NM_019105	-0.801587895	0.024431445
IL1R1	NM_000877	-0.821758005	0.037521235
SERPING1	NM_000062	-0.838574616	0.006268419

SOX8	NM_014587	-0.850123672	0.007868495
PPAP2A	NM_176895	-0.851090257	0.047366767
TLR5	NM_003268	-0.878496644	0.003509155
CARD6	NM_032587	-0.8990957	0.007444408
CARD6	NM_032587	-0.909958278	0.006531218
AHR	NM_001621	-0.914099147	0.031844827
IL17RD	NM_017563	-0.926613824	0.001438242
CEBPA	NM_004364	-0.954942289	0.003777009
TGFBR3	NM_003243	-1.023538968	0.006413539
HS3ST1	NM_005114	-1.064625386	0.044932133
CTSF	NM_003793	-1.08076549	0.001046566
COL14A1	NM_021110	-1.109826361	0.016821248
MBP	NM_001025100	-1.122017559	0.029049186
CD14	NM_001174104	-1.155681702	0.02058822
PDGFD	NM_025208	-1.195125954	0.029539258
PDGFRL	NM_006207	-1.218453339	0.026277182
C1R	NM_001733	-1.226709758	0.017171372
CDKN1C	NM_000076	-1.233684271	0.033230212
FIGF	NM_004469	-1.27411166	0.008619174
CTSK	NM_000396	-1.296814752	0.001010483
INHBB	NM_002193	-1.366965237	0.013903843
CXCL14	NM_004887	-1.377085922	0.026895624
TNFAIP2	NM_006291	-1.438484749	0.000603516
FGF9	NM_002010	-1.482823589	0.001968263
CD34	NM_001025109	-1.509492961	0.000594591
CD36	NM_001001547	-1.511082863	0.030167033
S100B	NM_006272	-1.546791482	6.34E-05
CD34	NM_001773	-1.732455958	0.001544958
C3	NM_000064	-1.879637878	0.004332559
CCRL1	NM_178445	-1.990620448	0.006301474
TNFSF15	NM_005118	-2.136303659	0.001046832
C3	NM_000064	-2.282625996	0.005655134
CTSH	NM_004390	-2.446473497	0.002219418

Table 83: Selected significantly differentially expressed genes between Res and Norm fibroblasts in the unstimulated high serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

8.8 Differentially expressed gene lists from the unstimulated low serum samples in the microarray experiments

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
ALDH3A2	NM_001031806	1.56734237	0.048275941
SERPINE2	NM_006216	1.518726505	0.035960938
TAGLN	NM_001001522	1.509365466	0.03283886
TAGLN	NM_001001522	1.144203865	0.016734525
ADH4	NM_000670	0.931777928	0.000333424
POGZ	ENST00000437847	0.560305661	0.00832172
DUSP1	NM_004417	-1.384786331	0.03283886
ADAM12	NM_003474	-1.802686277	0.035960938

Table 84: Selected significantly differentially expressed genes between Jrep and estRA fibroblasts in the unstimulated low serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
INHBB	NM_002193	1.90394431	0.006666342
ALDH3A2	NM_000382	1.671966203	0.000335378
ALDH3A2	NM_001031806	1.668092684	0.000566102
MAPK10	NM_138980	1.290702139	0.001795459
SERPINI1	NM_005025	1.25744147	0.008518211
ALDH3A1	NM_001135168	1.244444239	2.07E-05
TAGLN	NM_001001522	1.229314657	0.007416524
ALDH3A1	NM_000691	1.182079196	5.89E-05
TAGLN	NM_001001522	1.023381598	0.000872247
IL16	NM_172217	1.005194553	0.00018112
CDHR4	NM_001007540	0.918381004	0.004381714
AHRR	NM_020731	0.90801188	0.028035917
IL16	NM_004513	0.907563497	0.002101519
LTC4S	NM_145867	0.864950958	0.012483713
HLF	NM_002126	0.86154971	0.023209213
ADH4	NM_000670	0.859955388	7.87E-06
COL9A3	NM_001853	0.848752473	0.049175173
LSP1	NM_001013254	0.796417526	0.026762295
LGI1	NM_005097	0.680432554	0.025462939
XIST	NR_001564	0.668653494	0.026486061
XIST	NR_001564	0.661652254	0.037733154
POGZ	ENST00000437847	0.657705393	1.96E-06
XIST	NR_001564	0.648945231	0.042781295
RIIAD1	NM_001144956	0.612153773	0.018610847
COX11	AK124809	0.557628502	0.000422741
SERPINB1	NM_030666	0.510784758	0.010789715
VWF	NM_000552	0.42235587	0.023796776
HDAC6	NM_006044	0.372579294	0.044060068

IL6R	NM_001206866	0.347316626	0.023476767
ZFAT	ENST00000522974	0.29003749	0.02427553
MMP11	NM_005940	-0.417029515	0.020020696
COL4A3BP	NM_001130105	-0.463290217	0.009827908
MAPK12	NM_002969	-0.533152425	0.028569609
JMJD6	NM_015167	-0.542083821	0.014564147
CD276	NM_001024736	-0.554900142	0.047739687
MAPK6	NM_002748	-0.562005639	0.001198541
IGSF3	NM_001542	-0.573560847	0.031925704
IER5	NM_016545	-0.58613413	0.027115432
ADAMTS6	NM_197941	-0.59629777	0.040630281
ADAMTS6	NM_197941	-0.602396967	0.029530116
COQ10B	NM_025147	-0.616508358	0.005054796
IL1R2	NM_004633	-0.621442493	0.034813872
TGFB1	NM_000660	-0.625815216	0.029621744
CD276	NM_001024736	-0.632529072	0.009827908
IRAK1	NM_001569	-0.640863242	0.003423184
CTSA	NM_001127695	-0.640935118	0.029455012
IRAK1	NM_001569	-0.713782767	0.00205189
BCL2	NM_000633	-0.722730239	0.043102922
CD164	NM_006016	-0.778736748	0.006532227
SOX30	NM_007017	-0.798928899	0.040723891
MAPK6	NM_002748	-0.810530257	1.13E-05
ZFP36	NM_003407	-0.838247894	0.045239978
DUSP14	NM_007026	-0.861332572	0.002840002
JUN	NM_002228	-0.963175308	0.015370867
DUSP2	NM_004418	-1.14916995	0.01162803
ATF3	NM_001674	-1.160016517	0.006931131
COL4A1	NM_001845	-1.190971414	0.033859171
BMP2	NM_001200	-1.22242828	0.016829038
DUSP5	NM_004419	-1.225991169	0.026063976
FAP	NM_004460	-1.288443373	0.017241072
CXCL2	NM_002089	-1.32356235	0.038394835
ATF3	NM_001040619	-1.394853971	0.002176527
ADAM12	NM_003474	-1.487703792	0.007650797
MME	NM_007289	-1.495042886	0.030286342
FOS	NM_005252	-1.573048744	0.016549859
COL4A2	NM_001846	-1.625229126	0.002522366
DUSP1	NM_004417	-1.657100399	2.65E-05
EGR1	NM_001964	-1.785919272	0.000566102
MME	NM_007289	-1.794213622	0.020861032
MYC	NM_002467	-1.851574417	2.13E-05
MME	NM_007289	-1.862144536	0.011009178
IER3	NM_003897	-1.992504148	0.00018112
EGR2	NM_000399	-2.070960425	0.011044416
NFKBIZ	NM_031419	-2.087211603	1.96E-06
ATF3	NM_001040619	-2.105046969	0.001378265
PTGS2	NM_000963	-2.176271809	0.01369418
IER3	NM_003897	-2.220127414	0.000268428
FOSB	NM_006732	-2.293312419	0.000883401
CXCL2	NM_002089	-2.781839154	0.003423184

Table 85: Selected significantly differentially expressed genes between Jrep and veRA fibroblasts in the unstimulated low serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
INHBB	NM_002193	2.096671364	0.002199433
ALDH3A2	NM_000382	1.503319809	0.001468246
ALDH3A2	NM_001031806	1.437083617	0.003853169
SERPINE2	NM_006216	1.155894026	0.01623734
ALDH3A1	NM_001135168	1.073366064	0.000312449
ALDH3A1	NM_000691	1.056700458	0.000364906
SERPINI1	NM_005025	1.007642761	0.046083117
IL16	NM_172217	1.005578304	0.000180079
AHRR	NM_020731	1.001762444	0.012911774
IL16	NM_004513	0.981158369	0.000635411
ADH4	NM_000670	0.919956556	4.15E-07
IGFBP3	NM_001013398	0.821850406	0.047262991
LTC4S	NM_145867	0.81309832	0.020243105
CDHR4	NM_001007540	0.757335654	0.024679147
TAGLN	NM_001001522	0.670268733	0.049401089
POGZ	ENST00000437847	0.619511706	4.16E-06
COL4A5	NM_033380	0.57076289	0.044118034
COX11	AK124809	0.398680844	0.019004971
ZFAT	ENST00000522974	0.314426909	0.01249894
MMP11	NM_005940	-0.374635966	0.044859807
COL4A3BP	NM_001130105	-0.380877033	0.044118034
IRAK1	NM_001569	-0.522970408	0.02300501
MAPK6	NM_002748	-0.584066683	0.002199433
IRAK1	NM_001569	-0.624373215	0.008332486
DUSP14	NM_007026	-0.696513766	0.021396681
CTSA	NM_001127695	-0.789372428	0.00498065
BCL2	NM_000633	-0.819288785	0.018898426
JUN	NM_002228	-0.904927545	0.025218145
ICAM1	NM_000201	-1.031154979	0.018278848
DUSP1	NM_004417	-1.084599717	0.010032025

EGR1	NM_001964	-1.265911753	0.023505824
MYC	NM_002467	-1.318129973	0.003995484
NFKBIZ	NM_031419	-1.384646922	0.002208265
COL4A2	NM_001846	-1.400173682	0.010779314
ATF3	NM_001040619	-1.468767449	0.042035215
FIGF	NM_004469	-1.550350067	0.01167747
CXCL2	NM_002089	-2.168425817	0.032607725

Table 86: Selected significantly differentially expressed genes between Jrep and Res fibroblasts in the unstimulated low serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
TAGLN	NM_001001522	2.333200054	1.61E-07
COL8A1	NM_001850	2.141182109	0.024460067
SERPINE2	NM_006216	1.979583476	3.05E-05
ITGA10	NM_003637	1.913507668	0.003300872
TAGLN	NM_001001522	1.898278629	8.73E-10
CDH13	NM_001257	1.847248441	0.020467002
MAPK10	NM_138980	1.2021	0.007960021
LTC4S	NM_145867	1.187771969	0.000645619
AHRR	NM_020731	1.187288697	0.005223668
IGFBP3	NM_001013398	1.168547262	0.004381821
SERPINI1	NM_005025	1.163722257	0.028179184
ADAMTS4	NM_005099	1.136163009	0.026015709
CDH11	NM_001797	1.023945909	0.032692945
MAP3K5	NM_005923	0.968565045	0.04589645
PDGFC	NM_016205	0.942501786	0.040728128
IL16	NM_004513	0.928901763	0.003536073
ITGA4	NM_000885	0.901742066	0.036154693
ALDH3A1	NM_000691	0.887078725	0.007796841
IL16	NM_172217	0.885027284	0.002854544
COL6A1	NM_001848	0.878440109	0.001287536
ALDH3A1	NM_001135168	0.812243282	0.017057309
ADH4	NM_000670	0.771661992	0.000171219
IL1RAP	NM_002182	0.761741345	0.026744124
MAP3K5	NM_005923	0.752732322	0.033895809
PDGFA	NM_002607	0.667950681	0.040368113
SYNGR1	NM_145731	0.640880551	0.047361325
POGZ	ENST00000437847	0.619747607	3.05E-05
IL1RAP	NM_134470	0.573344482	0.036628258
IRF2BPL	NM_024496	0.571998076	0.006555971
COX11	AK124809	0.452435564	0.011407861
ZFAT	ENST00000522974	0.3416579	0.011528001
DUSP16	NM_030640	-0.370878056	0.033895809
COL4A3BP	NM_001130105	-0.447453797	0.022680341
VEGFB	NM_003377	-0.549880171	0.00445237
CTSA	NM_000308	-0.602683271	0.030324267
ZAK	NM_016653	-0.614034768	0.00258847
CD44	NM_000610	-0.629713508	0.014048041
MAPK6	NM_002748	-0.638264518	0.001648396
CTSA	NM_001127695	-0.665559861	0.039699706
ZAK	NM_016653	-0.753596567	0.002672441
MMP2	NM_004530	-0.763734165	0.016565824
DUSP8	NM_004420	-0.777123232	0.007893328
IRAK1	NM_001569	-0.831310285	0.000607322
IRAK1	NM_001569	-0.842629288	0.000182746
SOX30	NM_007017	-0.87162515	0.038962645
DUSP14	NM_007026	-0.896146625	0.00373103
BCL2	NM_000657	-0.961378995	0.001569513
NFKBIZ	NM_031419	-1.06375703	0.044196282
BCL2	NM_000633	-1.248076917	0.000309651
CTSK	NM_000396	-1.264099578	0.017107967
EGR1	NM_001964	-1.401876999	0.017858101
CCL18	NM_002988	-1.461682047	0.006348621
CILP	NM_003613	-1.599474994	0.048854546
CTSH	NM_004390	-3.226650996	0.001838196
FIGF	NM_004469	-3.314412236	2.91E-08

Table 87: Selected significantly differentially expressed genes between Jrep and Norm fibroblasts in the unstimulated low serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
LGI1	NM_005097	-1.096769962	0.018868119
FIGF	NM_004469	-2.735315848	0.000308702

Table 88: Selected significantly differentially expressed genes between estRA and Norm fibroblasts in the unstimulated low serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
RIIAD1	NM_001144956	-0.72606679	0.023453095

Table 89: Selected significantly differentially expressed genes between veRA and Res fibroblasts in the unstimulated low serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
PTGS2	NM_000963	2.30601602	0.012021238
IL6	NM_000600	2.238818238	0.021768154
ITGA10	NM_003637	1.915065416	0.003775361
ID1	NM_002165	1.674500883	0.007348784
LIF	NM_002309	1.485265582	0.045249235
SERPINE1	NM_000602	1.434020996	0.009224673
IER3	NM_003897	1.427687886	0.045849919
ADAMTS4	NM_005099	1.32955867	0.005912306
TNFRSF12A	NM_016639	1.169537578	0.005296758
CDH11	NM_001797	1.147989536	0.010721839
HS3ST3B1	ENST00000360954	1.14074101	0.038958069
SERPINE2	NM_006216	1.133342825	0.028412968
MYC	NM_002467	1.128207399	0.026733655
TAGLN	NM_001001522	1.103885397	0.027851664
NFKBIZ	NM_031419	1.023454573	0.049322907
HIF1A	NM_181054	0.983023676	0.010987424
TAGLN	NM_001001522	0.874897031	0.009224673
RUNX1	NM_001122607	0.746054464	0.044119169
SMAD7	NM_005904	0.721509262	0.034542581
CD164	NM_006016	0.659625523	0.036785184
CLEC2D	NM_001004419	0.535847596	0.022791438
CLEC2D	NM_001004419	0.524617779	0.033749409
JMJD6	NM_015167	0.516459498	0.031491725
COQ10B	NM_025147	0.478153868	0.034542581
VEGFB	NM_003377	-0.426106709	0.034542581
HLA-A	NM_002116	-0.581916574	0.041671712
DUSP8	NM_004420	-0.641171448	0.034473059
HLA-A	NM_002116	-0.70743281	0.02710327
HLA-G	NM_002127	-0.710891751	0.011590118
SMAD3	NM_005902	-0.741503221	0.040597208
IL17RD	NM_017563	-0.855038006	0.032366949
TLR5	NM_003268	-0.863934157	0.038958069
LGI1	NM_005097	-1.031103743	0.000836877
CD34	NM_001025109	-1.244265529	0.034500725
FGF9	NM_002010	-1.396302674	0.034542581
INHBB	NM_002193	-1.549186416	0.048633353
CD34	NM_001773	-1.70662611	0.020924614
FIGF	NM_004469	-2.137874139	0.000836877
C3	NM_000064	-2.341739161	0.043160544
CTSH	NM_004390	-2.646091674	0.012521152

Table 90: Selected significantly differentially expressed genes between veRA and Norm fibroblasts in the unstimulated low serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

HGNC symbol	Database ID	Log2 fold-change	Adjusted P Value
TAGLN	NM_001001522	1.482312554	0.002273658
TAGLN	NM_001001522	1.228009896	2.43E-05
CDH11	NM_001797	1.13999156	0.031375864
LGI1	NM_005097	-0.907878903	0.006115137
INHBB	NM_002193	-1.741913471	0.049320232
FIGF	NM_004469	-1.76406217	0.010757188

Table 91: Selected significantly differentially expressed genes between Res and Norm fibroblasts in the unstimulated low serum treatment group sorted by log2 fold change. P values adjusted using the Benjamini-Hochberg method.

9 REFERENCES

A-Gonzalez, N., Guillen, J.A., Gallardo, G., Diaz, M., De La Rosa, J.V., Hernandez, I.H., Casanova-Acebes, M., Lopez, F., Tabraue, C., Beceiro, S., Hong, C., Lara, P.C., Andujar, M., Arai, S., Miyazaki, T., Li, S., Corbi, A.L., Tontonoz, P., Hidalgo, A. and Castrillo, A. (2013) 'The nuclear receptor LXRα controls the functional specialization of splenic macrophages'. *Nature Immunology,* 14: 831.

Abe, R., Donnelly, S.C., Peng, T., Bucala, R. and Metz, C.N. (2001) 'Peripheral blood fibrocytes: differentiation pathway and migration to wound sites'. *J Immunol,* 166: (12): 7556-7562.

Acharya, P.S., Majumdar, S., Jacob, M., Hayden, J., Mrass, P., Weninger, W., Assoian, R.K. and Pure, E. (2008) 'Fibroblast migration is mediated by CD44-dependent TGF beta activation'. *J Cell Sci,* 121: (Pt 9): 1393-1402.

Aicher, W.K., Alexander, D., Haas, C., Kuchen, S., Pagenstecher, A., Gay, S., Peter, H.H. and Eibel, H. (2003) 'Transcription factor early growth response 1 activity up-regulates expression of tissue inhibitor of metalloproteinases 1 in human synovial fibroblasts'. *Arthritis Rheum,* 48: (2): 348-359.

Ainscough, J.S., Gerberick, G.F., Kimber, I. and Dearman, R.J. (2015) 'Interleukin-1β Processing Is Dependent on a Calcium-mediated Interaction with Calmodulin'. *The Journal of Biological Chemistry,* 290: (52): 31151-31161.

Aird, D., Ross, M.G., Chen, W.-S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D.B., Nusbaum, C. and Gnirke, A. (2011) 'Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries'. *Genome Biology,* 12: (2): R18.

Al-Saadany, H.M., Hussein, M.S., Gaber, R.A. and Zaytoun, H.A. (2016) 'Th-17 cells and serum IL-17 in rheumatoid arthritis patients: Correlation with disease activity and severity'. *The Egyptian Rheumatologist,* 38: (1): 1-7.

Alarcon, B., Berkhout, B., Breitmeyer, J. and Terhorst, C. (1988) 'Assembly of the human T cell receptor-CD3 complex takes place in the endoplasmic reticulum and involves intermediary complexes between the CD3-gamma.delta.epsilon core and single T cell receptor alpha or beta chains'. *J Biol Chem,* 263: (6): 2953-2961.

Aletaha, D., Neogi, T., Silman, A.J., Funovits, J., Felson, D.T., Bingham, C.O., 3rd, Birnbaum, N.S., Burmester, G.R., Bykerk, V.P., Cohen, M.D., Combe, B., Costenbader, K.H., Dougados, M., Emery, P., Ferraccioli, G., Hazes, J.M., Hobbs, K., Huizinga, T.W., Kavanaugh, A., Kay, J., Kvien, T.K., Laing, T., Mease, P., Menard, H.A., Moreland, L.W., Naden, R.L., Pincus, T., Smolen, J.S., Stanislawska-Biernat, E., Symmons, D., Tak, P.P., Upchurch, K.S., Vencovsky, J., Wolfe, F. and Hawker, G. (2010) '2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative'. *Ann Rheum Dis,* 69: (9): 1580-1588.

Allard, S.A., Maini, R.N. and Muirden, K.D. (1988) 'Cells and matrix expressing cartilage components in fibroblastic tissue in rheumatoid pannus'. *Scand J Rheumatol Suppl,* 76: 125-129.

Alonzi, T., Fattori, E., Lazzaro, D., Costa, P., Probert, L., Kollias, G., De Benedetti, F., Poli, V. and Ciliberto, G. (1998) 'Interleukin 6 is required for the development of collagen-induced arthritis'. *J Exp Med,* 187: (4): 461-468.

Ambarus, C.A., Krausz, S., Van Eijk, M., Hamann, J., Radstake, T.R.D.J., Reedquist, K.A., Tak, P.P. and Baeten, D.L.P. (2012a) 'Systematic validation of specific phenotypic markers for in vitro polarized human macrophages'. *Journal of Immunological Methods,* 375: (1–2): 196-206.

Ambarus, C.A., Noordenbos, T., De Hair, M.J., Tak, P.P. and Baeten, D.L. (2012b) 'Intimal lining layer macrophages but not synovial sublining macrophages display an IL-10 polarized-like phenotype in chronic synovitis'. *Arthritis Res Ther,* 14: (2): R74.

Ancuta, P., Liu, K.-Y., Misra, V., Wacleche, V., Gosselin, A., Zhou, X. and Gabuzda, D. (2009) 'Transcriptional profiling reveals developmental relationship and distinct biological functions of CD16+ and CD16- monocyte subsets'. *BMC Genomics,* 10: (1): 403.

Anderson, A.C. (2014) 'Tim-3: An Emerging Target in the Cancer Immunotherapy Landscape'. *Cancer Immunology Research,* 2: (5): 393-398.

Angiolillo, A.L., Sgadari, C., Taub, D.D., Liao, F., Farber, J.M., Maheshwari, S., Kleinman, H.K., Reaman, G.H. and Tosato, G. (1995) 'Human interferon-inducible protein 10 is a potent inhibitor of angiogenesis in vivo'. *J Exp Med,* 182: (1): 155-162.

Apostolopoulos, J., Davenport, P. and Tipping, P.G. (1996) 'Interleukin-8 Production by Macrophages From Atheromatous Plaques'. *Arteriosclerosis, Thrombosis, and Vascular Biology,* 16: (8): 1007-1012.

Armaka, M., Gkretsi, V., Kontoyiannis, D. and Kollias, G. (2009) 'A standardized protocol for the isolation and culture of normal and arthritogenic murine synovial fibroblasts'.

Aruffo, A., Stamenkovic, I., Melnick, M., Underhill, C.B. and Seed, B. (1990) 'CD44 is the principal cell surface receptor for hyaluronate'. *Cell,* 61: (7): 1303-1313.

Assefnia, S., Dakshanamurthy, S., Auvil, J.M.G., Hampel, C., Anastasiadis, P.Z., Kallakury, B., Uren, A., Foley, D.W., Brown, M.L., Shapiro, L., Brenner, M., Haigh, D. and Byers, S.W. (2014) 'Cadherin-11 in poor prognosis malignancies and rheumatoid arthritis: common target, common therapies'. *Oncotarget,* 5: (6): 1458-1474.

Aubin, J.E. (1979) 'Autofluorescence of viable cultured mammalian cells'. *J Histochem Cytochem,* 27: (1): 36-43.

Augsten, M., Hägglöf, C., Olsson, E., Stolz, C., Tsagozis, P., Levchenko, T., Frederick, M.J., Borg, Å., Micke, P., Egevad, L. and Östman, A. (2009) 'CXCL14 is an autocrine growth factor for fibroblasts and acts as a multi-modal stimulator of prostate tumor growth'. *Proceedings of the National Academy of Sciences,* 106: (9): 3414-3419.

Baeten, D., Demetter, P., Cuvelier, C., Van Den Bosch, F., Kruithof, E., Van Damme, N., Verbruggen, G., Mielants, H., Veys, E. and De Keyser, F. (2000) 'Comparative study of the synovial histology in rheumatoid arthritis, spondyloarthropathy, and osteoarthritis: influence of disease duration and activity'. *Annals of the Rheumatic Diseases,* 59: (12): 945-953.

Baeten, D., Demetter, P., Cuvelier, C.A., Kruithof, E., Van Damme, N., De Vos, M., Veys, E.M. and De Keyser, F. (2002) 'Macrophages expressing the scavenger receptor CD163: a link between immune alterations of the gut and synovial inflammation in spondyloarthropathy'. *J Pathol,* 196: (3): 343-350.

Baeten, D., Jon Møller, H., Delanghe, J., Veys, E.M., Moestrup, S.K. and De Keyser, F. (2004) 'Association of CD163+ macrophages and local production of soluble CD163 with decreased lymphocyte activation in spondylarthropathy synovitis'. *Arthritis & Rheumatism,* 50: (5): 1611-1623.

Ballara, S., Taylor, P.C., Reusch, P., Marme, D., Feldmann, M., Maini, R.N. and Paleolog, E.M. (2001) 'Raised serum vascular endothelial growth factor levels are associated with destructive change in inflammatory arthritis'. *Arthritis Rheum,* 44: (9): 2055-2064.

Bandura, D.R., Baranov, V.I., Ornatsky, O.I., Antonov, A., Kinach, R., Lou, X., Pavlov, S., Vorobiev, S., Dick, J.E. and Tanner, S.D. (2009) 'Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry'. *Anal Chem,* 81: (16): 6813-6822.

Barland, P., Novikoff, A.B. and Hamerman, D. (1962) 'Electron microscopy of the human synovial membrane'. *J Cell Biol,* 14: 207-220.

Bauer, S., Jendro, M., Wadle, A., Kleber, S., Stenner, F., Dinser, R., Reich, A., Faccin, E., Godde, S., Dinges, H., Muller-Ladner, U. and Renner, C. (2006) 'Fibroblast activation protein is expressed by rheumatoid myofibroblast-like synoviocytes'. *Arthritis Research & Therapy,* 8: (6): R171.

Bayley, R., Kite, K.A., Mcgettrick, H.M., Smith, J.P., Kitas, G.D., Buckley, C.D. and Young, S.P. (2014) 'The autoimmune-associated genetic variant PTPN22 R620W enhances neutrophil activation and function in patients with rheumatoid arthritis and healthy individuals'. *Ann Rheum Dis*.

Begovich, A.B., Carlton, V.E.H., Honigberg, L.A., Schrodi, S.J., Chokkalingam, A.P., Alexander, H.C., Ardlie, K.G., Huang, Q., Smith, A.M., Spoerke, J.M., Conn, M.T., Chang, M., Chang, S.-Y.P., Saiki, R.K., Catanese, J.J., Leong, D.U., Garcia, V.E., Mcallister, L.B., Jeffery, D.A., Lee, A.T., Batliwalla, F., Remmers, E., Criswell, L.A., Seldin, M.F., Kastner, D.L., Amos, C.I., Sninsky, J.J. and Gregersen, P.K. (2004) 'A Missense Single-Nucleotide Polymorphism in a Gene Encoding a Protein Tyrosine Phosphatase (PTPN22) Is Associated with Rheumatoid Arthritis'. *The American Journal of Human Genetics,* 75: (2): 330-337.

Benito-Miguel, M., García-Carmona, Y., Balsa, A., Bautista-Caro, M.-B., Arroyo-Villa, I., Cobo-Ibáñez, T., Bonilla-Hernán, M.G., De Ayala, C.P., Sánchez-Mateos, P., Martín-Mola, E. and Miranda-Carús, M.-E. (2012) 'IL-15 Expression on RA Synovial Fibroblasts Promotes B Cell Survival'. *PLoS ONE,* 7: (7): e40620.

Benjamini, Y. and Hochberg, Y. (1995) 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing'. *Journal of the Royal Statistical Society. Series B (Methodological),* 57: (1): 289-300.

Bettelli, E., Carrier, Y., Gao, W., Korn, T., Strom, T.B., Oukka, M., Weiner, H.L. and Kuchroo, V.K. (2006) 'Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells'. *Nature,* 441: (7090): 235-238.

Bombardieri, M., Kam, N.W., Brentano, F., Choi, K., Filer, A., Kyburz, D., Mcinnes, I.B., Gay, S., Buckley, C. and Pitzalis, C. (2011) 'A BAFF/APRIL-dependent TLR3-stimulated pathway enhances the capacity of rheumatoid synovial fibroblasts to induce AID expression and Ig class-switching in B cells'. *Ann Rheum Dis,* 70: (10): 1857-1865.

Bottini, N. and Firestein, G.S. (2013) 'Duality of fibroblast-like synoviocytes in RA: passive responders and imprinted aggressors'. *Nat Rev Rheumatol,* 9: (1): 24-33.

Bradfield, P.F., Amft, N., Vernon-Wilson, E., Exley, A.E., Parsonage, G., Rainger, G.E., Nash, G.B., Thomas, A.M., Simmons, D.L., Salmon, M. and Buckley, C.D. (2003) 'Rheumatoid fibroblast-like synoviocytes overexpress the chemokine stromal cell-derived factor 1 (CXCL12), which supports distinct patterns and rates of CD4+ and CD8+ T cell migration within synovial tissue'. *Arthritis Rheum,* 48: (9): 2472-2482.

Buechler, C., Ritter, M., Orsó, E., Langmann, T., Klucken, J. and Schmitz, G. (2000) 'Regulation of scavenger receptor CD163 expression in human monocytes and macrophages by pro- and antiinflammatory stimuli'. *Journal of Leukocyte Biology,* 67: (1): 97-103.

Bukhari, M., Lunt, M., Harrison, B.J., Scott, D.G., Symmons, D.P. and Silman, A.J. (2002) 'Rheumatoid factor is the major predictor of increasing severity of radiographic erosions in rheumatoid arthritis: results from the Norfolk Arthritis Register Study, a large inception cohort'. *Arthritis Rheum,* 46: (4): 906-912.

Burger, J.A., Zvaifler, N.J., Tsukada, N., Firestein, G.S. and Kipps, T.J. (2001) 'Fibroblast-like synoviocytes support B-cell pseudoemperipolesis via a stromal cell-derived factor-1- and CD106 (VCAM-1)-dependent mechanism'. *J Clin Invest,* 107: (3): 305-315.

Burns, J.M., Summers, B.C., Wang, Y., Melikian, A., Berahovich, R., Miao, Z., Penfold, M.E., Sunshine, M.J., Littman, D.R., Kuo, C.J., Wei, K., Mcmaster, B.E., Wright, K., Howard, M.C. and Schall, T.J. (2006) 'A novel chemokine receptor for SDF-1 and I-TAC involved in cell survival, cell adhesion, and tumor development'. *J Exp Med,* 203: (9): 2201-2213.

Bustamante, M.F., Garcia-Carbonell, R., Whisenant, K.D. and Guma, M. (2017) 'Fibroblast-like synoviocyte metabolism in the pathogenesis of rheumatoid arthritis'. *Arthritis Research & Therapy,* 19: (1): 110.

Buwitt, U., Flohr, T. and Bottger, E.C. (1992) 'Molecular cloning and characterization of an interferon induced human cDNA with sequence homology to a mammalian peptide chain release factor'. *EMBO J,* 11: (2): 489-496.

Calabrese, C., Kirchner, E., Kontzias, K., Velcheti, V. and Calabrese, L.H. (2017) 'Rheumatic immune-related adverse events of checkpoint therapy for cancer: case series of a new nosological entity'. *RMD Open,* 3: (1).

Campbell, J.J., Hedrick, J., Zlotnik, A., Siani, M.A., Thompson, D.A. and Butcher, E.C. (1998) 'Chemokines and the Arrest of Lymphocytes Rolling Under Flow Conditions'. *Science,* 279: (5349): 381-384.

Campo, G.M., Avenoso, A., Campo, S., Angela, D., Ferlazzo, A.M. and Calatroni, A. (2006) 'TNF-alpha, IFN-gamma, and IL-1beta modulate hyaluronan synthase expression in human skin fibroblasts: synergistic effect by concomital treatment with FeSO4 plus ascorbate'. *Mol Cell Biochem,* 292: (1-2): 169-178.

Canton, J., Neculai, D. and Grinstein, S. (2013) 'Scavenger receptors in homeostasis and immunity'. *Nat Rev Immunol,* 13: (9): 621-634.

Castell, J.V., Gomez-Lechon, M.J., David, M., Hirano, T., Kishimoto, T. and Heinrich, P.C. (1988)

'Recombinant human interleukin-6 (IL-6/BSF-2/HSF) regulates the synthesis of acute phase proteins in human hepatocytes'. *FEBS Lett,* 232: (2): 347-350.

Chabaud, M., Durand, J.M., Buchs, N., Fossiez, F., Page, G., Frappart, L. and Miossec, P. (1999) 'Human interleukin-17: A T cell–derived proinflammatory cytokine produced by the rheumatoid synovium'. *Arthritis & Rheumatism,* 42: (5): 963-970.

Chan, A., Filer, A., Parsonage, G., Kollnberger, S., Gundle, R., Buckley, C.D. and Bowness, P. (2008) 'Mediation of the proinflammatory cytokine response in rheumatoid arthritis and spondylarthritis by interactions between fibroblast-like synoviocytes and natural killer cells'. *Arthritis Rheum,* 58: (3): 707-717.

Chan, M.M., Kefford, R.F., Carlino, M., Clements, A. and Manolios, N. (2015) 'Arthritis and tenosynovitis associated with the anti-PD1 antibody pembrolizumab in metastatic melanoma'. *J Immunother,* 38: (1): 37-39.

Chang, C.B., Han, S.A., Kim, E.M., Lee, S., Seong, S.C. and Lee, M.C. (2013) 'Chondrogenic potentials of human synovium-derived cells sorted by specific surface markers'. *Osteoarthritis and Cartilage,* 21: (1): 190-199.

Chang, H.Y., Chi, J.-T., Dudoit, S., Bondre, C., Rijn, M.V.D., Botstein, D. and Brown, P.O. (2002) 'Diversity, Topographic Differentiation, and Positional Memory in Human Fibroblasts'. *Proceedings of the National Academy of Sciences of the United States of America,* 99: (20): 12877-12882.

Chapuis, F., Rosenzwajg, M., Yagello, M., Ekman, M., Biberfeld, P. and Gluckman, J.C. (1997) 'Differentiation of human dendritic cells from monocytes in vitro'. *Eur J Immunol,* 27: (2): 431-441.

Chen, V., Croft, D., Purkis, P. and Kramer, I.M. (1998) 'Co-culture of synovial fibroblasts and differentiated U937 cells is sufficient for high interleukin-6 but not interleukin-1beta or tumour necrosis factor-alpha release'. *Rheumatology,* 37: (2): 148-156.

Chen, Y., Wang, S., Yi, Z., Tian, H., Aliyari, R., Li, Y., Chen, G., Liu, P., Zhong, J., Chen, X., Du, P., Su, L., Qin, F.X., Deng, H. and Cheng, G. (2014) 'Interferon-inducible cholesterol-25-hydroxylase inhibits hepatitis C virus replication via distinct mechanisms'. *Sci Rep,* 4: 7242.

Chen, Y.F., Jobanputra, P., Barton, P., Jowett, S., Bryan, S., Clark, W., Fry-Smith, A. and Burls, A. (2006) 'A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness'. *Health Technol Assess,* 10: (42): iii-iv, xi-xiii, 1-229.

Chiu, R., Boyle, W.J., Meek, J., Smeal, T., Hunter, T. and Karin, M. (1988) 'The c-Fos protein interacts with c-Jun/AP-1 to stimulate transcription of AP-1 responsive genes'. *Cell,* 54: (4): 541-552.

Choi, I.Y., Karpus, O.N., Turner, J.D., Hardie, D., Marshall, J.L., De Hair, M.J.H., Maijer, K.I., Tak, P.P., Raza, K., Hamann, J., Buckley, C.D., Gerlag, D.M. and Filer, A. (2017) 'Stromal cell markers are differentially expressed in the synovial tissue of patients with early arthritis'. *PLoS ONE,* 12: (8): e0182751.

Chomarat, P., Banchereau, J., Davoust, J. and Palucka, A.K. (2000) 'IL-6 switches the

differentiation of monocytes from dendritic cells to macrophages'. *Nat Immunol,* 1: (6): 510-514.

Chomarat, P., Rissoan, M.C., Pin, J.J., Banchereau, J. and Miossec, P. (1995) 'Contribution of IL-1, CD14, and CD13 in the increased IL-6 production induced by in vitro monocyte-synoviocyte interactions'. *J Immunol,* 155: (7): 3645-3652.

Choy, E.H., Isenberg, D.A., Garrood, T., Farrow, S., Ioannou, Y., Bird, H., Cheung, N., Williams, B., Hazleman, B., Price, R., Yoshizaki, K., Nishimoto, N., Kishimoto, T. and Panayi, G.S. (2002) 'Therapeutic benefit of blocking interleukin-6 activity with an anti-interleukin-6 receptor monoclonal antibody in rheumatoid arthritis: a randomized, double-blind, placebo-controlled, dose-escalation trial'. *Arthritis Rheum,* 46: (12): 3143-3150.

Ciechomska, M., O'reilly, S., Przyborski, S., Oakley, F., Bogunia-Kubik, K. and Van Laar, J.M. (2016) 'Histone Demethylation and Toll-like Receptor 8–Dependent Cross-Talk in Monocytes Promotes Transdifferentiation of Fibroblasts in Systemic Sclerosis Via Fra-2'. *Arthritis & Rheumatology,* 68: (6): 1493-1504.

Cleveland, W.S. (1979) 'Robust Locally Weighted Regression and Smoothing Scatterplots'. *Journal of the American Statistical Association,* 74: (368): 829-836.

Cole, K.E., Strick, C.A., Paradis, T.J., Ogborne, K.T., Loetscher, M., Gladue, R.P., Lin, W., Boyd, J.G., Moser, B., Wood, D.E., Sahagan, B.G. and Neote, K. (1998) 'Interferon-inducible T cell alpha chemoattractant (I-TAC): a novel non-ELR CXC chemokine with potent activity on activated T cells through selective high affinity binding to CXCR3'. *J Exp Med,* 187: (12): 2009-2021.

Coller, J.A. (2009) "Microarray Platforms and Aspects of Experimental Variation". Batch Effects and Noise in Microarray Experiments. John Wiley & Sons, Ltd 5-17.

Cooper, D.L., Martin, S.G., Robinson, J.I., Mackie, S.L., Charles, C.J., Nam, J., Isaacs, J.D., Emery, P. and Morgan, A.W. (2012) 'FcgammaRIIIa expression on monocytes in rheumatoid arthritis: role in immune-complex stimulated TNF production and non-response to methotrexate therapy'. *PLoS ONE,* 7: (1): e28918.

Crocker, P.R. and Gordon, S. (1986) 'Properties and distribution of a lectin-like hemagglutinin differentially expressed by murine stromal tissue macrophages'. *J Exp Med,* 164: (6): 1862-1875.

Croft, A.P., Naylor, A.J., Marshall, J.L., Hardie, D.L., Zimmermann, B., Turner, J., Desanti, G., Adams, H., Yemm, A.I. and Müller-Ladner, U. (2016) 'Rheumatoid synovial fibroblasts differentiate into distinct subsets in the presence of cytokines and cartilage'. *Arthritis Research & Therapy,* 18: (1): 270.

Dayer, J.M., Krane, S.M., Russell, R.G. and Robinson, D.R. (1976) 'Production of collagenase and prostaglandins by isolated adherent rheumatoid synovial cells'. *Proceedings of the National Academy of Sciences of the United States of America,* 73: (3): 945-949.

De Ceulaer, C., Papazoglou, S. and Whaley, K. (1980) 'Increased biosynthesis of complement components by cultured monocytes, synovial fluid macrophages and skynovial membrane cells from patients with rheumatoid arthritis'. *Immunology,* 41: (1): 37-43.

De Hair, M.J.H., Harty, L.C., Gerlag, D.M., Pitzalis, C., Veale, D.J. and Tak, P.P. (2011) 'Synovial Tissue Analysis for the Discovery of Diagnostic and Prognostic Biomarkers in Patients with Early Arthritis'. *The Journal of Rheumatology,* 38: (9): 2068-2072.

De La Motte, C.A., Hascall, V.C., Calabro, A., Yen-Lieberman, B. and Strong, S.A. (1999) 'Mononuclear Leukocytes Preferentially Bind via CD44 to Hyaluronan on Human Intestinal Mucosal Smooth Muscle Cells after Virus Infection or Treatment with Poly(I·C)'. *Journal of Biological Chemistry,* 274: (43): 30747-30755.

De La Puerta, M.L., Trinidad, A.G., Rodríguez, M.D.C., De Pereda, J.M., Sánchez Crespo, M., Bayón, Y. and Alonso, A. (2013) 'The Autoimmunity Risk Variant LYP-W620 Cooperates with CSK in the Regulation of TCR Signaling'. *PLoS ONE,* 8: (1): e54569.

De La Rica, L., Urquiza, J.M., Gómez-Cabrero, D., Islam, A.B.M.M.K., López-Bigas, N., Tegnér, J., Toes, R.E.M. and Ballestar, E. (2013) 'Identification of novel markers in rheumatoid arthritis through integrated analysis of DNA methylation and microRNA expression'. *Journal of Autoimmunity,* 41: (0): 6-16.

De Waal Malefyt, R., Abrams, J., Bennett, B., Figdor, C.G. and De Vries, J.E. (1991) 'Interleukin 10(IL-10) inhibits cytokine synthesis by human monocytes: an autoregulatory role of IL-10 produced by monocytes'. *J Exp Med,* 174: (5): 1209-1220.

Deighton, C.M., Walker, D.J., Griffiths, I.D. and Roberts, D.F. (1989) 'The contribution of H LA to rheumatoid arthritis'. *Clinical Genetics,* 36: (3): 178-182.

Del Rey, M.J., Faré, R., Izquierdo, E., Usategui, A., Rodríguez-Fernández, J.L., Suárez-Fueyo, A., Cañete, J.D. and Pablos, J.L. (2014) 'Clinicopathological Correlations of Podoplanin (gp38) Expression in Rheumatoid Synovium and Its Potential Contribution to Fibroblast Platelet Crosstalk'. *PLoS ONE,* 9: (6): e99607.

Dennis, G., Holweg, C.T., Kummerfeld, S.K., Choy, D.F., Setiadi, A.F., Hackney, J.A., Haverty, P.M., Gilbert, H., Lin, W.Y., Diehl, L., Fischer, S., Song, A., Musselman, D., Klearman, M., Gabay, C., Kavanaugh, A., Endres, J., Fox, D.A., Martin, F. and Townsend, M.J. (2014) 'Synovial phenotypes in rheumatoid arthritis correlate with response to biologic therapeutics'. *Arthritis Research & Therapy,* 16: (2): R90.

Dieu, M.-C., Vanbervliet, B., Vicari, A., Bridon, J.-M., Oldham, E., Aït-Yahia, S., Brière, F., Zlotnik, A., Lebecque, S. and Caux, C. (1998) 'Selective Recruitment of Immature and Mature Dendritic Cells by Distinct Chemokines Expressed in Different Anatomic Sites'. *The Journal of Experimental Medicine,* 188: (2): 373-386.

Dolhain, R.J., Ter Haar, N.T., Hoefakker, S., Tak, P.P., De Ley, M., Claassen, E., Breedveld, F.C. and Miltenburg, A.M. (1996) 'Increased expression of interferon (IFN)-gamma together with IFN-gamma receptor in the rheumatoid synovial membrane compared with synovium of patients with osteoarthritis'. *Br J Rheumatol,* 35: (1): 24-32.

Donlin, L.T., Jayatilleke, A., Giannopoulou, E.G., Kalliolias, G.D. and Ivashkiv, L.B. (2014) 'Modulation of TNF-induced macrophage polarization by synovial fibroblasts'. *J Immunol,* 193: (5): 2373-2383.

Dooley, S., Herlitzka, I., Hanselmann, R., Ermis, A., Henn, W., Remberger, K., Hopf, T. and Welter, C. (1996) 'Constitutive expression of c-fos and c-jun, overexpression of ets-2, and reduced expression of metastasis suppressor gene nm23-H1 in rheumatoid arthritis'. *Annals of the Rheumatic Diseases,* 55: (5): 298-304.

Dougados, M., Soubrier, M., Antunez, A., Balint, P., Balsa, A., Buch, M., Casado, G., Detert, J., El-Zorkany, B., Emery, P., Hajjaj-Hassouni, N., Harigai, M., Luo, S.-F., Kurucz, R., Maciel, G., Mola, E.M., Montecucco, C.M., Mcinnes, I., Radner, H., Smolen, J., Song, Y.-W., Vonkeman, H.E., Winthrop, K. and Kay, J. (2013) 'Prevalence of comorbidities in rheumatoid arthritis and evaluation of their monitoring: results of an international, cross-sectional study (COMORA)'. *Annals of the Rheumatic Diseases*.

Dufour, J.H., Dziejman, M., Liu, M.T., Leung, J.H., Lane, T.E. and Luster, A.D. (2002a) 'IFN-gamma-inducible protein 10 (IP-10; CXCL10)-deficient mice reveal a role for IP-10 in effector T cell generation and trafficking'. *J Immunol,* 168: (7): 3195-3204.

Dufour, J.H., Dziejman, M., Liu, M.T., Leung, J.H., Lane, T.E. and Luster, A.D. (2002b) 'IFN-γ-Inducible Protein 10 (IP-10; CXCL10)-Deficient Mice Reveal a Role for IP-10 in Effector T Cell Generation and Trafficking'. *The Journal of Immunology,* 168: (7): 3195-3204.

Durand, M.K., Bodker, J.S., Christensen, A., Dupont, D.M., Hansen, M., Jensen, J.K., Kjelgaard, S., Mathiasen, L., Pedersen, K.E., Skeldal, S., Wind, T. and Andreasen, P.A. (2004) 'Plasminogen activator inhibitor-I and tumour growth, invasion, and metastasis'. *Thromb Haemost,* 91: (3): 438-449.

Edfors, F., Danielsson, F., Hallström, B.M., Käll, L., Lundberg, E., Pontén, F., Forsström, B. and Uhlén, M. (2016) 'Gene-specific correlation of RNA and protein levels in human cells and tissues'. *Mol Syst Biol,* 12: (10).

Ehrenstein, M.R., Evans, J.G., Singh, A., Moore, S., Warnes, G., Isenberg, D.A. and Mauri, C. (2004) 'Compromised function of regulatory T cells in rheumatoid arthritis and reversal by anti-TNFalpha therapy'. *J Exp Med,* 200: (3): 277-285.

Ekwall, A.-K., Eisler, T., Anderberg, C., Jin, C., Karlsson, N., Brisslert, M. and Bokarewa, M. (2011) 'The tumour-associated glycoprotein podoplanin is expressed in fibroblast-like synoviocytes of the hyperplastic synovial lining layer in rheumatoid arthritis'. *Arthritis Research & Therapy,* 13: (2): R40.

Elomaa, O., Sankala, M., Pikkarainen, T., Bergmann, U., Tuuttila, A., Raatikainen-Ahokas, A., Sariola, H. and Tryggvason, K. (1998) 'Structure of the human macrophage MARCO receptor and characterization of its bacteria-binding region'. *J Biol Chem,* 273: (8): 4530-4538.

Epelman, S., Lavine, K.J., Beaudin, A.E., Sojka, D.K., Carrero, J.A., Calderon, B., Brija, T., Gautier, E.L., Ivanov, S., Satpathy, A.T., Schilling, J.D., Schwendener, R., Sergin, I., Razani, B., Forsberg, E.C., Yokoyama, W.M., Unanue, E.R., Colonna, M., Randolph, G.J. and Mann, D.L. (2014) 'Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation'. *Immunity,* 40: (1): 91-104.

Fahy, N., De Vries-Van Melle, M.L., Lehmann, J., Wei, W., Grotenhuis, N., Farrell, E., Van Der Kraan, P.M., Murphy, J.M., Bastiaansen-Jenniskens, Y.M. and Van Osch, G.J.V.M. (2014) 'Human osteoarthritic synovium impacts chondrogenic differentiation of mesenchymal stem cells via macrophage polarisation state'. *Osteoarthritis and Cartilage,* 22: (8): 1167-1175.

Fassbender, H.G. and Simmling-Annefeld, M. (1983) 'The potential aggressiveness of synovial tissue in rheumatoid arthritis'. *J Pathol,* 139: (3): 399-406.

Fava, R.A., Olsen, N.J., Spencer-Green, G., Yeo, K.T., Yeo, T.K., Berse, B., Jackman, R.W., Senger,

D.R., Dvorak, H.F. and Brown, L.F. (1994) 'Vascular permeability factor/endothelial growth factor (VPF/VEGF): accumulation and expression in human synovial fluids and rheumatoid synovial tissue'. *J Exp Med,* 180: (1): 341-346.

Filer, A. (2013) 'The fibroblast as a therapeutic target in rheumatoid arthritis'. *Current Opinion in Pharmacology,* 13: (3): 413-419.

Filer, A., Antczak, P., Parsonage, G.N., Legault, H.M., O'toole, M., Pearson, M.J., Thomas, A.M., Scheel-Toellner, D., Raza, K., Buckley, C.D. and Falciani, F. (2015) 'Stromal transcriptional profiles reveal hierarchies of anatomical site, serum response and disease and identify disease specific pathways'. *PLoS ONE,* 10: (3): e0120917.

Filer, A., Parsonage, G., Smith, E., Osborne, C., Thomas, A.M., Curnow, S.J., Rainger, G.E., Raza, K., Nash, G.B., Lord, J., Salmon, M. and Buckley, C.D. (2006) 'Differential survival of leukocyte subsets mediated by synovial, bone marrow, and skin fibroblasts: site-specific versus activation-dependent survival of T cells and neutrophils'. *Arthritis Rheum,* 54: (7): 2096-2108.

Filer, A., Ward, L.S.C., Kemble, S., Davies, C.S., Munir, H., Rogers, R., Raza, K., Buckley, C.D., Nash, G.B. and Mcgettrick, H.M. (2017) 'Identification of a transitional fibroblast function in very early rheumatoid arthritis'. *Annals of the Rheumatic Diseases*.

Fleckner, J., Rasmussen, H.H. and Justesen, J. (1991) 'Human interferon gamma potently induces the synthesis of a 55-kDa protein (gamma 2) highly homologous to rabbit peptide chain release factor and bovine tryptophanyl-tRNA synthetase'. *Proceedings of the National Academy of Sciences of the United States of America,* 88: (24): 11520-11524.

Fleetwood, A.J., Dinh, H., Cook, A.D., Hertzog, P.J. and Hamilton, J.A. (2009) 'GM-CSF- and M-CSF-dependent macrophage phenotypes display differential dependence on Type I interferon signaling'. *Journal of Leukocyte Biology,* 86: (2): 411-421.

Fleetwood, A.J., Lawrence, T., Hamilton, J.A. and Cook, A.D. (2007) 'Granulocyte-Macrophage Colony-Stimulating Factor (CSF) and Macrophage CSF-Dependent Macrophage Phenotypes Display Differences in Cytokine Profiles and Transcription Factor Activities: Implications for CSF Blockade in Inflammation'. *The Journal of Immunology,* 178: (8): 5245-5252.

Flores-Borja, F., Jury, E.C., Mauri, C. and Ehrenstein, M.R. (2008) 'Defects in CTLA-4 are associated with abnormal regulatory T cell function in rheumatoid arthritis'. *Proceedings of the National Academy of Sciences,* 105: (49): 19396-19401.

Fonseca, J.E., Edwards, J.C., Blades, S. and Goulding, N.J. (2002) 'Macrophage subpopulations in rheumatoid synovium: reduced CD163 expression in CD4+ T lymphocyte-rich microenvironments'. *Arthritis Rheum,* 46: (5): 1210-1216.

Frank-Bertoncelj, M., Trenkmann, M., Klein, K., Karouzakis, E., Rehrauer, H., Bratus, A., Kolling, C., Armaka, M., Filer, A., Michel, B.A., Gay, R.E., Buckley, C.D., Kollias, G., Gay, S. and Ospelt, C. (2017) 'Epigenetically-driven anatomical diversity of synovial fibroblasts guides joint-specific fibroblast functions'. *Nat Commun,* 8: 14852.

Ganster, R.W., Taylor, B.S., Shao, L. and Geller, D.A. (2001) 'Complex regulation of human inducible nitric oxide synthase gene transcription by Stat 1 and NF-kappa B'. *Proc Natl Acad Sci U S A,* 98: (15): 8638-8643.

Garcia-Carbonell, R., Divakaruni, A.S., Lodi, A., Vicente-Suarez, I., Saha, A., Cheroutre, H., Boss, G.R., Tiziani, S., Murphy, A.N. and Guma, M. (2016) 'Critical Role of Glucose Metabolism in Rheumatoid Arthritis Fibroblast-like Synoviocytes'. *Arthritis Rheumatol,* 68: (7): 1614-1626.

Gautier, E.L., Shay, T., Miller, J., Greter, M., Jakubzick, C., Ivanov, S., Helft, J., Chow, A., Elpek, K.G., Gordonov, S., Mazloom, A.R., Ma'ayan, A., Chua, W.-J., Hansen, T.H., Turley, S.J., Merad, M. and Randolph, G.J. (2012) 'Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages'. *Nat Immunol,* 13: (11): 1118-1128.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A.J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J.Y. and Zhang, J. (2004) 'Bioconductor: open software development for computational biology and bioinformatics'. *Genome Biol,* 5: (10): R80.

Gerlag, D.M., Haringman, J.J., Smeets, T.J., Zwinderman, A.H., Kraan, M.C., Laud, P.J., Morgan, S., Nash, A.F. and Tak, P.P. (2004) 'Effects of oral prednisolone on biomarkers in synovial tissue and clinical improvement in rheumatoid arthritis'. *Arthritis Rheum,* 50: (12): 3783-3791.

Ghosh, A.K. and Vaughan, D.E. (2012) 'PAI-1 in tissue fibrosis'. *J Cell Physiol,* 227: (2): 493-507.

Gibofsky, A., Winchester, R.J., Patarroyo, M., Fotino, M. and Kunkel, H.G. (1978) 'Disease associations of the Ia-like human alloantigens. Contrasting patterns in rheumatoid arthritis and systemic lupus erythematosus'. *J Exp Med,* 148: (6): 1728-1732.

Ginhoux, F., Greter, M., Leboeuf, M., Nandi, S., See, P., Gokhan, S., Mehler, M.F., Conway, S.J., Ng, L.G., Stanley, E.R., Samokhvalov, I.M. and Merad, M. (2010) 'Fate mapping analysis reveals that adult microglia derive from primitive macrophages'. *Science,* 330: (6005): 841-845.

Gioud-Paquet, M., Auvinet, M., Raffin, T., Girard, P., Bouvier, M., Lejeune, E. and Monier, J.C. (1987) 'IgM rheumatoid factor (RF), IgA RF, IgE RF, and IgG RF detected by ELISA in rheumatoid arthritis'. *Ann Rheum Dis,* 46: (1): 65-71.

Gitter, B.D., Labus, J.M., Lees, S.L. and Scheetz, M.E. (1989) 'Characteristics of human synovial fibroblast activation by IL-1 beta and TNF alpha'. *Immunology,* 66: (2): 196-200.

Goldbach-Mansky, R., Lee, J.M., Hoxworth, J.M., Smith, D., Duray, P., Schumacher, H.R., Yarboro, C.H., Klippel, J., Kleiner, D. and El-Gabalawy, H.S. (2000) 'Active synovial matrix metalloproteinase-2 is associated with radiographic erosions in patients with early synovitis'. *Arthritis Research & Therapy,* 2: (2): 145.

Gonzalo-Gil, E., Criado, G., Santiago, B., Dotor, J., Pablos, J.L. and Galindo, M. (2013) 'Transforming growth factor (TGF)-beta signalling is increased in rheumatoid synovium but TGF-beta blockade does not modify experimental arthritis'. *Clin Exp Immunol,* 174: (2): 245-255.

Gosselin, D., Link, V.M., Romanoski, Casey e., Fonseca, Gregory j., Eichenfield, Dawn z., Spann, Nathanael j., Stender, Joshua d., Chun, Hyun b., Garner, H., Geissmann, F. and Glass, Christopher k. (2014) 'Environment Drives Selection and Function of Enhancers Controlling Tissue-Specific Macrophage Identities'. *Cell,* 159: (6): 1327-1340.

Grant, E.P., Picarella, D., Burwell, T., Delaney, T., Croci, A., Avitahl, N., Humbles, A.A., Gutierrez-

Ramos, J.-C., Briskin, M., Gerard, C. and Coyle, A.J. (2002) 'Essential Role for the C5a Receptor in Regulating the Effector Phase of Synovial Infiltration and Joint Destruction in Experimental Arthritis'. *The Journal of Experimental Medicine,* 196: (11): 1461-1471.

Greenberg, J.D., Kishimoto, M., Strand, V., Cohen, S.B., Olenginski, T.P., Harrington, T., Kafka, S.P., Reed, G. and Kremer, J.M. (2008) 'Tumor necrosis factor antagonist responsiveness in a United States rheumatoid arthritis cohort'. *Am J Med,* 121: (6): 532-538.

Gregersen, P.K., Silver, J. and Winchester, R.J. (1987) 'The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis'. *Arthritis Rheum,* 30: (11): 1205-1213.

Greisen, S.R., Moller, H.J., Stengaard-Pedersen, K., Hetland, M.L., Horslev-Petersen, K., Jorgensen, A., Hvid, M. and Deleuran, B. (2011) 'Soluble macrophage-derived CD163 is a marker of disease activity and progression in early rheumatoid arthritis'. *Clinical and experimental rheumatology,* 29: (4): 689-692.

Greisen, S.R., Moller, H.J., Stengaard-Pedersen, K., Hetland, M.L., Horslev-Petersen, K., Junker, P., Ostergaard, M., Hvid, M. and Deleuran, B. (2015) 'Macrophage activity assessed by soluble CD163 in early rheumatoid arthritis: association with disease activity but different response patterns to synthetic and biologic DMARDs'. *Clinical and experimental rheumatology*.

Guilliams, M., Dutertre, C.-A., Scott, Charlotte l., Mcgovern, N., Sichien, D., Chakarov, S., Van gassen, S., Chen, J., Poidinger, M., De prijck, S., Tavernier, Simon j., Low, I., Irac, Sergio e., Mattar, Citra n., Sumatoh, Hermi r., Low, Gillian hui l., Chung, Tam john k., Chan, Dedrick kok h., Tan, Ker k., Hon, Tony lim k., Fossum, E., Bogen, B., Choolani, M., Chan, Jerry kok y., Larbi, A., Luche, H., Henri, S., Saeys, Y., Newell, Evan w., Lambrecht, Bart n., Malissen, B. and Ginhoux, F. (2016) 'Unsupervised High-Dimensional Analysis Aligns Dendritic Cells across Tissues and Species'. *Immunity,* 45: (3): 669-684.

Gullo, F. and De Bari, C. (2013) 'Prospective purification of a subpopulation of human synovial mesenchymal stem cells with enhanced chondro-osteogenic potency'. *Rheumatology (Oxford),* 52: (10): 1758-1768.

Hahn, G., Stuhlmuller, B., Hain, N., Kalden, J.R., Pfizenmaier, K. and Burmester, G.R. (1993) 'Modulation of monocyte activation in patients with rheumatoid arthritis by leukapheresis therapy'. *J Clin Invest,* 91: (3): 862-870.

Hao, L., Zhu, G., Lu, Y., Wang, M., Jules, J., Zhou, X. and Chen, W. (2015) 'Deficiency of cathepsin K prevents inflammation and bone erosion in rheumatoid arthritis and periodontitis and reveals its shared osteoimmune role'. *FEBS Lett,* 589: (12): 1331-1339.

Harada, M., Mitsuyama, K., Yoshida, H., Sakisaka, S., Taniguchi, E., Kawaguchi, T., Ariyoshi, M., Saiki, T., Sakamoto, M., Nagata, K., Sata, M., Matsuo, K. and Tanikawa, K. (1998) 'Vascular endothelial growth factor in patients with rheumatoid arthritis'. *Scand J Rheumatol,* 27: (5): 377-380.

Hardy, R., Hulso, C., Liu, Y., Gasparini, S., Fong-Yee, C., Tu, J., Stoner, S., Stewart, P., Raza, K., Cooper, M., Seibel, M. and Zhou, H. (2013) 'Characterisation of fibroblast-like synoviocytes from a murine model of joint inflammation'. *Arthritis Research & Therapy,* 15: (1): R24.

Haringman, J.J., Gerlag, D.M., Zwinderman, A.H., Smeets, T.J., Kraan, M.C., Baeten, D., Mcinnes,

I.B., Bresnihan, B. and Tak, P.P. (2005) 'Synovial tissue macrophages: a sensitive biomarker for response to treatment in patients with rheumatoid arthritis'. *Ann Rheum Dis,* 64: (6): 834-838.

Harre, U., Georgess, D., Bang, H., Bozec, A., Axmann, R., Ossipova, E., Jakobsson, P.-J., Baum, W., Nimmerjahn, F., Szarka, E., Sarmay, G., Krumbholz, G., Neumann, E., Toes, R., Scherer, H.-U., Catrina, A.I., Klareskog, L., Jurdic, P. and Schett, G. (2012) 'Induction of osteoclastogenesis and bone loss by human autoantibodies against citrullinated vimentin'. *The Journal of Clinical Investigation,* 122: (5): 1791-1802.

Hartnell, A., Steel, J., Turley, H., Jones, M., Jackson, D.G. and Crocker, P.R. (2001) 'Characterization of human sialoadhesin, a sialic acid binding receptor expressed by resident and inflammatory macrophage populations'. *Blood,* 97: (1): 288-296.

Hartung, H.P. and Hadding, U. (1983) 'Synthesis of complement by macrophages and modulation of their functions through complement activation'. *Springer Seminars in Immunopathology,* 6: (4): 283-326.

Hashimoto, D., Chow, A., Noizat, C., Teo, P., Beasley, Mary b., Leboeuf, M., Becker, Christian d., See, P., Price, J., Lucas, D., Greter, M., Mortha, A., Boyer, Scott w., Forsberg, E.C., Tanaka, M., Van rooijen, N., García-Sastre, A., Stanley, E.R., Ginhoux, F., Frenette, Paul s. and Merad, M. (2013) 'Tissue-Resident Macrophages Self-Maintain Locally throughout Adult Life with Minimal Contribution from Circulating Monocytes'. *Immunity,* 38: (4): 792-804.

Hayashida, K., Nanki, T., Girschick, H., Yavuz, S., Ochi, T. and Lipsky, P. (2001) 'Synovial stromal cells from rheumatoid arthritis patients attract monocytes by producing MCP-1 and IL-8'. *Arthritis Res,* 3: (2): 1-9.

He, L. and Hannon, G.J. (2004) 'MicroRNAs: small RNAs with a big role in gene regulation'. *Nat Rev Genet,* 5: (7): 522-531.

Herzing, L.B., Romer, J.T., Horn, J.M. and Ashworth, A. (1997) 'Xist has properties of the X-chromosome inactivation centre'. *Nature,* 386: (6622): 272-275.

Hidalgo, L.G., Einecke, G., Allanach, K. and Halloran, P.F. (2008) 'The Transcriptome of Human Cytotoxic T Cells: Similarities and Disparities Among Allostimulated CD4+ CTL, CD8+ CTL and NK cells'. *American Journal of Transplantation,* 8: (3): 627-636.

Hill, J.A., Southwood, S., Sette, A., Jevnikar, A.M., Bell, D.A. and Cairns, E. (2003) 'Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule'. *J Immunol,* 171: (2): 538-541.

Hjelmstrom, P., Fjell, J., Nakagawa, T., Sacca, R., Cuff, C.A. and Ruddle, N.H. (2000) 'Lymphoid tissue homing chemokines are expressed in chronic inflammation'. *Am J Pathol,* 156: (4): 1133-1138.

Holt, D.J., Chamberlain, L.M. and Grainger, D.W. (2010) 'Cell-cell signaling in co-cultures of macrophages and fibroblasts'. *Biomaterials,* 31: (36): 9382-9394.

Holt, R.A. and Jones, S.J.M. (2008) 'The new paradigm of flow cell sequencing'. *Genome Research,* 18: (6): 839-846.

Hoppe, B., Haupl, T., Gruber, R., Kiesewetter, H., Burmester, G.R., Salama, A. and Dorner, T. (2006) 'Detailed analysis of the variability of peptidylarginine deiminase type 4 in German patients with rheumatoid arthritis: a case-control study'. *Arthritis Res Ther,* 8: (2): R34.

Horiuchi, M., Morinobu, A., Chin, T., Sakai, Y., Kurosaka, M. and Kumagai, S. (2009) 'Expression and function of histone deacetylases in rheumatoid arthritis synovial fibroblasts'. *J Rheumatol,* 36: 1580 - 1589.

Hotelling, H. (1933) 'Analysis of a complex of statistical variables into principal components'. *Journal of Educational Psychology,* 24: (7): 498-520.

Hou, N., Zou, Y., Piao, X., Liu, S., Wang, L., Li, S. and Chen, Q. (2016) 'T-Cell Immunoglobulin- and Mucin-Domain-Containing Molecule 3 Signaling Blockade Improves Cell-Mediated Immunity Against Malaria'. *J Infect Dis,* 214: (10): 1547-1556.

Hrdlickova, R., Toloue, M. and Tian, B. (2017) 'RNA-Seq methods for transcriptome analysis'. *Wiley Interdisciplinary Reviews: RNA,* 8: (1): e1364-n/a.

Huang, W., Mabrouk, M.E., Sylvester, J., Dehnade, F. and Zafarullah, M. (2011) 'Enhanced expression of tissue inhibitor of metalloproteinases-4 gene in human osteoarthritic synovial membranes and its differential regulation by cytokines in chondrocytes'. *Open Rheumatol J,* 5: 81-87.

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

Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K.D., Irizarry, R.A., Lawrence, M., Love, M.I., Macdonald, J., Obenchain, V., Oles, A.K., Pages, H., Reyes, A., Shannon, P., Smyth, G.K., Tenenbaum, D., Waldron, L. and Morgan, M. (2015) 'Orchestrating high-throughput genomic analysis with Bioconductor'. *Nat Meth,* 12: (2): 115-121.

Huizinga, T., Amos, C., Van Der Helm-Van Mil, A., Chen, W., Van Gaalen, F., Jawaheer, D., Schreuder, G., Wener, M., Breedveld, F. and Ahmad, N. (2005) 'Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins'. *Arthritis Rheum,* 52: 3433 - 3438.

Huynh, P.T., Beswick, E.J., Coronado, Y.A., Johnson, P., O'connell, M.R., Watts, T., Singh, P., Qiu, S., Morris, K., Powell, D.W. and Pinchuk, I.V. (2016) 'CD90(+) stromal cells are the major source of IL-6, which supports cancer stem-like cells and inflammation in colorectal cancer'. *Int J Cancer,* 138: (8): 1971-1981.

Hwang, S.Y., Kim, J.Y., Kim, K.W., Park, M.K., Moon, Y., Kim, W.U. and Kim, H.Y. (2004) 'IL-17 induces production of IL-6 and IL-8 in rheumatoid arthritis synovial fibroblasts via NF-kappaB- and PI3-kinase/Akt-dependent pathways'. *Arthritis Res Ther,* 6: (2): R120-128.

Hyrich, K.L., Watson, K.D., Silman, A.J. and Symmons, D.P. (2006) 'Predictors of response to anti-TNF-alpha therapy among patients with rheumatoid arthritis: results from the British Society for Rheumatology Biologics Register'. *Rheumatology (Oxford),* 45: (12): 1558-1565.

Inazuka, M., Tahira, T., Horiuchi, T., Harashima, S., Sawabe, T., Kondo, M., Miyahara, H. and

Hayashi, K. (2000) 'Analysis of p53 tumour suppressor gene somatic mutations in rheumatoid arthritis synovium'. *Rheumatology,* 39: (3): 262-266.

Ioan-Facsinay, A., El-Bannoudi, H., Scherer, H.U., Van Der Woude, D., Menard, H.A., Lora, M., Trouw, L.A., Huizinga, T.W. and Toes, R.E. (2011) 'Anti-cyclic citrullinated peptide antibodies are a collection of anti-citrullinated protein antibodies and contain overlapping and non-overlapping reactivities'. *Ann Rheum Dis,* 70: (1): 188-193.

Isomäki, P., Junttila, I., Vidqvist, K.-L., Korpela, M. and Silvennoinen, O. (2015) 'The activity of JAK-STAT pathways in rheumatoid arthritis: constitutive activation of STAT3 correlates with interleukin 6 levels'. *Rheumatology,* 54: (6): 1103-1113.

Isozaki, T., Amin, M.A., Ruth, J.H., Campbell, P.L., Tsou, P.-S., Ha, C.M., Stinson, W.A., Domino, S.E. and Koch, A.E. (2014a) 'Fucosyltransferase 1 Mediates Angiogenesis in Rheumatoid Arthritis'. *Arthritis & Rheumatology,* 66: (8): 2047-2058.

Isozaki, T., Ruth, J.H., Amin, M.A., Campbell, P.L., Tsou, P.-S., Ha, C.M., Haines, G.K., Edhayan, G. and Koch, A.E. (2014b) 'Fucosyltransferase 1 mediates angiogenesis, cell adhesion and rheumatoid arthritis synovial tissue fibroblast proliferation'. *Arthritis Research & Therapy,* 16: (1): R28.

Jablonski, K.A., Amici, S.A., Webb, L.M., Ruiz-Rosado, J.D.D., Popovich, P.G., Partida-Sanchez, S. and Guerau-De-Arellano, M. (2015) 'Novel Markers to Delineate Murine M1 and M2 Macrophages'. *PLoS ONE,* 10: (12): e0145342.

Jaguin, M., Houlbert, N., Fardel, O. and Lecureur, V. (2013) 'Polarization profiles of human M-CSF-generated macrophages and comparison of M1-markers in classically activated macrophages from GM-CSF and M-CSF origin'. *Cell Immunol,* 281: (1): 51-61.

James, E.A., Moustakas, A.K., Bui, J., Papadopoulos, G.K., Bondinas, G., Buckner, J.H. and Kwok, W.W. (2010) 'HLA-DR1001 presents "altered-self" peptides derived from joint-associated proteins by accepting citrulline in three of its binding pockets'. *Arthritis Rheum,* 62: (10): 2909-2918.

Janusz, M.J. and Hare, M. (1993) 'Cartilage degradation by cocultures of transformed macrophage and fibroblast cell lines. A model of metalloproteinase-mediated connective tissue degradation'. *J Immunol,* 150: (5): 1922-1931.

Jaskowiak, P.A., Campello, R.J. and Costa, I.G. (2014) 'On the selection of appropriate distances for gene expression data clustering'. *BMC Bioinformatics,* 15: (2): S2.

Johansen, F.E., Braathen, R. and Brandtzaeg, P. (2000) 'Role of J chain in secretory immunoglobulin formation'. *Scand J Immunol,* 52: (3): 240-248.

Johnson, D.B., Balko, J.M., Compton, M.L., Chalkias, S., Gorham, J., Xu, Y., Hicks, M., Puzanov, I., Alexander, M.R., Bloomer, T.L., Becker, J.R., Slosky, D.A., Phillips, E.J., Pilkinton, M.A., Craig-Owens, L., Kola, N., Plautz, G., Reshef, D.S., Deutsch, J.S., Deering, R.P., Olenchock, B.A., Lichtman, A.H., Roden, D.M., Seidman, C.E., Koralnik, I.J., Seidman, J.G., Hoffman, R.D., Taube, J.M., Diaz, L.a.J., Anders, R.A., Sosman, J.A. and Moslehi, J.J. (2016) 'Fulminant Myocarditis with Combination Immune Checkpoint Blockade'. *New England Journal of Medicine,* 375: (18): 1749-1755.

Johnson, W.E., Li, C. and Rabinovic, A. (2007) 'Adjusting batch effects in microarray expression data using empirical Bayes methods'. *Biostatistics,* 8: (1): 118-127.

Johnston, J.A., Kawamura, M., Kirken, R.A., Chen, Y.Q., Blake, T.B., Shibuya, K., Ortaldo, J.R., Mcvicar, D.W. and O'shea, J.J. (1994) 'Phosphorylation and activation of the Jak-3 Janus kinase in response to interleukin-2'. *Nature,* 370: (6485): 151-153.

Jones, C., Virji, M. and Crocker, P.R. (2003) 'Recognition of sialylated meningococcal lipopolysaccharide by siglecs expressed on myeloid cells leads to enhanced bacterial uptake'. *Mol Microbiol,* 49: (5): 1213-1225.

Jongbloed, S.L., Lebre, M.C., Fraser, A.R., Gracie, J.A., Sturrock, R.D., Tak, P.P. and Mcinnes, I.B. (2006) 'Enumeration and phenotypical analysis of distinct dendritic cell subsets in psoriatic arthritis and rheumatoid arthritis'. *Arthritis Res Ther,* 8: (1): R15.

Jorgensen, I., Kos, J., Krasovec, M., Troelsen, L., Klarlund, M., Jensen, T.W., Hansen, M.S. and Jacobsen, S. (2011) 'Serum cysteine proteases and their inhibitors in rheumatoid arthritis: relation to disease activity and radiographic progression'. *Clin Rheumatol,* 30: (5): 633-638.

Juarez, M., Filer, A. and Buckley, C.D. (2012) 'Fibroblasts as therapeutic targets in rheumatoid arthritis and cancer'. *Swiss Med Wkly,* 142: w13529.

Juárez Pérez, M.J.A. (2014) *Functional and transcriptional characterisation of synovial fibroblasts in early inflammatory arthritis and established rheumatoid arthritis.* . Birmingham : University of Birmingham, 2014.

Jungel, A., Baresova, V., Ospelt, C., Simmen, B., Michel, B., Gay, R., Gay, S., Seemayer, C. and Neidhart, M. (2006) 'Trichostatin A sensitises rheumatoid arthritis synovial fibroblasts for TRAIL-induced apoptosis'. *Ann Rheum Dis,* 65: 910 - 912.

Källberg, H., Padyukov, L., Plenge, R.M., Rönnelid, J., Gregersen, P.K., Van Der Helm-Van Mil, A.H.M., Toes, R.E.M., Huizinga, T.W., Klareskog, L. and Alfredsson, L. (2007) 'Gene-Gene and Gene-Environment Interactions Involving HLA-DRB1, PTPN22, and Smoking in Two Subsets of Rheumatoid Arthritis'. *The American Journal of Human Genetics,* 80: (5): 867-875.

Kalluri, R. (2016) 'The biology and function of fibroblasts in cancer'. *Nat Rev Cancer,* 16: (9): 582-598.

Kang, C.P., Lee, H.-S., Ju, H., Cho, H., Kang, C. and Bae, S.-C. (2006) 'A functional haplotype of the PADI4 gene associated with increased rheumatoid arthritis susceptibility in Koreans'. *Arthritis & Rheumatism,* 54: (1): 90-96.

Karonitsch, T., Dalwigk, K., Byrne, R., Niedereiter, B., Cetin, E., Wanivenhaus, A., Scheinecker, C., Smolen, J.S. and Kiener, H.P. (2010) 'IFN-gamma promotes fibroblast-like synoviocytes motility'. *Annals of the Rheumatic Diseases,* 69: (Suppl 2): A63-A63.

Karouzakis, E., Gay, R., Michel, B., Gay, S. and Neidhart, M. (2009) 'DNA hypomethylation in rheumatoid arthritis synovial fibroblasts'. *Arthritis Rheum,* 60: 3613 - 3622.

Kassem, A., Henning, P., Kindlund, B., Lindholm, C. and Lerner, U.H. (2015) 'TLR5, a novel mediator of innate immunity-induced osteoclastogenesis and bone loss'. *FASEB J,* 29: (11): 4449-4460.

Katzmann, D.J., Stefan, C.J., Babst, M. and Emr, S.D. (2003) 'Vps27 recruits ESCRT machinery to endosomes during MVB sorting'. *The Journal of Cell Biology,* 162: (3): 413-423.

Kauffmann, A., Gentleman, R. and Huber, W. (2009) 'arrayQualityMetrics--a bioconductor package for quality assessment of microarray data'. *Bioinformatics,* 25: (3): 415-416.

Kawanaka, N., Yamamura, M., Aita, T., Morita, Y., Okamoto, A., Kawashima, M., Iwahashi, M., Ueno, A., Ohmoto, Y. and Makino, H. (2002) 'CD14+,CD16+ blood monocytes and joint inflammation in rheumatoid arthritis'. *Arthritis Rheum,* 46: (10): 2578-2586.

Khandpur, R., Carmona-Rivera, C., Vivekanandan-Giri, A., Gizinski, A., Yalavarthi, S., Knight, J.S., Friday, S., Li, S., Patel, R.M., Subramanian, V., Thompson, P., Chen, P., Fox, D.A., Pennathur, S. and Kaplan, M.J. (2013) 'NETs are a source of citrullinated autoantigens and stimulate inflammatory responses in rheumatoid arthritis'. *Sci Transl Med,* 5: (178): 178ra140.

Kiener, H.P., Lee, D.M., Agarwal, S.K. and Brenner, M.B. (2006) 'Cadherin-11 Induces Rheumatoid Arthritis Fibroblast-Like Synoviocytes to Form Lining Layers in Vitro'. *The American Journal of Pathology,* 168: (5): 1486-1499.

Kiener, H.P., Niederreiter, B., Lee, D.M., Jimenez-Boj, E., Smolen, J.S. and Brenner, M.B. (2009) 'Cadherin 11 promotes invasive behavior of fibroblast-like synoviocytes'. *Arthritis & Rheumatism,* 60: (5): 1305-1310.

Kim, K.W., Cho, M.L., Kim, H.R., Ju, J.H., Park, M.K., Oh, H.J., Kim, J.S., Park, S.H., Lee, S.H. and Kim, H.Y. (2007a) 'Up-regulation of stromal cell-derived factor 1 (CXCL12) production in rheumatoid synovial fibroblasts through interactions with T lymphocytes: role of interleukin-17 and CD40L-CD40 interaction'. *Arthritis Rheum,* 56: (4): 1076-1086.

Kim, K.W., Cho, M.L., Lee, S.H., Oh, H.J., Kang, C.M., Ju, J.H., Min, S.Y., Cho, Y.G., Park, S.H. and Kim, H.Y. (2007b) 'Human rheumatoid synovial fibroblasts promote osteoclastogenic activity by activating RANKL via TLR-2 and TLR-4 activation'. *Immunol Lett,* 110: (1): 54-64.

Kim, K.W., Cho, M.L., Oh, H.J., Kim, H.R., Kang, C.M., Heo, Y.M., Lee, S.H. and Kim, H.Y. (2009) 'TLR-3 enhances osteoclastogenesis through upregulation of RANKL expression from fibroblast-like synoviocytes in patients with rheumatoid arthritis'. *Immunol Lett,* 124: (1): 9-17.

Kinloch, A., Tatzer, V., Wait, R., Peston, D., Lundberg, K., Donatien, P., Moyes, D., Taylor, P. and Venables, P. (2005) 'Identification of citrullinated alpha-enolase as a candidate autoantigen in rheumatoid arthritis'. *Arthritis Research & Therapy,* 7: (6): R1421 - R1429.

Kisielow, M., Kisielow, J., Capoferri-Sollami, G. and Karjalainen, K. (2005) 'Expression of lymphocyte activation gene 3 (LAG-3) on B cells is induced by T cells'. *Eur J Immunol,* 35: (7): 2081-2088.

Koch, A.E., Harlow, L.A., Haines, G.K., Amento, E.P., Unemori, E.N., Wong, W.L., Pope, R.M. and Ferrara, N. (1994) 'Vascular endothelial growth factor. A cytokine modulating endothelial function in rheumatoid arthritis'. *J Immunol,* 152: (8): 4149-4156.

Kohyama, M., Ise, W., Edelson, B.T., Wilker, P.R., Hildner, K., Mejia, C., Frazier, W.A., Murphy, T.L. and Murphy, K.M. (2008) 'Role for Spi-C in the development of red pulp macrophages and splenic iron homeostasis'. *Nature,* 457: 318.

Kotrych, D., Dziedziejko, V., Safranow, K., Drozdzik, M. and Pawlik, A. (2015) 'CXCL9 and CXCL10 gene polymorphisms in patients with rheumatoid arthritis'. *Rheumatol Int,* 35: (8): 1319-1323.

Koussounadis, A., Langdon, S.P., Um, I.H., Harrison, D.J. and Smith, V.A. (2015) 'Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system'. *Sci Rep,* 5: 10775.

Kruskal, J.B. (1964) 'Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis'. *Psychometrika,* 29: (1): 1-27.

Kunisch, E., Fuhrmann, R., Roth, A., Winter, R., Lungershausen, W. and Kinne, R.W. (2004) 'Macrophage specificity of three anti-CD68 monoclonal antibodies (KP1, EBM11, and PGM1) widely used for immunohistochemistry and flow cytometry'. *Ann Rheum Dis,* 63: (7): 774-784.

Kurreeman, F.a.S., Padyukov, L., Marques, R.B., Schrodi, S.J., Seddighzadeh, M., Stoeken-Rijsbergen, G., Van Der Helm-Van Mil, A.H.M., Allaart, C.F., Verduyn, W., Houwing-Duistermaat, J., Alfredsson, L., Begovich, A.B., Klareskog, L., Huizinga, T.W.J. and Toes, R.E.M. (2007) 'A Candidate Gene Approach Identifies the TRAF1/C5 Region as a Risk Factor for Rheumatoid Arthritis'. *PLOS Medicine,* 4: (9): e278.

Kyburz, D., Gabay, C., Michel, B.A. and Finckh, A. (2011) 'The long-term impact of early treatment of rheumatoid arthritis on radiographic progression: a population-based cohort study'. *Rheumatology (Oxford),* 50: (6): 1106-1110.

Lacey, D.C., Achuthan, A., Fleetwood, A.J., Dinh, H., Roiniotis, J., Scholz, G.M., Chang, M.W., Beckman, S.K., Cook, A.D. and Hamilton, J.A. (2012) 'Defining GM-CSF- and macrophage-CSF-dependent macrophage responses by in vitro models'. *J Immunol,* 188: (11): 5752-5765.

Langmead, B. and Salzberg, S.L. (2012) 'Fast gapped-read alignment with Bowtie 2'. *Nat Meth,* 9: (4): 357-359.

Laragione, T., Brenner, M., Sherry, B. and Gulko, P.S. (2011) 'CXCL10 and its receptor CXCR3 regulate synovial fibroblast invasion in rheumatoid arthritis'. *Arthritis Rheum,* 63: (11): 3274-3283.

Laverman, P., Van Der Geest, T., Terry, S.Y., Gerrits, D., Walgreen, B., Helsen, M.M., Nayak, T.K., Freimoser-Grundschober, A., Waldhauer, I., Hosse, R.J., Moessner, E., Umana, P., Klein, C., Oyen, W.J., Koenders, M.I. and Boerman, O.C. (2015) 'Immuno-PET and Immuno-SPECT of Rheumatoid Arthritis with Radiolabeled Anti-Fibroblast Activation Protein Antibody Correlates with Severity of Arthritis'. *J Nucl Med,* 56: (5): 778-783.

Lavin, Y., Winter, D., Blecher-Gonen, R., David, E., Keren-Shaul, H., Merad, M., Jung, S. and Amit, I. (2014) 'Tissue-Resident Macrophage Enhancer Landscapes Are Shaped by the Local Microenvironment'. *Cell,* 159: (6): 1312-1326.

Lee, C.H., Shah, B., Moioli, E.K. and Mao, J.J. (2010) 'CTGF directs fibroblast differentiation from human mesenchymal stem/stromal cells and defines connective tissue healing in a rodent injury model'. *The Journal of Clinical Investigation,* 120: (9): 3340-3349.

Lee, D.M., Kiener, H.P., Agarwal, S.K., Noss, E.H., Watts, G.F.M., Chisaka, O., Takeichi, M. and Brenner, M.B. (2007) 'Cadherin-11 in Synovial Lining Formation and Pathology in Arthritis'.

Science, 315: (5814): 1006-1010.

Lee, J.-H., Kim, B., Jin, W.J., Kim, H.-H., Ha, H. and Lee, Z.H. (2017) 'Pathogenic roles of CXCL10 signaling through CXCR3 and TLR4 in macrophages and T cells: relevance for arthritis'. *Arthritis Research & Therapy,* 19: (1): 163.

Lee, S.S., Joo, Y.S., Kim, W.U., Min, D.J., Min, J.K., Park, S.H., Cho, C.S. and Kim, H.Y. (2001) 'Vascular endothelial growth factor levels in the serum and synovial fluid of patients with rheumatoid arthritis'. *Clinical and experimental rheumatology,* 19: (3): 321-324.

Lee, S.Y., Yoon, B.Y., Kim, J.I., Heo, Y.M., Woo, Y.J., Park, S.H., Kim, H.Y., Kim, S.I. and Cho, M.L. (2014) 'Interleukin-17 increases the expression of Toll-like receptor 3 via the STAT3 pathway in rheumatoid arthritis fibroblast-like synoviocytes'. *Immunology,* 141: (3): 353-361.

Leek, J.T. (2014) 'svaseq: removing batch effects and other unwanted noise from sequencing data'. *Nucleic Acids Res,* 42: (21): e161-e161.

Leek, J.T. and Storey, J.D. (2007) 'Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis'. *PLOS Genetics,* 3: (9): e161.

Leek, J.T. and Storey, J.D. (2008) 'A general framework for multiple testing dependence'. *Proceedings of the National Academy of Sciences,* 105: (48): 18718-18723.

Lefevre, S., Knedla, A., Tennie, C., Kampmann, A., Wunrau, C., Dinser, R., Korb, A., Schnaker, E.M., Tarner, I.H., Robbins, P.D., Evans, C.H., Sturz, H., Steinmeyer, J., Gay, S., Scholmerich, J., Pap, T., Muller-Ladner, U. and Neumann, E. (2009) 'Synovial fibroblasts spread rheumatoid arthritis to unaffected joints'. *Nat Med,* 15: (12): 1414-1420.

Leonard, W.J. and O'shea, J.J. (1998) 'Jaks and STATs: biological implications'. *Annu Rev Immunol,* 16: 293-322.

Lin, J., Huo, R., Xiao, L., Zhu, X., Xie, J., Sun, S., He, Y., Zhang, J., Sun, Y., Zhou, Z., Wu, P., Shen, B., Li, D. and Li, N. (2014) 'A Novel p53/microRNA-22/Cyr61 Axis in Synovial Cells Regulates Inflammation in Rheumatoid Arthritis'. *Arthritis & Rheumatology,* 66: (1): 49-59.

Listing, J., Gerhold, K. and Zink, A. (2013) 'The risk of infections associated with rheumatoid arthritis, with its comorbidity and treatment'. *Rheumatology (Oxford),* 52: (1): 53-61.

Liu, X. and Piela-Smith, T.H. (2000) 'Fibrin(ogen)-Induced Expression of ICAM-1 and Chemokines in Human Synovial Fibroblasts'. *The Journal of Immunology,* 165: (9): 5255-5261.

Liu, X., Zhang, H., Chang, X., Shen, J., Zheng, W., Xu, Y., Wang, J., Gao, W. and He, S. (2017) 'Upregulated expression of CCR3 in rheumatoid arthritis and CCR3-dependent activation of fibroblast-like synoviocytes'. *Cell Biol Toxicol,* 33: (1): 15-26.

Liu, Y., Beyer, A. and Aebersold, R. (2016) 'On the Dependency of Cellular Protein Levels on mRNA Abundance'. *Cell,* 165: (3): 535-550.

Long, L., Yu, P., Liu, Y., Wang, S., Li, R., Shi, J., Zhang, X., Li, Y., Sun, X., Zhou, B., Cui, L. and Li, Z. (2013) 'Upregulated microRNA-155 expression in peripheral blood mononuclear cells and fibroblast-like synoviocytes in rheumatoid arthritis'. *Clin Dev Immunol,* 2013: 296139.

Lubberts, E. (2015) 'The IL-23-IL-17 axis in inflammatory arthritis'. *Nat Rev Rheumatol,* 11: (7): 415-429.

Lundberg, K., Kinloch, A., Fisher, B.A., Wegner, N., Wait, R., Charles, P., Mikuls, T.R. and Venables, P.J. (2008) 'Antibodies to citrullinated alpha-enolase peptide 1 are specific for rheumatoid arthritis and cross-react with bacterial enolase'. *Arthritis Rheum,* 58: (10): 3009-3019.

Macfadyen, J.R., Haworth, O., Roberston, D., Hardie, D., Webster, M.-T., Morris, H.R., Panico, M., Sutton-Smith, M., Dell, A., Van Der Geer, P., Wienke, D., Buckley, C.D. and Isacke, C.M. (2005) 'Endosialin (TEM1, CD248) is a marker of stromal fibroblasts and is not selectively expressed on tumour endothelium'. *FEBS Letters,* 579: (12): 2569-2575.

Macgregor, A.J., Snieder, H., Rigby, A.S., Koskenvuo, M., Kaprio, J., Aho, K. and Silman, A.J. (2000) 'Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins'. *Arthritis & Rheumatism,* 43: (1): 30-37.

Machold, K.P., Stamm, T.A., Eberl, G.J.M., Nell, V.K.P., Dunky, A., Uffmann, M. and Smolen, J.S. (2002) 'Very recent onset arthritis--clinical, laboratory, and radiological findings during the first year of disease'. *The Journal of Rheumatology,* 29: (11): 2278-2287.

Maia, M., De Vriese, A., Janssens, T., Moons, M., Van Landuyt, K., Tavernier, J., Lories, R.J. and Conway, E.M. (2010) 'CD248 and its cytoplasmic domain: a therapeutic target for arthritis'. *Arthritis Rheum,* 62: (12): 3595-3606.

Majeed, M., Mcqueen, F., Yeoman, S. and Mclean, L. (2004) 'Relationship between serum hyaluronic acid level and disease activity in early rheumatoid arthritis'. *Annals of the Rheumatic Diseases,* 63: (9): 1166-1168.

Malavez, Y., Voss, O.H., Gonzalez-Mejia, M.E., Parihar, A. and Doseff, A.I. (2015) 'Distinct contribution of protein kinase Cdelta and protein kinase Cepsilon in the lifespan and immune response of human blood monocyte subpopulations'. *Immunology,* 144: (4): 611-620.

Manabe, N., Oda, H., Nakamura, K., Kuga, Y., Uchida, S. and Kawaguchi, H. (1999) 'Involvement of fibroblast growth factor-2 in joint destruction of rheumatoid arthritis patients'. *Rheumatology,* 38: (8): 714-720.

Mantovani, A., Sica, A., Sozzani, S., Allavena, P., Vecchi, A. and Locati, M. (2004) 'The chemokine system in diverse forms of macrophage activation and polarization'. *Trends in Immunology,* 25: (12): 677-686.

Mantovani, A., Sozzani, S., Locati, M., Allavena, P. and Sica, A. (2002) 'Macrophage polarization: tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes'. *Trends in Immunology,* 23: (11): 549-555.

Martin, M. (2011) 'Cutadapt removes adapter sequences from high-throughput sequencing reads'. *2011,* 17: (1).

Marudamuthu, A.S., Shetty, S.K., Bhandary, Y.P., Karandashova, S., Thompson, M., Sathish, V., Florova, G., Hogan, T.B., Pabelick, C.M., Prakash, Y.S., Tsukasaki, Y., Fu, J., Ikebe, M., Idell, S. and Shetty, S. (2015) 'Plasminogen activator inhibitor-1 suppresses profibrotic responses in fibroblasts from fibrotic lungs'. *J Biol Chem,* 290: (15): 9428-9441.

Massey, H.M. and Flanagan, A.M. (1999) 'Human osteoclasts derive from CD14-positive monocytes'. *Br J Haematol,* 106: (1): 167-170.

Masson-Bessiere, C., Sebbag, M., Girbal-Neuhauser, E., Nogueira, L., Vincent, C., Senshu, T. and Serre, G. (2001) 'The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the alpha- and beta-chains of fibrin'. *J Immunol,* 166: (6): 4177-4184.

Mcgovern, D.P.B., Jones, M.R., Taylor, K.D., Marciante, K., Yan, X., Dubinsky, M., Ippoliti, A., Vasiliauskas, E., Berel, D., Derkowski, C., Dutridge, D., Fleshner, P., Shih, D.Q., Melmed, G., Mengesha, E., King, L., Pressman, S., Haritunians, T., Guo, X., Targan, S.R. and Rotter, J.I. (2010) 'Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease'. *Human Molecular Genetics,* 19: (17): 3468-3476.

Mckenney, J.K., Weiss, S.W. and Folpe, A.L. (2001) 'CD31 expression in intratumoral macrophages: a potential diagnostic pitfall'. *Am J Surg Pathol,* 25: (9): 1167-1173.

Meerschaert, J. and Furie, M.B. (1994) 'Monocytes use either CD11/CD18 or VLA-4 to migrate across human endothelium in vitro'. *J Immunol,* 152: (4): 1915-1926.

Meininger, C., Yano, H., Rottapel, R., Bernstein, A., Zsebo, K. and Zetter, B. (1992) 'The c-kit receptor ligand functions as a mast cell chemoattractant'. *Blood,* 79: (4): 958-963.

Michou, L., Lasbleiz, S., Rat, A.C., Migliorini, P., Balsa, A., Westhovens, R., Barrera, P., Alves, H., Pierlot, C., Glikmans, E., Garnier, S., Dausset, J., Vaz, C., Fernandes, M., Petit-Teixeira, E., Lemaire, I., Pascual-Salcedo, D., Bombardieri, S., Dequeker, J., Radstake, T.R., Van Riel, P., Van De Putte, L., Lopes-Vaz, A., Prum, B., Bardin, T., Dieude, P. and Cornelis, F. (2007) 'Linkage proof for PTPN22, a rheumatoid arthritis susceptibility gene and a human autoimmunity gene'. *Proc Natl Acad Sci U S A,* 104: (5): 1649-1654.

Micklem, K., Rigney, E., Cordell, J., Simmons, D., Stross, P., Turley, H., Seed, B. and Mason, D. (1989) 'A human macrophage-associated antigen (CD68) detected by six different monoclonal antibodies'. *British Journal of Haematology,* 73: (1): 6-11.

Mikita, J., Dubourdieu-Cassagno, N., Deloire, M.S., Vekris, A., Biran, M., Raffard, G., Brochet, B., Canron, M.-H., Franconi, J.-M., Boiziau, C. and Petry, K.G. (2011) 'Altered M1/M2 activation patterns of monocytes in severe relapsing experimental rat model of multiple sclerosis. Amelioration of clinical status by M2 activated monocyte administration'. *Multiple Sclerosis Journal,* 17: (1): 2-15.

Mikuls, T.R., Payne, J.B., Yu, F., Thiele, G.M., Reynolds, R.J., Cannon, G.W., Markt, J., Mcgowan, D., Kerr, G.S., Redman, R.S., Reimold, A., Griffiths, G., Beatty, M., Gonzalez, S.M., Bergman, D.A., Hamilton, B.C., 3rd, Erickson, A.R., Sokolove, J., Robinson, W.H., Walker, C., Chandad, F. and O'dell, J.R. (2014) 'Periodontitis and Porphyromonas gingivalis in patients with rheumatoid arthritis'. *Arthritis Rheumatol,* 66: (5): 1090-1100.

Misharin, Alexander v., Cuda, Carla m., Saber, R., Turner, Jason d., Gierut, Angelica k., Haines, G.K., Iii, Berdnikovs, S., Filer, A., Clark, Andrew r., Buckley, Christopher d., Mutlu, Gökhan m., Budinger, G.R.S. and Perlman, H. (2014) 'Nonclassical Ly6C− Monocytes Drive the Development of Inflammatory Arthritis in Mice'. *Cell Reports,* 9: (2): 591-604.

Mizoguchi, F., Slowikowski, K., Marshall, J.L., Wei, K., Rao, D.A., Chang, S.K., Nguyen, H.N., Noss,

E.H., Turner, J.D., Earp, B.E., Blazar, P.E., Wright, J., Simmons, B.P., Donlin, L.T., Kalliolias, G.D., Goodman, S.M., Bykerk, V.P., Ivashkiv, L.B., Lederer, J.A., Hacohen, N., Nigrovic, P.A., Filer, A., Buckley, C.D., Raychaudhuri, S. and Brenner, M.B. (2017) 'Single Cell Transcriptomics And Flow Cytometry Reveal Disease-Associated Fibroblast Subsets In Rheumatoid Arthritis'. *bioRxiv*.

Mizuno, T., Chou, M.Y. and Inouye, M. (1984) 'A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA)'. *Proceedings of the National Academy of Sciences of the United States of America,* 81: (7): 1966-1970.

Mizutani, M., Pino, P.A., Saederup, N., Charo, I.F., Ransohoff, R.M. and Cardona, A.E. (2012) 'The fractalkine receptor but not CCR2 is present on microglia from embryonic development throughout adulthood'. *J Immunol,* 188: (1): 29-36.

Mohamadzadeh, M., Degrendele, H., Arizpe, H., Estess, P. and Siegelman, M. (1998) 'Proinflammatory stimuli regulate endothelial hyaluronan expression and CD44/HA-dependent primary adhesion'. *J Clin Invest,* 101: (1): 97-108.

Morko, J., Kiviranta, R., Joronen, K., Saamanen, A.M., Vuorio, E. and Salminen-Mankonen, H. (2005) 'Spontaneous development of synovitis and cartilage degeneration in transgenic mice overexpressing cathepsin K'. *Arthritis Rheum,* 52: (12): 3713-3717.

Moura, C.S., Abrahamowicz, M., Beauchamp, M.E., Lacaille, D., Wang, Y., Boire, G., Fortin, P.R., Bessette, L., Bombardier, C., Widdifield, J., Hanly, J.G., Feldman, D., Maksymowych, W., Peschken, C., Barnabe, C., Edworthy, S. and Bernatsky, S. (2015) 'Early medication use in new-onset rheumatoid arthritis may delay joint replacement: results of a large population-based study'. *Arthritis Res Ther,* 17: 197.

Mulherin, D., Fitzgerald, O. and Bresnihan, B. (1996) 'Synovial tissue macrophage populations and articular damage in rheumatoid arthritis'. *Arthritis Rheum,* 39: (1): 115-124.

Muller-Alouf, H., Alouf, J.E., Gerlach, D., Ozegowski, J.H., Fitting, C. and Cavaillon, J.M. (1994) 'Comparative study of cytokine release by human peripheral blood mononuclear cells stimulated with Streptococcus pyogenes superantigenic erythrogenic toxins, heat-killed streptococci, and lipopolysaccharide'. *Infect Immun,* 62: (11): 4915-4921.

Muller-Ladner, U., Kriegsmann, J., Franklin, B.N., Matsumoto, S., Geiler, T., Gay, R.E. and Gay, S. (1996) 'Synovial fibroblasts of patients with rheumatoid arthritis attach to and invade normal human cartilage when engrafted into SCID mice'. *Am J Pathol,* 149: (5): 1607-1615.

Murphy, C.A., Langrish, C.L., Chen, Y., Blumenschein, W., Mcclanahan, T., Kastelein, R.A., Sedgwick, J.D. and Cua, D.J. (2003) 'Divergent pro- and antiinflammatory roles for IL-23 and IL-12 in joint autoimmune inflammation'. *J Exp Med,* 198: (12): 1951-1957.

Murray, P.J., Allen, J.E., Biswas, S.K., Fisher, E.A., Gilroy, D.W., Goerdt, S., Gordon, S., Hamilton, J.A., Ivashkiv, L.B., Lawrence, T., Locati, M., Mantovani, A., Martinez, F.O., Mege, J.L., Mosser, D.M., Natoli, G., Saeij, J.P., Schultze, J.L., Shirey, K.A., Sica, A., Suttles, J., Udalova, I., Van Ginderachter, J.A., Vogel, S.N. and Wynn, T.A. (2014) 'Macrophage activation and polarization: nomenclature and experimental guidelines'. *Immunity,* 41: (1): 14-20.

Nagayoshi, R., Nagai, T., Matsushita, K., Sato, K., Sunahara, N., Matsuda, T., Nakamura, T., Komiya, S., Onda, M. and Matsuyama, T. (2005) 'Effectiveness of anti-folate receptor beta antibody conjugated with truncated Pseudomonas exotoxin in the targeting of rheumatoid

arthritis synovial macrophages'. *Arthritis Rheum,* 52: (9): 2666-2675.

Nakano, K., Okada, Y., Saito, K. and Tanaka, Y. (2004) 'Induction of RANKL expression and osteoclast maturation by the binding of fibroblast growth factor 2 to heparan sulfate proteoglycan on rheumatoid synovial fibroblasts'. *Arthritis & Rheumatism,* 50: (8): 2450-2458.

Nasu, Y., Nishida, K., Miyazawa, S., Komiyama, T., Kadota, Y., Abe, N., Yoshida, A., Hirohata, S., Ohtsuka, A. and Ozaki, T. (2008) 'Trichostatin A, a histone deacetylase inhibitor, suppresses synovial inflammation and subsequent cartilage destruction in a collagen antibody-induced arthritis mouse model'. *Osteoarthritis Cartilage,* 16: 723 - 732.

Naylor, A.J., Filer, A. and Buckley, C.D. (2013) 'The role of stromal cells in the persistence of chronic inflammation'. *Clin Exp Immunol,* 171: (1): 30-35.

Nell, V.P.K., Machold, K.P., Eberl, G., Stamm, T.A., Uffmann, M. and Smolen, J.S. (2004) 'Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis'. *Rheumatology,* 43: (7): 906-914.

Nockher, W.A. and Scherberich, J.E. (1998) 'Expanded CD14+ CD16+ monocyte subpopulation in patients with acute and chronic infections undergoing hemodialysis'. *Infect Immun,* 66: (6): 2782-2790.

Okamoto, H., Yamamura, M., Morita, Y., Harada, S., Makino, H. and Ota, Z. (1997) 'The synovial expression and serum levels of interleukin-6, interleukin-11, leukemia inhibitory factor, and oncostatin M in rheumatoid arthritis'. *Arthritis Rheum,* 40: (6): 1096-1105.

Onishi, R.M. and Gaffen, S.L. (2010) 'Interleukin-17 and its target genes: mechanisms of interleukin-17 function in disease'. *Immunology,* 129: (3): 311-321.

Oshlack, A. and Wakefield, M.J. (2009) 'Transcript length bias in RNA-seq data confounds systems biology'. *Biology Direct,* 4: (1): 14.

Ospelt, C., Brentano, F., Rengel, Y., Stanczyk, J., Kolling, C., Tak, P.P., Gay, R.E., Gay, S. and Kyburz, D. (2008) 'Overexpression of toll-like receptors 3 and 4 in synovial tissue from patients with early rheumatoid arthritis: Toll-like receptor expression in early and longstanding arthritis'. *Arthritis & Rheumatism,* 58: (12): 3684-3692.

Page, C.E., Smale, S., Carty, S.M., Amos, N., Lauder, S.N., Goodfellow, R.M., Richards, P.J., Jones, S.A., Topley, N. and Williams, A.S. (2010) 'Interferon-γ inhibits interleukin-1β-induced matrix metalloproteinase production by synovial fibroblasts and protects articular cartilage in early arthritis'. *Arthritis Research & Therapy,* 12: (2): R49.

Pandya, J.M., Lundell, A.-C., Andersson, K., Nordström, I., Theander, E. and Rudin, A. (2017) 'Blood chemokine profile in untreated early rheumatoid arthritis: CXCL10 as a disease activity marker'. *Arthritis Research & Therapy,* 19: (1): 20.

Pankow, W., Neumann, K., Ruschoff, J. and Von Wichert, P. (1995) 'Human alveolar macrophages: comparison of cell size, autofluorescence, and HLA-DR antigen expression in smokers and nonsmokers'. *Cancer Detect Prev,* 19: (3): 268-273.

Pap, T., Aupperle, K.R., Gay, S., Firestein, G.S. and Gay, R.E. (2001) 'Invasiveness of synovial fibroblasts is regulated by p53 in the SCID mouse in vivo model of cartilage invasion'. *Arthritis*

Rheum, 44: (3): 676-681.

Park, K. and Scott, A.L. (2010) 'Cholesterol 25-hydroxylase production by dendritic cells and macrophages is regulated by type I interferons'. *J Leukoc Biol,* 88: (6): 1081-1087.

Parsonage, G., Falciani, F., Burman, A., Filer, A., Ross, E., Bofill, M., Martin, S., Salmon, M. and Buckley, C.D. (2003) 'Global gene expression profiles in fibroblasts from synovial, skin and lymphoid tissue reveals distinct cytokine and chemokine expression patterns'. *Thromb Haemost,* 90: (4): 688-697.

Parsonage, G., Filer, A., Bik, M., Hardie, D., Lax, S., Howlett, K., Church, L.D., Raza, K., Wong, S.H., Trebilcock, E., Scheel-Toellner, D., Salmon, M., Lord, J.M. and Buckley, C.D. (2008) 'Prolonged, granulocyte-macrophage colony-stimulating factor-dependent, neutrophil survival following rheumatoid synovial fibroblast activation by IL-17 and TNFalpha'. *Arthritis Res Ther,* 10: (2): R47.

Parsonage, G., Filer, A.D., Haworth, O., Nash, G.B., Rainger, G.E., Salmon, M. and Buckley, C.D. (2005) 'A stromal address code defined by fibroblasts'. *Trends Immunol,* 26: (3): 150-156.

Parwaresch, M.R., Radzun, H.J., Kreipe, H., Hansmann, M.L. and Barth, J. (1986) 'Monocyte/macrophage-reactive monoclonal antibody Ki-M6 recognizes an intracytoplasmic antigen'. *Am J Pathol,* 125: (1): 141-151.

Pedersen, M., Jacobsen, S., Klarlund, M., Pedersen, B., Wiik, A., Wohlfahrt, J. and Frisch, M. (2006) 'Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides'. *Arthritis Research & Therapy,* 8: (4): R133.

Perdiguero, E.G., Klapproth, K., Schulz, C., Busch, K., Azzoni, E., Crozet, L., Garner, H., Trouillet, C., De Bruijn, M.F., Geissmann, F. and Rodewald, H.-R. (2014) 'Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors'. *Nature,* advance online publication.

Pettit, A.R., Macdonald, K.P., O'sullivan, B. and Thomas, R. (2000) 'Differentiated dendritic cells expressing nuclear RelB are predominantly located in rheumatoid synovial tissue perivascular mononuclear cell aggregates'. *Arthritis Rheum,* 43: (4): 791-800.

Pierer, M., Rethage, J., Seibl, R., Lauener, R., Brentano, F., Wagner, U., Hantzschel, H., Michel, B.A., Gay, R.E., Gay, S. and Kyburz, D. (2004) 'Chemokine Secretion of Rheumatoid Arthritis Synovial Fibroblasts Stimulated by Toll-Like Receptor 2 Ligands'. *The Journal of Immunology,* 172: (2): 1256-1265.

Pilling, D., Fan, T., Huang, D., Kaul, B. and Gomer, R.H. (2009) 'Identification of markers that distinguish monocyte-derived fibrocytes from monocytes, macrophages, and fibroblasts'. *PLoS ONE,* 4: (10): e7475.

Plenge , R.M., Seielstad , M., Padyukov , L., Lee , A.T., Remmers , E.F., Ding , B., Liew , A., Khalili , H., Chandrasekaran , A., Davies , L.R.L., Li , W., Tan , A.K.S., Bonnard , C., Ong , R.T.H., Thalamuthu , A., Pettersson , S., Liu , C., Tian , C., Chen , W.V., Carulli , J.P., Beckman , E.M., Altshuler , D., Alfredsson , L., Criswell , L.A., Amos , C.I., Seldin , M.F., Kastner , D.L., Klareskog , L. and Gregersen , P.K. (2007) 'TRAF1–C5 as a Risk Locus for Rheumatoid Arthritis — A Genomewide Study'. *New England Journal of Medicine,* 357: (12): 1199-1209.

Portela, A. and Esteller, M. (2010) 'Epigenetic modifications and human disease'. *Nat Biotech,*

28: (10): 1057-1068.

Priller, J. and Bottcher, C. (2017) 'Patrolling monocytes sense peripheral infection and induce cytokine-mediated neuronal dysfunction'. *Nat Med,* 23: (6): 659-661.

Pulford, K.A., Rigney, E.M., Micklem, K.J., Jones, M., Stross, W.P., Gatter, K.C. and Mason, D.Y. (1989) 'KP1: a new monoclonal antibody that detects a monocyte/macrophage associated antigen in routinely processed tissue sections'. *J Clin Pathol,* 42: (4): 414-421.

Pulford, K.A., Sipos, A., Cordell, J.L., Stross, W.P. and Mason, D.Y. (1990) 'Distribution of the CD68 macrophage/myeloid associated antigen'. *Int Immunol,* 2: (10): 973-980.

Ragan, C. and Meyer, K. (1949) 'THE HYALURONIC ACID OF SYNOVIAL FLUID IN RHEUMATOID ARTHRITIS'. *Journal of Clinical Investigation,* 28: (1): 56-59.

Raschke, W.C., Baird, S., Ralph, P. and Nakoinz, I. (1978) 'Functional macrophage cell lines transformed by abelson leukemia virus'. *Cell,* 15: (1): 261-267.

Reich, N., Maurer, B., Akhmetshina, A., Venalis, P., Dees, C., Zerr, P., Palumbo, K., Zwerina, J., Nevskaya, T., Gay, S., Distler, O., Schett, G. and Distler, J.H. (2010) 'The transcription factor Fra-2 regulates the production of extracellular matrix in systemic sclerosis'. *Arthritis Rheum,* 62: (1): 280-290.

Rey-Giraud, F., Hafner, M. and Ries, C.H. (2012) 'Generation of Monocyte-Derived Macrophages under Serum-Free Conditions Improves Their Tumor Promoting Functions'. *PLoS ONE,* 7: (8): e42656.

Ribeiro, R.A., Flores, C.A., Cunha, F.Q. and Ferreira, S.H. (1991) 'IL-8 causes in vivo neutrophil migration by a cell-dependent mechanism'. *Immunology,* 73: (4): 472-477.

Rinn, J.L., Bondre, C., Gladstone, H.B., Brown, P.O. and Chang, H.Y. (2006) 'Anatomic Demarcation by Positional Variation in Fibroblast Gene Expression Programs'. *PLoS Genet,* 2: (7): e119.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W. and Smyth, G.K. (2015) 'limma powers differential expression analyses for RNA-sequencing and microarray studies'. *Nucleic Acids Res,* 43: (7): e47.

Ritchie, M.E., Silver, J., Oshlack, A., Holmes, M., Diyagama, D., Holloway, A. and Smyth, G.K. (2007) 'A comparison of background correction methods for two-colour microarrays'. *Bioinformatics,* 23: (20): 2700-2707.

Robinson, M.D., Mccarthy, D.J. and Smyth, G.K. (2010) 'edgeR: a Bioconductor package for differential expression analysis of digital gene expression data'. *Bioinformatics,* 26: (1): 139-140.

Rogacev, K.S., Cremers, B., Zawada, A.M., Seiler, S., Binder, N., Ege, P., Grosse-Dunker, G., Heisel, I., Hornof, F., Jeken, J., Rebling, N.M., Ulrich, C., Scheller, B., Bohm, M., Fliser, D. and Heine, G.H. (2012) 'CD14++CD16+ monocytes independently predict cardiovascular events: a cohort study of 951 patients referred for elective coronary angiography'. *J Am Coll Cardiol,* 60: (16): 1512-1520.

Rollins, B., Walz, A. and Baggiolini, M. (1991) 'Recombinant human MCP-1/JE induces

chemotaxis, calcium flux, and the respiratory burst in human monocytes'. *Blood,* 78: (4): 1112-1116.

Rossol, M., Kraus, S., Pierer, M., Baerwald, C. and Wagner, U. (2012) 'The CD14(bright) CD16+ monocyte subset is expanded in rheumatoid arthritis and promotes expansion of the Th17 cell population'. *Arthritis Rheum,* 64: (3): 671-677.

Rupprecht, T.A., Plate, A., Adam, M., Wick, M., Kastenbauer, S., Schmidt, C., Klein, M., Pfister, H.W. and Koedel, U. (2009) 'The chemokine CXCL13 is a key regulator of B cell recruitment to the cerebrospinal fluid in acute Lyme neuroborreliosis'. *J Neuroinflammation,* 6: 42.

Saalbach, A., Aneregg, U., Bruns, M., Schnabel, E., Herrmann, K. and Haustein, U.F. (1996) 'Novel fibroblast-specific monoclonal antibodies: properties and specificities'. *J Invest Dermatol,* 106: (6): 1314-1319.

Saalbach, A., Kraft, R., Herrmann, K., Haustein, U.F. and Anderegg, U. (1998) 'The monoclonal antibody AS02 recognizes a protein on human fibroblasts being highly homologous to Thy-1'. *Arch Dermatol Res,* 290: (7): 360-366.

Sakaguchi, S. (2004) 'Naturally arising CD4+ regulatory t cells for immunologic self-tolerance and negative control of immune responses'. *Annu Rev Immunol,* 22: 531-562.

Sallusto, F. and Lanzavecchia, A. (1994) 'Efficient presentation of soluble antigen by cultured human dendritic cells is maintained by granulocyte/macrophage colony-stimulating factor plus interleukin 4 and downregulated by tumor necrosis factor alpha'. *J Exp Med,* 179: (4): 1109-1118.

Sallusto, F., Schaerli, P., Loetscher, P., Schaniel, C., Lenig, D., Mackay, C.R., Qin, S. and Lanzavecchia, A. (1998) 'Rapid and coordinated switch in chemokine receptor expression during dendritic cell maturation'. *Eur J Immunol,* 28: (9): 2760-2769.

Santiago-Schwarz, F., Anand, P., Liu, S. and Carsons, S.E. (2001) 'Dendritic cells (DCs) in rheumatoid arthritis (RA): progenitor cells and soluble factors contained in RA synovial fluid yield a subset of myeloid DCs that preferentially activate Th1 inflammatory-type responses'. *J Immunol,* 167: (3): 1758-1768.

Sasai, M., Saeki, Y., Ohshima, S., Nishioka, K., Mima, T., Tanaka, T., Katada, Y., Yoshizaki, K., Suemura, M. and Kishimoto, T. (1999) 'Delayed onset and reduced severity of collagen-induced arthritis in interleukin-6-deficient mice'. *Arthritis Rheum,* 42: (8): 1635-1643.

Sasmono, R.T., Oceandy, D., Pollard, J.W., Tong, W., Pavli, P., Wainwright, B.J., Ostrowski, M.C., Himes, S.R. and Hume, D.A. (2003) 'A macrophage colony-stimulating factor receptor-green fluorescent protein transgene is expressed throughout the mononuclear phagocyte system of the mouse'. *Blood,* 101: (3): 1155-1163.

Sato, M., Mori, Y., Sakurada, A., Fujimura, S. and Horii, A. (1998) 'The H-cadherin (CDH13) gene is inactivated in human lung cancer'. *Hum Genet,* 103: (1): 96-101.

Scaife, S., Brown, R., Kellie, S., Filer, A., Martin, S., Thomas, A.M.C., Bradfield, P.F., Amft, N., Salmon, M. and Buckley, C.D. (2004) 'Detection of differentially expressed genes in synovial fibroblasts by restriction fragment differential display'. *Rheumatology,* 43: (11): 1346-1352.

Schall, T.J., Bacon, K., Toy, K.J. and Goeddel, D.V. (1990) 'Selective attraction of monocytes and

T lymphocytes of the memory phenotype by cytokine RANTES'. *Nature,* 347: (6294): 669-671.

Schellekens, G.A., De Jong, B.A., Van Den Hoogen, F.H., Van De Putte, L.B. and Van Venrooij, W.J. (1998) 'Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies'. *J Clin Invest,* 101: (1): 273-281.

Schuierer, S. and Roma, G. (2016) 'The exon quantification pipeline (EQP): a comprehensive approach to the quantification of gene, exon and junction expression from RNA-seq data'. *Nucleic Acids Res,* 44: (16): e132-e132.

Schulz, C., Gomez Perdiguero, E., Chorro, L., Szabo-Rogers, H., Cagnard, N., Kierdorf, K., Prinz, M., Wu, B., Jacobsen, S.E., Pollard, J.W., Frampton, J., Liu, K.J. and Geissmann, F. (2012) 'A lineage of myeloid cells independent of Myb and hematopoietic stem cells'. *Science,* 336: (6077): 86-90.

Scott, B.B., Weisbrot, L.M., Greenwood, J.D., Bogoch, E.R., Paige, C.J. and Keystone, E.C. (1997) 'Rheumatoid arthritis synovial fibroblast and U937 macrophage/monocyte cell line interaction in cartilage degradation'. *Arthritis Rheum,* 40: (3): 490-498.

Seemayer, C.A., Kuchen, S., Neidhart, M., Kuenzler, P., Rihoskova, V., Neumann, E., Pruschy, M., Aicher, W.K., Muller-Ladner, U., Gay, R.E., Michel, B.A., Firestein, G.S. and Gay, S. (2003) 'p53 in rheumatoid arthritis synovial fibroblasts at sites of invasion'. *Ann Rheum Dis,* 62: (12): 1139-1144.

Sellam, J., Rouanet, S., Hendel-Chavez, H., Miceli-Richard, C., Combe, B., Sibilia, J., Le Loet, X., Tebib, J., Jourdan, R., Dougados, M., Taoufik, Y. and Mariette, X. (2013) 'CCL19, a B cell chemokine, is related to the decrease of blood memory B cells and predicts the clinical response to rituximab in patients with rheumatoid arthritis'. *Arthritis Rheum,* 65: (9): 2253-2261.

Shields, S.B., Oestreich, A.J., Winistorfer, S., Nguyen, D., Payne, J.A., Katzmann, D.J. and Piper, R. (2009) 'ESCRT ubiquitin-binding domains function cooperatively during MVB cargo sorting'. *The Journal of Cell Biology,* 185: (2): 213-224.

Shipman, W.D., Dasoveanu, D.C. and Lu, T.T. (2017) 'Tertiary lymphoid organs in systemic autoimmune diseases: pathogenic or protective?'. *F1000Research,* 6: 196.

Silman, A., Newman, J. and Macgregor, A. (1996) 'Cigarette smoking increases the risk of rheumatoid arthritis. Results from a nationwide study of disease-discordant twins'. *Arthritis Rheum,* 39: 732 - 735.

Simpkins, S.A., Hanby, A.M., Holliday, D.L. and Speirs, V. (2012) 'Clinical and functional significance of loss of caveolin-1 expression in breast cancer-associated fibroblasts'. *J Pathol,* 227: (4): 490-498.

Skoumal, M., Haberhauer, G., Kolarz, G., Hawa, G., Woloszczuk, W. and Klingler, A. (2004) 'Serum cathepsin K levels of patients with longstanding rheumatoid arthritis: correlation with radiological destruction'. *Arthritis Res Ther,* 7: (1): R65.

Skrzeczynska-Moncznik, J., Bzowska, M., Loseke, S., Grage-Griebenow, E., Zembala, M. and Pryjma, J. (2008) 'Peripheral blood CD14high CD16+ monocytes are main producers of IL-10'. *Scand J Immunol,* 67: (2): 152-159.

Sloan, E.K., Ciocca, D.R., Pouliot, N., Natoli, A., Restall, C., Henderson, M.A., Fanelli, M.A., Cuello-

Carrion, F.D., Gago, F.E. and Anderson, R.L. (2009) 'Stromal cell expression of caveolin-1 predicts outcome in breast cancer'. *Am J Pathol,* 174: (6): 2035-2043.

Smeets, T.J.M., Kraan, M.C., Van Loon, M.E. and Tak, P.-P. (2003) 'Tumor necrosis factor α blockade reduces the synovial cell infiltrate early after initiation of treatment, but apparently not by induction of apoptosis in synovial tissue'. *Arthritis & Rheumatism,* 48: (8): 2155-2162.

Smith, M.D. (2011) 'The normal synovium'. *Open Rheumatol J,* 5: 100-106.

Smolen, J.S., Aletaha, D. and Mcinnes, I.B. (2016) 'Rheumatoid arthritis'. *Lancet,* 388: (10055): 2023-2038.

Smolen, J.S., Landewé, R., Bijlsma, J., Burmester, G., Chatzidionysiou, K., Dougados, M., Nam, J., Ramiro, S., Voshaar, M., Van Vollenhoven, R., Aletaha, D., Aringer, M., Boers, M., Buckley, C.D., Buttgereit, F., Bykerk, V., Cardiel, M., Combe, B., Cutolo, M., Van Eijk-Hustings, Y., Emery, P., Finckh, A., Gabay, C., Gomez-Reino, J., Gossec, L., Gottenberg, J.-E., Hazes, J.M.W., Huizinga, T., Jani, M., Karateev, D., Kouloumas, M., Kvien, T., Li, Z., Mariette, X., Mcinnes, I., Mysler, E., Nash, P., Pavelka, K., Poór, G., Richez, C., Van Riel, P., Rubbert-Roth, A., Saag, K., Da Silva, J., Stamm, T., Takeuchi, T., Westhovens, R., De Wit, M. and Van Der Heijde, D. (2017) 'EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update'. *Annals of the Rheumatic Diseases,* 76: (6): 960-977.

Smyth, G.K. and Speed, T. (2003) 'Normalization of cDNA microarray data'. *Methods,* 31: (4): 265-273.

Snapper, C.M., Mcintyre, T.M., Mandler, R., Pecanha, L.M., Finkelman, F.D., Lees, A. and Mond, J.J. (1992) 'Induction of IgG3 secretion by interferon gamma: a model for T cell-independent class switching in response to T cell-independent type 2 antigens'. *J Exp Med,* 175: (5): 1367-1371.

Sohar, N., Hammer, H. and Sohar, I. (2002) 'Lysosomal peptidases and glycosidases in rheumatoid arthritis'. *Biol Chem,* 383: (5): 865-869.

Sotgia, F., Martinez-Outschoorn, U.E., Pavlides, S., Howell, A., Pestell, R.G. and Lisanti, M.P. (2011) 'Understanding the Warburg effect and the prognostic value of stromal caveolin-1 as a marker of a lethal tumor microenvironment'. *Breast Cancer Res,* 13: (4): 213.

Spurlock, C.F., 3rd, Tossberg, J.T., Fuchs, H.A., Olsen, N.J. and Aune, T.M. (2012) 'Methotrexate increases expression of cell cycle checkpoint genes via JNK activation'. *Arthritis Rheum,* 64: (6): 1780-1789.

Srirangan, S. and Choy, E.H. (2010) 'The Role of Interleukin 6 in the Pathophysiology of Rheumatoid Arthritis'. *Therapeutic Advances in Musculoskeletal Disease,* 2: (5): 247-256.

Stanczyk, J., Kowalski, M.L., Grzegorczyk, J., Szkudlinska, B., Jarzebska, M., Marciniak, M. and Synder, M. (2005) 'RANTES and Chemotactic Activity in Synovial Fluids From Patients With Rheumatoid Arthritis and Osteoarthritis'. *Mediators Inflamm,* 2005: (6).

Stanczyk, J., Pedrioli, D.M.L., Brentano, F., Sanchez-Pernaute, O., Kolling, C., Gay, R.E., Detmar, M., Gay, S. and Kyburz, D. (2008) 'Altered expression of MicroRNA in synovial fibroblasts and synovial tissue in rheumatoid arthritis'. *Arthritis & Rheumatism,* 58: (4): 1001-1009.

Stastny, P. (1978) 'Association of the B-Cell Alloantigen DRw4 with Rheumatoid Arthritis'. *New England Journal of Medicine,* 298: (16): 869-871.

Steinhauser, M.L., Kunkel, S.L., Hogaboam, C.M., Evanoff, H., Strieter, R.M. and Lukacs, N.W. (1998) 'Macrophage/fibroblast coculture induces macrophage inflammatory protein-1alpha production mediated by intercellular adhesion molecule-1 and oxygen radicals'. *J Leukoc Biol,* 64: (5): 636-641.

Sternberg, Z., Ghanim, H., Gillotti, K.M., Tario Jr, J.D., Munschauer, F., Curl, R., Noor, S., Yu, J., Ambrus Sr, J.L., Wallace, P. and Dandona, P. (2013) 'Flow cytometry and gene expression profiling of immune cells of the carotid plaque and peripheral blood'. *Atherosclerosis,* 229: (2): 338-347.

Strand, V., Kimberly, R. and Isaacs, J.D. (2007) 'Biologic therapies in rheumatology: lessons learned, future directions'. *Nat Rev Drug Discov,* 6: (1): 75-92.

Strauss-Ayali, D., Conrad, S.M. and Mosser, D.M. (2007) 'Monocyte subpopulations and their differentiation patterns during infection'. *J Leukoc Biol,* 82: (2): 244-252.

Stuchell, M.D., Garrus, J.E., Muller, B., Stray, K.M., Ghaffarian, S., Mckinnon, R., Krausslich, H.G., Morham, S.G. and Sundquist, W.I. (2004) 'The human endosomal sorting complex required for transport (ESCRT-I) and its role in HIV-1 budding'. *J Biol Chem,* 279: (34): 36059-36071.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. and Mesirov, J.P. (2005) 'Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles'. *Proceedings of the National Academy of Sciences,* 102: (43): 15545-15550.

Sumagin, R., Prizant, H., Lomakina, E., Waugh, R.E. and Sarelius, I.H. (2010) 'LFA-1 and Mac-1 Define Characteristically Different Intralumenal Crawling and Emigration Patterns for Monocytes and Neutrophils In Situ'. *The Journal of Immunology,* 185: (11): 7057-7066.

Sun, H., Xia, Y., Wang, L., Wang, Y. and Chang, X. (2013) 'PSORS1C1 may be involved in rheumatoid arthritis'. *Immunol Lett,* 153: (1-2): 9-14.

Sur Chowdhury, C., Giaglis, S., Walker, U.A., Buser, A., Hahn, S. and Hasler, P. (2014) 'Enhanced neutrophil extracellular trap generation in rheumatoid arthritis: analysis of underlying signal transduction pathways and potential diagnostic utility'. *Arthritis Res Ther,* 16: (3): R122.

Suzuki, A., Yamada, R., Chang, X., Tokuhiro, S., Sawada, T., Suzuki, M., Nagasaki, M., Nakayama-Hamada, M., Kawaida, R., Ono, M., Ohtsuki, M., Furukawa, H., Yoshino, S., Yukioka, M., Tohma, S., Matsubara, T., Wakitani, S., Teshima, R., Nishioka, Y., Sekine, A., Iida, A., Takahashi, A., Tsunoda, T., Nakamura, Y. and Yamamoto, K. (2003) 'Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis'. *Nat Genet,* 34: (4): 395-402.

Symmons, D., Turner, G., Webb, R., Asten, P., Barrett, E., Lunt, M., Scott, D. and Silman, A. (2002) 'The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century'. *Rheumatology,* 41: (7): 793-800.

Tak, P.P., Smeets, T.J., Daha, M.R., Kluin, P.M., Meijers, K.A., Brand, R., Meinders, A.E. and

Breedveld, F.C. (1997) 'Analysis of the synovial cell infiltrate in early rheumatoid synovial tissue in relation to local disease activity'. *Arthritis Rheum,* 40: (2): 217-225.

Takeuchi, O. and Akira, S. (2010) 'Pattern recognition receptors and inflammation'. *Cell,* 140: (6): 805-820.

Takeuchi, T., Liang, S.B., Matsuyoshi, N., Zhou, S., Miyachi, Y., Sonobe, H. and Ohtsuki, Y. (2002) 'Loss of T-cadherin (CDH13, H-cadherin) expression in cutaneous squamous cell carcinoma'. *Lab Invest,* 82: (8): 1023-1029.

Takeuchi, T., Misaki, A., Liang, S.B., Tachibana, A., Hayashi, N., Sonobe, H. and Ohtsuki, Y. (2000) 'Expression of T-cadherin (CDH13, H-Cadherin) in human brain and its characteristics as a negative growth regulator of epidermal growth factor in neuroblastoma cells'. *J Neurochem,* 74: (4): 1489-1497.

Tavian, M. and Peault, B. (2005) 'Embryonic development of the human hematopoietic system'. *Int J Dev Biol,* 49: (2-3): 243-250.

Taylor, L.H., Twigg, S., Worthington, J., Emery, P., Morgan, A.W., Wilson, A.G. and Teare, M.D. (2013) 'Metaanalysis of the Association of Smoking and PTPN22 R620W Genotype on Autoantibody Status and Radiological Erosions in Rheumatoid Arthritis'. *The Journal of Rheumatology,* 40: (7): 1048-1053.

Thompson, O., Moghraby, J.S., Ayscough, K.R. and Winder, S.J. (2012) 'Depletion of the actin bundling protein SM22/transgelin increases actin dynamics and enhances the tumourigenic phenotypes of cells'. *BMC Cell Biology,* 13: 1-1.

Topalian, S.L., Hodi, F.S., Brahmer, J.R., Gettinger, S.N., Smith, D.C., Mcdermott, D.F., Powderly, J.D., Carvajal, R.D., Sosman, J.A., Atkins, M.B., Leming, P.D., Spigel, D.R., Antonia, S.J., Horn, L., Drake, C.G., Pardoll, D.M., Chen, L., Sharfman, W.H., Anders, R.A., Taube, J.M., Mcmiller, T.L., Xu, H., Korman, A.J., Jure-Kunkel, M., Agrawal, S., Mcdonald, D., Kollia, G.D., Gupta, A., Wigginton, J.M. and Sznol, M. (2012) 'Safety, Activity, and Immune Correlates of Anti–PD-1 Antibody in Cancer'. *New England Journal of Medicine,* 366: (26): 2443-2454.

Totaro, M.C., Cattani, P., Ria, F., Tolusso, B., Gremese, E., Fedele, A.L., D'onghia, S., Marchetti, S., Sante, G.D., Canestri, S. and Ferraccioli, G. (2013) 'Porphyromonas gingivalis and the pathogenesis of rheumatoid arthritis: analysis of various compartments including the synovial tissue'. *Arthritis Res Ther,* 15: (3): R66.

Trabandt, A., Aicher, W.K., Gay, R.E., Sukhatme, V.P., Fassbender, H.G. and Gay, S. (1992) 'Spontaneous expression of immediately-early response genes c-fos and egr-1 in collagenase-producing rheumatoid synovial fibroblasts'. *Rheumatol Int,* 12: (2): 53-59.

Triebel, F., Jitsukawa, S., Baixeras, E., Roman-Roman, S., Genevee, C., Viegas-Pequignot, E. and Hercend, T. (1990) 'LAG-3, a novel lymphocyte activation gene closely related to CD4'. *J Exp Med,* 171: (5): 1393-1405.

Tsark, E.C., Wang, W., Teng, Y.C., Arkfeld, D., Dodge, G.R. and Kovats, S. (2002) 'Differential MHC class II-mediated presentation of rheumatoid arthritis autoantigens by human dendritic cells and macrophages'. *J Immunol,* 169: (11): 6625-6633.

Tsou, P.S., Ruth, J.H., Campbell, P.L., Isozaki, T., Lee, S., Marotte, H., Domino, S.E., Koch, A.E. and

Amin, M.A. (2013) 'A novel role for inducible Fut2 in angiogenesis'. *Angiogenesis,* 16: (1): 195-205.

Turman, M.A., Yabe, T., Mcsherry, C., Bach, F.H. and Houchins, J.P. (1993) 'Characterization of a novel gene (NKG7) on human chromosome 19 that is expressed in natural killer cells and T cells'. *Human Immunology,* 36: (1): 34-40.

Turner, J.D. and Filer, A. (2015) 'The role of the synovial fibroblast in rheumatoid arthritis pathogenesis'. *Current Opinion in Rheumatology,* 27: (2): 175-182.

Uccelli, A., Moretta, L. and Pistoia, V. (2008) 'Mesenchymal stem cells in health and disease'. *Nat Rev Immunol,* 8: (9): 726-736.

Umino, T., Skold, C.M., Pirruccello, S.J., Spurzem, J.R. and Rennard, S.I. (1999) 'Two-colour flow-cytometric analysis of pulmonary alveolar macrophages from smokers'. *Eur Respir J,* 13: (4): 894-899.

Unemori, E.N., Hibbs, M.S. and Amento, E.P. (1991) 'Constitutive expression of a 92-kD gelatinase (type V collagenase) by rheumatoid synovial fibroblasts and its induction in normal human fibroblasts by inflammatory cytokines'. *J Clin Invest,* 88: (5): 1656-1662.

Valencia, X., Higgins, J.M., Kiener, H.P., Lee, D.M., Podrebarac, T.A., Dascher, C.C., Watts, G.F., Mizoguchi, E., Simmons, B., Patel, D.D., Bhan, A.K. and Brenner, M.B. (2004) 'Cadherin-11 provides specific cellular adhesion between fibroblast-like synoviocytes'. *J Exp Med,* 200: (12): 1673-1679.

Van Aken, J., Lard, L.R., Le Cessie, S., Hazes, J.M.W., Breedveld, F.C. and Huizinga, T.W.J. (2004) 'Radiological outcome after four years of early versus delayed treatment strategy in patients with recent onset rheumatoid arthritis'. *Annals of the Rheumatic Diseases,* 63: (3): 274-279.

Van Den Berg, A., Freitas, J., Keles, F., Snoek, M., Van Marle, J., Jansen, H.M. and Lutter, R. (2006) 'Cytoskeletal architecture differentially controls post-transcriptional processing of IL-6 and IL-8 mRNA in airway epithelial-like cells'. *Exp Cell Res,* 312: (9): 1496-1506.

Van Der Heijden, J.W., Oerlemans, R., Dijkmans, B.a.C., Qi, H., Laken, C.J.V.D., Lems, W.F., Jackman, A.L., Kraan, M.C., Tak, P.P., Ratnam, M. and Jansen, G. (2009) 'Folate receptor β as a potential delivery route for novel folate antagonists to macrophages in the synovial tissue of rheumatoid arthritis patients'. *Arthritis & Rheumatism,* 60: (1): 12-21.

Van Der Helm-Van Mil, A.H. and Huizinga, T.W. (2008) 'Advances in the genetics of rheumatoid arthritis point to subclassification into distinct disease subsets'. *Arthritis Research & Therapy,* 10: (2): 205.

Van Der Helm-Van Mil, A.H., Verpoort, K.N., Breedveld, F.C., Huizinga, T.W., Toes, R.E. and De Vries, R.R. (2006) 'The HLA-DRB1 shared epitope alleles are primarily a risk factor for anti-cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis'. *Arthritis Rheum,* 54: (4): 1117-1121.

Van Der Maaten, L.J.P. and Hinton, G.E. (2008) 'Visualizing High-Dimensional Data Using t-SNE'. *Journal of Machine Learning Research,* 9: (Nov): 2579-2605.

Van Herwijnen, M.J.C., Wieten, L., Van Der Zee, R., Van Kooten, P.J., Wagenaar-Hilbers, J.P.,

Hoek, A., Den Braber, I., Anderton, S.M., Singh, M., Meiring, H.D., Van Els, C.a.C.M., Van Eden, W. and Broere, F. (2012) 'Regulatory T cells that recognize a ubiquitous stress-inducible self-antigen are long-lived suppressors of autoimmune arthritis'. *Proceedings of the National Academy of Sciences,* 109: (35): 14134-14139.

Van Landuyt, K.B., Jones, E.A., Mcgonagle, D., Luyten, F.P. and Lories, R.J. (2010) 'Flow cytometric characterization of freshly isolated and culture expanded human synovial cell populations in patients with chronic arthritis'. *Arthritis Res Ther,* 12: (1): R15.

Van Oene, M., Wintle, R.F., Liu, X., Yazdanpanah, M., Gu, X., Newman, B., Kwan, A., Johnson, B., Owen, J., Greer, W., Mosher, D., Maksymowych, W., Keystone, E., Rubin, L.A., Amos, C.I. and Siminovitch, K.A. (2005) 'Association of the lymphoid tyrosine phosphatase R620W variant with rheumatoid arthritis, but not Crohn's disease, in Canadian populations'. *Arthritis & Rheumatism,* 52: (7): 1993-1998.

Van Oosterhout, M., Bajema, I., Levarht, E.W.N., Toes, R.E.M., Huizinga, T.W.J. and Van Laar, J.M. (2008) 'Differences in synovial tissue infiltrates between anti–cyclic citrullinated peptide–positive rheumatoid arthritis and anti–cyclic citrullinated peptide–negative rheumatoid arthritis'. *Arthritis & Rheumatism,* 58: (1): 53-60.

Vang, T., Liu, W.H., Delacroix, L., Wu, S., Vasile, S., Dahl, R., Yang, L., Musumeci, L., Francis, D., Landskron, J., Tasken, K., Tremblay, M.L., Lie, B.A., Page, R., Mustelin, T., Rahmouni, S., Rickert, R.C. and Tautz, L. (2012) 'LYP inhibits T-cell activation when dissociated from CSK'. *Nat Chem Biol,* 8: (5): 437-446.

Veldhoen, M., Hocking, R.J., Atkins, C.J., Locksley, R.M. and Stockinger, B. (2006) 'TGFbeta in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells'. *Immunity,* 24: (2): 179-189.

Verreck, F.A., De Boer, T., Langenberg, D.M., Hoeve, M.A., Kramer, M., Vaisberg, E., Kastelein, R., Kolk, A., De Waal-Malefyt, R. and Ottenhoff, T.H. (2004) 'Human IL-23-producing type 1 macrophages promote but IL-10-producing type 2 macrophages subvert immunity to (myco)bacteria'. *Proc Natl Acad Sci U S A,* 101: (13): 4560-4565.

Vestweber, D. (2008) 'VE-Cadherin'. *The Major Endothelial Adhesion Molecule Controlling Cellular Junctions and Blood Vessel Formation,* 28: (2): 223-232.

Villani, A.-C., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., Griesbeck, M., Butler, A., Zheng, S., Lazo, S., Jardine, L., Dixon, D., Stephenson, E., Nilsson, E., Grundberg, I., Mcdonald, D., Filby, A., Li, W., De Jager, P.L., Rozenblatt-Rosen, O., Lane, A.A., Haniffa, M., Regev, A. and Hacohen, N. (2017) 'Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors'. *Science,* 356: (6335).

Vojinovic, J., Damjanov, N., D'urzo, C., Furlan, A., Susic, G., Pasic, S., Iagaru, N., Stefan, M. and Dinarello, C.A. (2011) 'Safety and efficacy of an oral histone deacetylase inhibitor in systemic-onset juvenile idiopathic arthritis'. *Arthritis & Rheumatism,* 63: (5): 1452-1458.

Vuorio, E., Einola, S., Hakkarainen, S. and Penttinen, R. (1982) 'Synthesis of underpolymerized hyaluronic acid by fibroblasts cultured from rheumatoid and non-rheumatoid synovitis'. *Rheumatol Int,* 2: (3): 97-102.

Wagner, U.G., Kurtin, P.J., Wahner, A., Brackertz, M., Berry, D.J., Goronzy, J.J. and Weyand, C.M.

(1998) 'The Role of CD8+ CD40L+ T Cells in the Formation of Germinal Centers in Rheumatoid Synovitis'. *The Journal of Immunology,* 161: (11): 6390-6397.

Wahl, S.M., Allen, J.B., Welch, G.R. and Wong, H.L. (1992) 'Transforming growth factor-beta in synovial fluids modulates Fc gamma RII (CD16) expression on mononuclear phagocytes'. *The Journal of Immunology,* 148: (2): 485-490.

Wang, Y., Kristan, J., Hao, L., Lenkoski, C.S., Shen, Y. and Matis, L.A. (2000) 'A Role for Complement in Antibody-Mediated Inflammation: C5-Deficient DBA/1 Mice Are Resistant to Collagen-Induced Arthritis'. *The Journal of Immunology,* 164: (8): 4340-4347.

Wang, Y., Rollins, S.A., Madri, J.A. and Matis, L.A. (1995) 'Anti-C5 monoclonal antibody therapy prevents collagen-induced arthritis and ameliorates established disease'. *Proceedings of the National Academy of Sciences,* 92: (19): 8955-8959.

Wang, Z., Gerstein, M. and Snyder, M. (2009) 'RNA-Seq: a revolutionary tool for transcriptomics'. *Nature Reviews Genetics,* 10: 57.

Wang, Z., Yin, N., Zhang, Z., Zhang, Y., Zhang, G. and Chen, W. (2017) 'Upregulation of T-cell Immunoglobulin and Mucin-Domain Containing-3 (Tim-3) in Monocytes/Macrophages Associates with Gastric Cancer Progression'. *Immunol Invest,* 46: (2): 134-148.

Weber, J.S., Dummer, R., De Pril, V., Lebbe, C. and Hodi, F.S. (2013) 'Patterns of onset and resolution of immune-related adverse events of special interest with ipilimumab: detailed safety analysis from a phase 3 trial in patients with advanced melanoma'. *Cancer,* 119: (9): 1675-1682.

Weiss, B., Davidkova, G. and Zhou, L.W. (1999) 'Antisense RNA gene therapy for studying and modulating biological processes'. *Cell Mol Life Sci,* 55: (3): 334-358.

Whitaker, J., Shoemaker, R., Boyle, D., Hillman, J., Anderson, D., Wang, W. and Firestein, G. (2013) 'An imprinted rheumatoid arthritis methylome signature reflects pathogenic phenotype'. *Genome Medicine,* 5: (4): 40.

Witkiewicz, A.K., Dasgupta, A., Sotgia, F., Mercier, I., Pestell, R.G., Sabel, M., Kleer, C.G., Brody, J.R. and Lisanti, M.P. (2009) 'An absence of stromal caveolin-1 expression predicts early tumor recurrence and poor clinical outcome in human breast cancers'. *Am J Pathol,* 174: (6): 2023-2034.

Wolfe, F. and Sharp, J.T. (1998) 'Radiographic outcome of recent-onset rheumatoid arthritis: A 19-year study of radiographic progression'. *Arthritis & Rheumatism,* 41: (9): 1571-1582.

Wong, K.L., Tai, J.J., Wong, W.C., Han, H., Sem, X., Yeap, W.H., Kourilsky, P. and Wong, S.C. (2011) 'Gene expression profiling reveals the defining features of the classical, intermediate, and nonclassical human monocyte subsets'. *Blood,* 118: (5): e16-31.

Woolhiser, M.R., Brockow, K. and Metcalfe, D.D. (2004) 'Activation of human mast cells by aggregated IgG through FcgammaRI: additive effects of C3a'. *Clin Immunol,* 110: (2): 172-180.

Workman, C.J., Wang, Y., El Kasmi, K.C., Pardoll, D.M., Murray, P.J., Drake, C.G. and Vignali, D.A. (2009) 'LAG-3 regulates plasmacytoid dendritic cell homeostasis'. *J Immunol,* 182: (4): 1885-1891.

Wu, S.C. and Zhang, Y. (2010) 'Active DNA demethylation: many roads lead to Rome'. *Nat Rev Mol Cell Biol,* 11: (9): 607-620.

Xavier, J.M., Shahram, F., Sousa, I., Davatchi, F., Matos, M., Abdollahi, B.S., Sobral, J., Nadji, A., Oliveira, M., Ghaderibarim, F., Shafiee, N.M. and Oliveira, S.A. (2015) 'FUT2: filling the gap between genes and environment in Behçet's disease?'. *Annals of the Rheumatic Diseases,* 74: (3): 618-624.

Xia, W., Hilgenbrink, A.R., Matteson, E.L., Lockwood, M.B., Cheng, J.X. and Low, P.S. (2009) 'A functional folate receptor is induced during macrophage activation and can be used to target drugs to activated macrophages'. *Blood,* 113: (2): 438-446.

Xiong, Y.S., Cheng, Y., Lin, Q.S., Wu, A.L., Yu, J., Li, C., Sun, Y., Zhong, R.Q. and Wu, L.J. (2014) 'Increased expression of Siglec-1 on peripheral blood monocytes and its role in mononuclear cell reactivity to autoantigen in rheumatoid arthritis'. *Rheumatology (Oxford),* 53: (2): 250-259.

Xue, C., Takahashi, M., Hasunuma, T., Aono, H., Yamamoto, K., Yoshino, S., Sumida, T. and Nishioka, K. (1997) 'Characterisation of fibroblast-like cells in pannus lesions of patients with rheumatoid arthritis sharing properties of fibroblasts and chondrocytes'. *Ann Rheum Dis,* 56: (4): 262-267.

Xue, J., Schmidt, S.V., Sander, J., Draffehn, A., Krebs, W., Quester, I., De Nardo, D., Gohel, T.D., Emde, M., Schmidleithner, L., Ganesan, H., Nino-Castro, A., Mallmann, M.R., Labzin, L., Theis, H., Kraut, M., Beyer, M., Latz, E., Freeman, T.C., Ulas, T. and Schultze, J.L. (2014a) 'Transcriptome-based network analysis reveals a spectrum model of human macrophage activation'. *Immunity,* 40: (2): 274-288.

Xue, M., Mckelvey, K., Shen, K., Minhas, N., March, L., Park, S.Y. and Jackson, C.J. (2014b) 'Endogenous MMP-9 and not MMP-2 promotes rheumatoid synovial fibroblast survival, inflammation and cartilage degradation'. *Rheumatology (Oxford),* 53: (12): 2270-2279.

Yamashita, A., Yonemitsu, Y., Okano, S., Nakagawa, K., Nakashima, Y., Irisa, T., Iwamoto, Y., Nagai, Y., Hasegawa, M. and Sueishi, K. (2002) 'Fibroblast Growth Factor-2 Determines Severity of Joint Disease in Adjuvant-Induced Arthritis in Rats'. *The Journal of Immunology,* 168: (1): 450-457.

Yang, C.-M., Luo, S.-F., Hsieh, H.-L., Chi, P.-L., Lin, C.-C., Wu, C.-C. and Hsiao, L.-D. (2010) 'Interleukin-1β induces ICAM-1 expression enhancing leukocyte adhesion in human rheumatoid arthritis synovial fibroblasts: Involvement of ERK, JNK, AP-1, and NF-κB'. *Journal of Cellular Physiology,* 224: (2): 516-526.

Yang, J.G., Deng, Y., Zhou, L.X., Li, X.Y., Sun, P.R. and Sun, N.X. (2013) 'Overexpression of CDKN1B inhibits fibroblast proliferation in a rabbit model of experimental glaucoma filtration surgery'. *Invest Ophthalmol Vis Sci,* 54: (1): 343-352.

Yang, M.-H., Wu, F.-X., Xie, C.-M., Qing, Y.-F., Wang, G.-R., Guo, X.-L., Tang, Z., Zhou, J.-G. and Yuan, G.-H. (2009) 'Expression of CC Chemokine Ligand 5 in Patients with Rheumatoid Arthritis and Its Correlation with Disease Activity and Medication'. *Chinese Medical Sciences Journal,* 24: (1): 50-54.

Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J. and Speed, T.P. (2002) 'Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide

systematic variation'. *Nucleic Acids Res,* 30: (4): e15.

Yeo, L., Adlard, N., Biehl, M., Juarez, M., Smallie, T., Snow, M., Buckley, C.D., Raza, K., Filer, A. and Scheel-Toellner, D. (2015a) 'Expression of chemokines CXCL4 and CXCL7 by synovial macrophages defines an early stage of rheumatoid arthritis'. *Annals of the Rheumatic Diseases*.

Yeo, L., Lom, H., Juarez, M., Snow, M., Buckley, C.D., Filer, A., Raza, K. and Scheel-Toellner, D. (2015b) 'Expression of FcRL4 defines a pro-inflammatory, RANKL-producing B cell subset in rheumatoid arthritis'. *Annals of the Rheumatic Diseases,* 74: (5): 928-935.

Yeo, L., Toellner, K.-M., Salmon, M., Filer, A., Buckley, C.D., Raza, K. and Scheel-Toellner, D. (2011) 'Cytokine mRNA profiling identifies B cells as a major source of RANKL in rheumatoid arthritis'. *Annals of the Rheumatic Diseases,* 70: (11): 2022-2028.

Yi, W., Zhang, P., Liang, Y., Zhou, Y., Shen, H., Fan, C., Moorman, J.P., Yao, Z.Q., Jia, Z. and Zhang, Y. (2017) 'T-bet-mediated Tim-3 expression dampens monocyte function during chronic hepatitis C virus infection'. *Immunology,* 150: (3): 301-311.

Young, C.L., Adamson, T.C., Vaughan, J.H. and Fox, R.I. (1984) 'Immunohistologic characterization of synovial membrane lymphocytes in rheumatoid arthritis'. *Arthritis & Rheumatism,* 27: (1): 32-39.

Young, M.D., Wakefield, M.J., Smyth, G.K. and Oshlack, A. (2010) 'Gene ontology analysis for RNA-seq: accounting for selection bias'. *Genome Biology,* 11: (2): R14.

Zhang, D.H., Yang, Z.L., Zhou, E.X., Miao, X.Y., Zou, Q., Li, J.H., Liang, L.F., Zeng, G.X. and Chen, S.L. (2016) 'Overexpression of Thy1 and ITGA6 is associated with invasion, metastasis and poor prognosis in human gallbladder carcinoma'. *Oncol Lett,* 12: (6): 5136-5144.

Zhang, L., Yu, M., Deng, J., Lv, X., Liu, J., Xiao, Y., Yang, W., Zhang, Y. and Li, C. (2015) 'Chemokine Signaling Pathway Involved in CCL2 Expression in Patients with Rheumatoid Arthritis'. *Yonsei Med J,* 56: (4): 1134-1142.

Zhao, J., Ouyang, Q., Hu, Z., Huang, Q., Wu, J., Wang, R. and Yang, M. (2016) 'A protocol for the culture and isolation of murine synovial fibroblasts'. *Biomedical Reports,* 5: (2): 171-175.

Zhou, L., Lee, D.H., Plescia, J., Lau, C.Y. and Altieri, D.C. (1994) 'Differential ligand binding specificities of recombinant CD11b/CD18 integrin I-domain'. *J Biol Chem,* 269: (25): 17075-17079.

Zhou, M., Qin, S., Chu, Y., Wang, F., Chen, L. and Lu, Y. (2014) 'Immunolocalization of MMP-2 and MMP-9 in human rheumatoid synovium'. *Int J Clin Exp Pathol,* 7: (6): 3048-3056.

Zickus, C., Kunkel, S.L., Simpson, K., Evanoff, H., Glass, M., Strieter, R.M. and Lukacs, N.W. (1998) 'Differential regulation of C-C chemokines during fibroblast-monocyte interactions: adhesion vs. inflammatory cytokine pathways'. *Mediators Inflamm,* 7: (4): 269-274.

Ziegler-Heitbrock, L., Ancuta, P., Crowe, S., Dalod, M., Grau, V., Hart, D.N., Leenen, P.J., Liu, Y.J., Macpherson, G., Randolph, G.J., Scherberich, J., Schmitz, J., Shortman, K., Sozzani, S., Strobl, H., Zembala, M., Austyn, J.M. and Lutz, M.B. (2010) 'Nomenclature of monocytes and dendritic cells in blood'. *Blood,* 116: (16): e74-80.

Ziegler-Heitbrock, L. and Hofer, T.P. (2013) 'Toward a refined definition of monocyte subsets'. *Front Immunol,* 4: 23.

Zimmermann, T., Kunisch, E., Pfeiffer, R., Hirth, A., Stahl, H.-D., Sack, U., Laube, A., Liesaus, E., Roth, A., Palombo-Kinne, E., Emmrich, F. and Kinne, R. (2000) 'Isolation and characterization of rheumatoid arthritis synovial fibroblasts from primary culture — primary culture cells markedly differ from fourth-passage cells'. *Arthritis Res,* 3: (1): 1-20.

Zrioual, S., Toh, M.-L., Tournadre, A., Zhou, Y., Cazalis, M.-A., Pachot, A., Miossec, V. and Miossec, P. (2008) 'IL-17RA and IL-17RC Receptors Are Essential for IL-17A-Induced ELR+ CXC Chemokine Expression in Synoviocytes and Are Overexpressed in Rheumatoid Blood'. *The Journal of Immunology,* 180: (1): 655-663.

Zvaifler, N.J., Steinman, R.M., Kaplan, G., Lau, L.L. and Rivelis, M. (1985) 'Identification of immunostimulatory dendritic cells in the synovial effusions of patients with rheumatoid arthritis'. *J Clin Invest,* 76: (2): 789-800.